6th BSC Severo Ochoa Doctoral Symposium

7th, 8th and 9th May, 2019

Book of Abstracts

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Book of Abstracts
6th BSC Severo Ochoa Doctoral Symposium

Editors
María José García Miraz
Carolina Olmo

Cover
Design based on artwork created by macrovector.com

This book is available at
https://www.bsc.es/education/predoctoral-phd/doctoral-symposium/6th-bsc-so-doctoral-symposium

published by
Barcelona Supercomputing Center

6th Edition, April 2019

ACKNOWLEDGEMENTS

EDITORIAL COMMENT

We are proud to present the Book of Abstracts for the 6th BSC Severo Ochoa Doctoral Symposium.

During more than ten years, the Barcelona Supercomputing Center has been receiving undergraduate, master and PhD students, and providing them training and skills to develop a successful career. Many of those students are now researchers and experts at BSC and in other international research institutions.

In fact, the number of students has never decreased. On the contrary, their number and research areas have grown and we noticed that these highly qualified students, especially the PhD candidates, needed a forum to present their findings and fruitfully exchange ideas. As a result, in 2014, the first BSC Doctoral Symposium was born.

In this 6th edition of the BSC Severo Ochoa Doctoral Symposium we are offering two keynote talks titled "Brain and Behaviour: to which extent are we responsible for who we are" and "Preparing for Extreme Heterogeneity in HPC" and one two session tutorial on the topic of Creativity and Innovation.

This year we have made a special effort to organize the presentations into interdisciplinary blocks. The presentations will be held in 6 different sessions and will tackle the following topics:
- Algorithms, plastic biodegradation, wave propagation and seismology
- Data Analytics, genetic variability, distributed computing and seismology
- Machine Learning, personalized medicine and air quality
- HPC, biological simulation and computer architecture
- Meshes and Fluids, Hemodynamics
- Models and simulation, dust and pathogenicity prediction

The posters will be exhibited and presented during three poster sessions that will give the authors the opportunity to explain their research and results.

WELCOME ADDRESS

I am delighted to welcome all the PhD students, Postdoc researchers, advisors, experts and attendees participating in the 6th BSC Severo Ochoa Doctoral Symposium.

The goal of the event continues to be providing a framework to share research results of the projects developed by PhD thesis that use High Performance Computing in some degree.

The symposium was conceived in the framework of the Severo Ochoa Program at BSC, following the project aims regarding the talent development and knowledge sharing. Keeping that in mind, the symposium provides an interactive forum for PhD students considering both the ones just beginning their research and others who have developed their research activities during several years.

As a consequence, I highly appreciate the support provided by BSC and the Severo Ochoa Center of Excellence Programme that make possible to celebrate this event.

I must add that I am very grateful to the BSC directors for supporting the symposium, to the group leaders and to the advisors for encouraging the participation of the students in the event. Moreover, I wish to specially thank the keynote speakers David Bueno i Torrens and Jeffrey Vetter and the for their willingness to share with us their knowledge and expertise.

And last but not least, I would like to thank all PhD students and Postdoc researchers for their presentations and effort. I wish you all the best for your career and I really hope you enjoy this great opportunity to meet other colleagues and share your experiences.

Dr. Maria Ribera Sancho
Manager of BSC Education & Training

KEYNOTE SPEAKER 1

David Bueno i Torrens

Doctor in biology and professor of genetics at the University of Barcelona.

Brain and behaviour: to which extent are we responsible for who we are?

The brain is the organ of thought. Its neural networks manage all our behaviours. Its ontogenetic origin, however, is dual. On one hand, its formation and its functioning are conditioned by a number of genes, which make each person more or less prone to any cognitive ability as well as for any behavioural response. On the other hand, the environment also influences how the synaptic conexions are established, which in turn sets the neural networks, whose activity will generate all the behaviours and learning. What is the relative influence of each of these factors? Can we contribute to the construction of our brain? In this keynote we will discuss to which extent we are responsible for being who we are.

David Bueno i Torrens (Barcelona, 1965) is a doctor in biology and professor of genetics at the University of Barcelona. His professional and academic career has been developed in Barcelona and Oxford, focusing on the genetics of development and neuroscience, and its relationship with human behavior. He teaches several subjects in the field of genetics and has published more than fifty scientific articles in specialized journals. In the field of scientific dissemination, he has published seven books to bring science closer to the public, as well as several textbooks.

KEYNOTE SPEAKER 2

Jeffrey Vetter

Jeffrey S. Vetter, Ph.D., is a Distinguished R&D Staff Member, and the founding group leader of the Future Technologies Group in the Computer Science and Mathematics Division of Oak Ridge National Laboratory (Tennessee, USA)

Preparing for Extreme Heterogeneity in High Performance Computing

While computing technologies have remained relatively stable for nearly two decades, new architectural features, such as heterogeneous cores, deep memory hierarchies, non-volatile memory (NVM), and near-memory processing, have emerged as possible solutions to address the concerns of energy-efficiency and cost. However, we expect this 'golden age' of architectural change to lead to extreme heterogeneity and it will have a major impact on software systems and applications. Software will need to be redesigned to exploit these new capabilities and provide some level of performance portability across these diverse architectures. In this talk, I will sample these emerging memory technologies, discuss their architectural and software implications, and describe several new approaches (e.g., domain specific languages, intelligent compilers and introspective runtime systems) to address these challenges.

Jeffrey Vetter, Ph.D., is a Distinguished Research and Development Staff Member as well as the founding group leader of the Future Technologies Group in the Computer Science and Mathematics Division at Oak Ridge National Laboratory (ORNL). In addition, he holds a joint appointment at the Electrical Engineering and Computer Science Department of the University of Tennessee-Knoxville. Vetter is a Fellow of the IEEE and a Distinguished Scientist Member of the ACM. In 2018, he was awarded the ORNL Director's Award for Outstanding Individual Accomplishment in Science and Technology. In 2010, as part of an interdisciplinary team from Georgia Tech, NYU, and ORNL, Vetter was awarded the ACM Gordon Bell Prize. In 2015, he served as the SC15 Technical Program Chair (the annual supercomputing conference with nearly 13,000 attendees). His recent books, entitled "Contemporary High Performance Computing: From Petascale toward Exascale (Vols. 1-3)," survey the international landscape of HPC. See his website for more information: https://ft.ornl.gov/~vetter/.

TUTORIALS

Creativity & Innovation: Why creativity is needed in a science career?

Goal: The goal of this workshop is to set the scene to value creativity and innovation in a scientific environment.

Contents:

- Creative Thinking:
    - Creativity vs. Innovation
    - Uses in the scientific environment
- Requirements and strategies to foster creativity and innovation
    - Barriers and blockages
- Practices and resources to develop creativity and innovation
- Collecting challenges to solve them with the help of creativity techniques

Creativity & Innovation: Working on real challenges with Design Thinking techniques

Goal: The goal of this workshop is to provide scientists with practical tools and resources that can be applied in their working environment to bring innovation to life.

Contents:

- Presenting challenges to be solved in teams
- The process of creative problem solving
- Introducing and practicing with applied creativity techniques:
    - Design Thinking (IDEO)
    - Design Sprint (GOOGLE)
- Participants present their creative solutions
- Individual Action Plan to transfer learning to the workplace

Alicia Marín Muniesa is a consultant, coach and trainer with 18 years of experience in multinational companies. She provides talent development and coaching for managers at international organizations, with a focus on Leadership Skills, Communication, Creativity Techniques, Cross Cultural Effectiveness, Self Management and Team Development. She has extensive experience in training members of the scientific and medical community in communication-related subjects, and has developed and run several workshops and leadership programs for women in science and education. She is also an Associate Professor lecturing on Cultural Intelligence at IE-Business School, Madrid, and Executive skills at The Ostelea School of Tourism and Hospitality, Barcelona.

Isabel Nogueroles is a professional specialized in Leadership and Organisational Change with wide experience in both international and national companies. She has worked extensively as a Change Management Consultant and as a Leadership Coach. She contributes to leadership performance by coaching Executives and Managers, Senior Management Committees, Work Groups, Start-ups, Organisations and young University Students. Accustomed to working with all hierarchical levels, she contributes to the clarification of vision, beliefs and values and develops capacities to increase leadership, relationship competence and influence to enable leaders to become the catalysts for change within their organizations. She is a professional who stands out for being strategic, creative, empathetic, positive, flexible and committed.

Judit Murlans is a HR consultant, trainer and coach. She has a Degree in Clinical and Organizational Psychology and several certifications as coach and trainer MBSR – Mindfulness. Her professional career includes 30 years of experience working in Human Resources, Advertising & Marketing. The last 15 years as a freelance. She assumed positions of Human Resources Management in several companies in Spain and Cuba. She founded and managed for over 10 years Velvet Mediendesign, an Advertising and Design Company located in Munich. She founded and led Little Che, an audio-visual production company, located in Barcelona. Her experience as a trainer is focused on Leadership Skills, Creativity and Innovation , Gamification, Communication and Stress reduction, working for national and international companies, public administration and universities.

PROGRAM

DAY 1 (May 7th)

Start time	Activity	Speaker/s	Chair
8.30h Registration			
9.00h	Welcome and opening	**Sergi Girona,** BSC Operations Director and CIO	Maria Ribera Sancho
9.20h	Keynote talk: Brain and behaviour: to which extent are we responsible for who we are?	**David Bueno i Torrens,** Universitat de Barcelona (UB)	
	Abstract: The brain is the organ of thought. Its neural networks manage all our behaviours. Its ontogenetic origin, however, is dual. On one hand, its formation and its functioning are conditioned by a number of genes, which make each person more or less prone to any cognitive ability as well as for any behavioural response. On the other hand, the environment also influences how the synaptic conexions are established, which in turn sets the neural networks, whose activity will generate all the behaviours and learning. What is the relative influence of each of these factors? Can we contribute to the construction of our brain? In this keynote we will discuss to which extent we are responsible for being who we are.		
10.30h Event Photo			
Coffee break & First Poster Session			
10.40h	Development of HPC Multiphysics Framework for HTS Magnets in Fusion, **José Lorenzo** Deciphering the interactions between the immune system and cancer cells to enable precision medicine, **Victoria Ruiz-Serra** Local traffic contribution to black carbon horizontal and vertical profiles in compact urban areas, **Jaime Benavides** De novo binding prediction of peptides to MHC class I, **Pep Amengual-Rigo** PluriZymes: new enzymes for new times, **Marc Domingo Cabasés**		
11.40h	**First Talk Session: Algorithms, plastic biodegradation, wave propagation and seismology**		
	1. Towards PET degradation engineering	**Sergi Rodà**	
	2. Maximal Entanglement in Quantum Computation	**Alba Cervera**	
	3. Solving Parameteric Wave Progapation Models with Domain Decomposed Reduced Order Methods	**Prattya Datta**	Victor Guallar
	4. Multifractal characterization of seismicity: the case of Carterbury region (New Zealand), 2000-2018	**Marisol Monterrubio Velasco**	
13.00h Lunch Break			
14.00h Tutorial 1st part			
	Creativity & Innovation: Why creativity is needed in a science career?	**Alicia Marín Muniesa, Isabel Nogueroles, Judit Murlans**	
	Goals & Content		

Goal: The goal of this workshop is to set the scene to value creativity and innovation in a scientific environment.

Contents:

Creative Thinking:
  Creativity vs. Innovation
  Uses in the scientific environment
Requirements and strategies to foster creativity and innovation
  Barriers and blockages
Practices and resources to develop creativity and innovation
Collecting challenges to solve them with the help of creativity techniques
Topics for this workshop:

- · Group awareness - Troubleshooting the challenges of early-career research

- · Communication - Understanding my audience and pitching my message

- · Project Management – How can I prioritise my tasks?

16.00h Adjourn

DAY 2 (May 8th)

Start time	Activity		Chair
9.00h	Opening of the second day		
9.10h	Keynote talk 2: Preparing for Extreme Heterogeneity in High Performance Computing	**Jeffrey Vetter**, Oak Ridge National Laboratory (ORNL)	Petar Radojkovic
	Abstract:While computing technologies have remained relatively stable for nearly two decades, new architectural features, such as heterogeneous cores, deep memory hierarchies, non-volatile memory (NVM), and near-memory processing, have emerged as possible solutions to address the concerns of energy-efficiency and cost. However, we expect this 'golden age' of architectural change to lead to extreme heterogeneity and it will have a major impact on software systems and applications. Software will need to be redesigned to exploit these new capabilities and provide some level of performance portability across these diverse architectures. In this talk, I will sample these emerging memory technologies, discuss their architectural and software implications, and describe several new approaches (e.g., domain specific languages, intelligent compilers and introspective runtime systems) to address these challenges.		

10.20h	**Coffee break & Second Poster Session:**
	An ILP-based Real-Time Scheduler for Distributed and Heterogeneous Computing Environments, **Eudald Sabaté Creixell**
	Techniques for reducing and bounding OpenMP dynamic memory, **Adrián Munera Sánchez**
	C/R Support for Heterogeneous HPC Applications, **Konstantinos Parasyris**
	High-Integrity GPU Designs for Critical Real-Time Automotive Systems, **Sergi Alcaide Portet**

11.30h	**Second Talk Session: Data Analytics, genetic variability, distributed computing and seismology**		
	1. Orchestration of Software Packages in Data Science Workflows	**Cristian Ramon-Cortes Vilarrodona**	
	2.Characterization of Structural Genomic Variability in Population Cohorts	**Jordi Valls**	Rosa Badia
	3. Cyberinfrastructure programming with COMPs	**Francesc Lordan**	
	4. Assessment of Damage Potential of Seismic Ground Motions	**Armando Aguilar Meléndez**	

13.00h	**Lunch Break**		
14.00h	**Third Talk Session: Machine Learning, personalized medicine and air quality**		
	1. Analysis of the interaction of genomic variants and their association to common diseases	**Lorena Alonso Parrilla**	
	2. Correcting Air Quality Forecasts with Machine Learning Algorithms	**Hervé Petetin**	Eduard Ayguadé
	3.TauRieL: Targeting Traveling Salesman Problem with deep reinforcement learning	**Gorker Alp Malazgirt**	
	4. Training CNNs using high-resolution images of variable shape	**Ferran Parés Pont**	

15.30h Tutorial session 2

Creativity & Innovation: Working on real challenges with Design Thinking techniques	**Alicia Marín Muniesa, Isabel Nogueroles, Judit Murlans**

Content&Goals Goal: The goal of this workshop is to provide scientists with practical tools and resources that can be applied in their working environment to bring innovation to life.

Contents:

    Presenting challenges to be solved in teams
    The process of creative problem solving
    Introducing and practicing with applied creativity techniques:
        Design Thinking (IDEO)
        Design Sprint (GOOGLE)
    Participants present their creative solutions
    Individual Action Plan to transfer learning to the workplace

17.30h   Adjourn

DAY 3 (May 9th)

Activity		Chair
9.00h Opening of the third day		

9.10h Fourth Talk Session: HPC, biological simulation and computer architecture

1. Enhancing Scheduling through Monitoring and Prediction Techniques	**Antoni Navarro Muñoz**	
2. Exploration of architectural parameters for future HPC systems	**Constantino Gómez**	
3. Supporting task creation inside FPGA devices	**Jaume Bosch**	Filippo Mantovani
4. Containers in HPC: A Scalability and Portability Study in Production Biological Simulations	**Oleksandr Rudyy**	
5. Experimental Study of Aggressive Undervolting in FPGAs	**Behzad Salami**	

10.50h Coffee break & Third Poster Session

A FM-index transformation to enable large k-steps, **Rubén Langarita**

A multilayer network approach to elucidate severity in Congenital Myasthenic Syndromes, **Iker Núñez**

Deep Learning Phase Picking of Large-N experiments, **Luis Fernández-Prieto**

Performance optimization of fully anisotropic elastic wave propagation on 2nd Generation Intel Xeon Phi processors, **Albert Farres**

11.50h Fifth Talk Session: Meshes and Fluids, Hemodynamics

1. Towards a low dissipation FE scheme for scale resolving turbulent compressible flows	**Lucas Gasparino**	
2. Local bisection for conformal refinement of unstructured 4D simplical meshes	**Guillem Belda Ferrín**	
3. Defining a stretching and alignment aware quality measure for linear and curved 2D meshes	**Guillermo Aparicio Estrems**	Arnau Folch
4. A One-Dimensional Finite Element Model for Human Circulatory Systems	**David Oks**	

13.15 Lunch

14.30h Sixth Talk Session: Models and simulation, dust and pathogenicity prediction

1. Correctly modeling IR spectra of astronomical interesting nanosilicate clusters	**Joan Mariñoso Guiu**	
2. How much soil dust aerosol is man-made?	**Martina Klose**	Oriol Jorba
3. Structural and Dynamics Analysis of Pyruvate Kinase from Erythrocytes: Implications in Pathology	**Luis Jordà**	

16.00h Adjourn/End of Doctoral Symposium

Talk Abstracts

Assessment of Damage Potential of Seismic Ground Motions

Armando Aguilar-Meléndez*†, Josep De la Puente*, Marisol Monterrubio-Velasco*, Otilio Rojas*
E. Ulises Moya-Sanchez‡, Octavio Castillo-Reyes*
*Barcelona Supercomputing Center, Barcelona, Spain
†Universidad Veracruzana, Poza Rica, Mexico
‡ Gobierno del Estado de Jalisco, Mexico
E-mail: {armando.aguilar, josep.delapuente, marisol.monterrubio, otilio.rojas, octavio.castillo}@bsc.es;
emoya@cs.upc.edu

Keywords—Seismic ground motions, CAV, seismic damage.

I. INTRODUCTION

One reason to study earthquakes is contributing to reduce or avoid the damage that some earthquakes trigger. In this topic, the analysis of seismic records is an important activity to obtain valuable information about the features of the seismic waves on a site. For instance, this analysis allows assessing diverse parameters that have been identified as a reasonable reference to assess potential damage of seismic ground motions. In this document, we described some relevant aspects of the analysis of seismic records with the purpose of assessing parameters related to the damage potential associated with earthquakes. For this objective, we applied the computer code Seismograms Analyzer-e [1]. On the other hand, the earthquakes that are analyzed in the present document have in common that in all of them occurred damage in buildings and unfortunately, people died.

II. ANALYSIS AND PROCESSING OF SEISMIC GROUND MOTIONS RELATED TO TWO DAMAGING EARTHQUAKES OF MEXICO

Mexico City recives seismic waves that are generated by earthquakes that occur in different regions of Mexico. For instance, Table I shows fundamental data about two damaging earthquakes that affected to Mexico City. The first earthquake is known as the Michoacán earthquake and it had its epicenter near to the Coast of Michoacán. The second event is Puebla earthquake and it had its epicenter in the south region of Puebla.

A. Michoacán earthquake of magnitude 8.1 (September 19, 1985)

According to Table I Michoacán earthquake of 1985 produced a maximum PGA (Peak Ground Acceleration) equal to 169 cm/s^2 at SCT station located at one of the diverse regions of soft soil in Mexico City. However, it is important to notice that the SCT station is located 425 km from the epicenter. Similarly, this earthquake produced a maximum PGA equal to 33.8 cm/s^2 at UNAM station which is located at a rock site in Mexico City (Table I). Therefore, for this earthquake, the PGA at the SCT station was 5 times greater than the PGA at the UNAM station. This type of behaviour during the Michoacán

earthquake was the key to confirm the importance of the local effects in the final features of the seismic waves that arrive at a specific site. The same earthquake produced a maximum PGA equal to 258 cm/s^2 at Zacatula, Michoacán located to 81 km from the epicenter.

B. Puebla earthquake of magnitude 7.1 (September 19, 2017)

One of the particularities of this earthquake is the fact that it also occurred on September 19 as in the case of the 1985 earthquake. However, in this case, the earthquake occurred 32 years after the 1985 earthquake. Unfortunately, in both

TABLE I.   BASIC DATA AND RESULTS RELATED TO SEISMIC RECORDS OF TWO DAMAGING EARTHQUAKES OF MEXICO

EQ	Seismic station	dist [km]	PGA & Comp. [cm/s2]	CAV_{STD} [cm/s]	Warning CAV_{STD}[1]
Michoacan Mexico	1.SCT Mexico City	425	96.7 NS	1630	Yes
09-19-1985 M=8.1			169 EW	1980	Yes
	2.UNAM Mexico City	419.5	31.1 NS	89.3	No
			33.8 EW	55.6	No
	3.MCH Zacatula	81	258 NS	2080	Yes
			171 EW	1720	Yes
Puebla Mexico	1.UNAM Mexico City	116.4	44.3 NS	183	Yes
09-19-2017 M=7.1			51.7 EW	213	Yes

[1] In this column the word Yes means that CAV_{STD} is greater than 156.96 cm/s (0.16 g.s) which is the threshold chosen for this study. Similarly, the word No means that CAV_{STD} is not greater than 156.96 cm/s

earthquakes the amount of damage triggered in Mexico City was important. According to Table I the Puebla earthquake of 2017 generated a maximum PGA equal to 51.7 cm/s^2 at the UNAM station. This station is located to 116.4 km from the epicenter and is situated in a rock site.

III. DAMAGE POTENTIAL OF GROUND MOTIONS

The seismic records give us relevant information. For instance, the seismic records are an essential data for the appropriate operation of Earthquake Early Warning systems, because the rapid detection of the size of an earthquake allows generating useful warnings to evacuate buildings, to activate special protection protocols, etcetera. At the same time, there is interest in to know the damage potential related to each seismic record. For this purpose different parameters have been proposed. One of these parameters is CAV_{STD} (standardized cumulative absolute velocity) [2]. The computation of this parameter can give us an idea of the damage potential of a ground motion. For instance, just when an earthquake occurred, we can process automatically the seismic record to compute the CAV_{STD} and the result can be included in a report that can be sent automatically to diverse stakeholders but essentially to authorities of civil protection. Then if values of CAV_{STD} exceed a threshold then the probabilities of damage in buildings can be considered as relevant. Therefore, in these cases, the authorities can obtain a rapid assessment of the urban area to identify the regions with more probabilities to suffer damage in their buildings and infrastructure. Consequently, this information can be used as a reference to organize emergency activities.

For instance, we can use 0.16g.s (156.96 cm/s) [3] for the threshold of CAV_{STD}. Therefore, according to this reference value and the data of Table I, we can identify that if a new earthquake similar to the Michoacán earthquake of 1985 occur
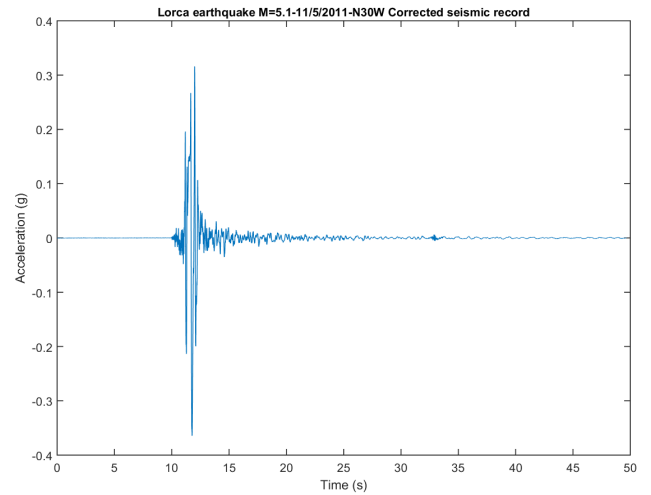

Fig. 1. The nort-30-west component of the LOR accelerogram obtained in the 2011 Lorca (Spain) earthquake. This seismic record was processed by Seismograms Analyzer-e software. [1].

in Mexico we would expect the following: a) the generation of significant seismic damage in the regions where both the SCT (Mexico City) and MCH (Zacatula, Michoacán) stations are located; b) a significantly greater concentration of damage in the SCT region of Mexico City than the concentration of damage in the UNAM region. This last affirmation is based on the fact that the maximum CAV_{STD} for the SCT station is equal to 1980 cm/s which is significantly greater than the threshold of 156.96 cm/s, and it is also based on the certainty that the maximum CAV_{STD} for the UNAM station is equal to 89.3 cm/s which is lower than the 156.96 cm/s. Therefore, if a new earthquake similar to the Michoacán earthquake occurs then the SCT region will require major emergency resources than the UNAM region.

On the other hand, for the Puebla earthquake of 2017, we would identify that seismic damage can occur at the buildings located in the same region than the UNAM station due to the fact that the maximum CAV_{STD} for this site is equal to 213 cm/s which exceeds the reference value of 156.96 cm/s.

The same previous criteria can be used for the earthquakes of Table II. For instance, in the Lorca earthquake case, it is possible to observe that CAV_{STD} value due to the E30N component does not exceed the reference value of 156.96 cm/s, but the CAV_{STD} value due to the N30W component is greater than the mentioned reference value. Therefore, in this case, the computation of these values suggest that the ground motion generated by the Lorca earthquake could generate seismic damage in buildings in the region where the LOR station is located. For this earthquake, the maximum PGA recorded was equal to 364 cm/s^2 (Fig. 1).

Additionally, we can observe that for the seismic records of the L'Aquila earthquake and of the Colima earthquake (Table II), the values of CAV_{STD} exceeds the reference value of 156 cm/s. Moreover, we also can observe that the maximum CAV_{STD} for the L'Aquila earthquake is equal to 1110 cm/s for a seismic record with a PGA equal to 612 cm/s^2 (Fig. 2). But, for the Colima earthquake, the maximum CAV_{STD} is equal to 407 cm/s for a seismic record with a PGA equal to 1250 cm/s^2. Therefore, according to the values

TABLE II. MAIN DATA OF THREE EARTHQUAKES THAT TRIGGERED DAMAGE IN URBAN AREAS

EQ	Seismic station	dist [km]	PGA & Comp. [cm/s^2]	CAV_{STD} [cm/s]	Warning CAV_{STD}[2]
Lorca Spain	1.LOR Lorca City	4.6	364 N30W	211	Yes
11-05-2011 M=5.1			154 E30N	113	No
LAquila Italy	1.Aterno LAquila City	4.85	612 EW	1110	Yes
04-06-2009 M=6.3					
Colima Mexico	1.COJ Cofradia	37	1250 NS	407	Yes
31-05-2007 M=5.1			680 EW	254	Yes

[2]

In this column the word Yes means that CAV_{STD} is greater than 156.96 cm/s (0.16 g.s) which is the threshold chosen for this study. Similarly, the word No means that CAV_{STD} is not greater than 156.96 cm/s
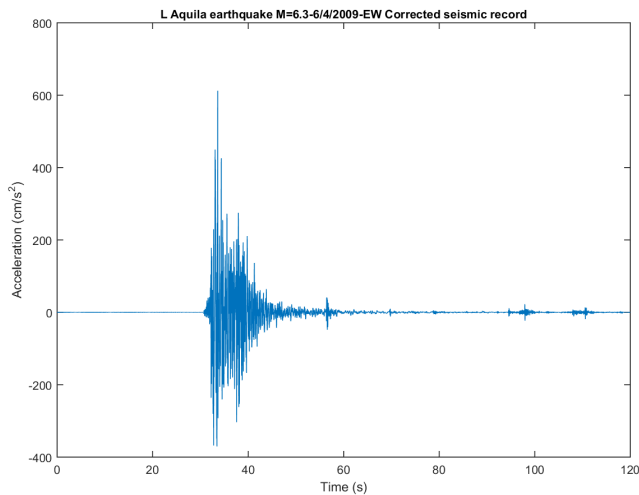
Fig. 2. The east-west component of the Aterno accelerogram obtained in the 2009 L'Aquila (Italy) earthquake. This seismic record was processed by Seismograms Analyzer-e software. [1].

of PGA and CAV_{STD} for both the L'Aquila and Colima earthquakes, respectively, we can observe that there is not a linear relationship between PGA and CAV_{STD}. However, in the cases analyzed in the present document the value of 156 cm/s was an appropriate threshold, because in the whole cases of the present study damage occurred when CAV_{STD} exceeded 156 cm/s.

IV. Conclusions and future works

According to the preliminary results obtained in the present study, it is possible to confirm that CAV_{STD} is a reasonable parameter to assess the damage potential of specific seismic ground motion. However, in order to do an extensive sensitivity analysis about the reliability in the threshold equal to 0.16g.s, it is convenient to compute CAV_{STD} using a significantly greater database of ground motions than the one used in this study. This new work will allow examining with major detail the reliability of CAV_{STD} as an estimator of the damage potential of seismic ground motions, and at the same time, this new work could be used to try to define more than one threshold in function of probabilities of occurrence of different levels of damage.

V. Acknowledgment

References

[1] A. Aguilar-Melendez *et al.*, "Seismograms Analyzer-e. Program for analysis of seismic records," 2018. [Online]. Available: https://sites.google.com/site/seismogramsanalyzere/home

[2] C. K. W. and B. Y., "Cumulative absolute velocity (CAV) and seismic intensity based on the PEER-NGA database," *Earthquake Spectra 28(2)*, 2012.

[3] EPRI, "Use of cumulative absolute velocity (cav) in determining effects of small magnitude earthquakes on seismic hazard analysis. in: Electric power research institute, electric power research institute, palo alto, ca, prepared by ares corporation inc and norm a abrahamson inc, report no: Rs1014099)." Springer., 2006.

Armando Aguilar is a Postdoctoral Researcher Conacyt-BSC in the CASE Department and a professor of the Faculty of Civil Engineering of the Universidad Veracruzana.He is a member of the Sistema Nacional de Investigadores (SNI-C).

Josep De la Puente received his PhD in the Ludwing Maxmillian University in 2007. His main topic research is about computational seismology. Currently, he is the manager of the geosciences application group of the Barcelona Supercomputing Center.

Marisol Monterrubio-Velasco received her PhD from the University Politécnica de Catalunya (Spain) and was a postdoctoral researcher at Geosciences center, UNAM. At present she is postdoc in the Barcelona Supercomputing Center. Her research interests focus on computational physics, statistical analysis and numerical simulation applied to Earth phenomena (fractal structures, earthquakes behavior, statistical analysis)

Otilio Rojas is a senior researcher at the Barcelona Supercomputing Center (BSC) and a professor at the Computer Science School of the Universidad Central de Venezuela (UCV).

E. Ulises Moya-Sanchez is currently, the Artificial Intelligence director of the State of Jalisco, Mexico. He is a member of the Sistema Nacional de Investigadores (SNI-C). He earned his Bachelor of Science (Physics) at Universidad de Guadalajara in 2004, his master degree (medical physics) at UNAM in 2007 and he received his PhD degree in 2014 from CINVESTAV. From 2017-2018 he was a postdoctoral researcher in High-Performance Artificial Intelligence group at Barcelona Supercomputing Center. His research interests span both high-performance computing and artificial intelligence.

Octavio Castillo Reyes is a postdoctoral researcher at the Barcelona Supercomputing Center (BSC) and a partial-time associate professor, Facultad de Informática, Universitat Politécnica de Catalunya.

Defining a stretching and alignment aware quality measure for linear and curved 2D meshes

Guillermo Aparicio-Estrems, Abel Gargallo-Peiró, Xevi Roca

Barcelona Supercomputing Center (BSC), Barcelona, Spain

E-mail: {guillermo.aparicio, abel.gargallo, xevi.roca}@bsc.es

Abstract—We define a regularized shape distortion (quality) measure for curved high-order 2D elements on a Riemannian plane. To this end, we measure the deviation of a given 2D element, straight-sided or curved, from the stretching and alignment determined by a target metric. The defined distortion (quality) is suitable to check the validity and the quality of straight-sided and curved elements on Riemannian planes determined by constant and point-wise varying metrics. The examples illustrate that the distortion can be minimized to curve (deform) the elements of a given high-order (linear) mesh and try to match with curved (linear) elements the point-wise alignment and stretching of an analytic target metric tensor.

I. Extended Abstract

A. Introduction

In the last decades, the utilization of unstructured meshes composed by highly stretched elements and aligned with dominant flow features, such as boundary layers and shock waves, have shown to be very advantageous. When compared with uniform refinement or with isotropic meshes with non-uniform sizing, anisotropic meshes lead to a significant reduction on the number of required degrees of freedom to obtain the same approximation accuracy. This allows performing simulations with a significantly reduced, and even unbeatable, computational cost.

The generation of anisotropic meshes requires to determine the location, stretching and alignment of the elements. These features can be prescribed manually with the help of the user interface of a mesh generation environment. They can also be prescribed imposing point-wise varying metric tensors obtained in an automatic and iterative adaption procedure based on error indicators or estimators. Then, an anisotropic mesher can be used to match the resolution, stretching and alignment determined by the target metric.

It is standard to use parallelotopes (quadrilaterals and hexahedra) to manually prescribe the alignment and stretching required to capture flow features such as boundary layers. Whereas the flexibility of simplices (triangles and tetrahedra) is the preferred one in automatic adaption iterations. Nevertheless, for both types of elements, the most mature anisotropic mesh generation techniques lead to meshes featuring second order elements such as multi-linear parallelotopes and linear simplices.

The utilization of curved anisotropic meshes composed by third order elements, such as multi-quadratic parallelotopes and quadratic simplices, or piece-wise polynomial elements of higher order has been mainly centered to curve, manually prescribed, straight-sided boundary layer meshes [1]–[9]. It has

not been until recently that the first metric based approaches have been explored to generate anisotropic meshes featuring straight-sided very high-order three dimensional approximations, curved quadratic triangles, and r-adapted curved high-order 2D elements. However, no specific efforts have been conducted to check the validity and measure the quality of curved high-order anisotropic meshes considering a prescribed metric tensor.

Our main contribution is to define a regularized shape distortion (quality) to measure the deviation of a given linear or high-order 2D element from the stretching and alignment determined by a target metric. The influence of the target metric on the element quality has only been considered in detail for linear elements and not for curved high-order elements. The defined distortion (quality) is suitable to check the validity and the quality of straight-sided and curved elements on Riemannian planes determined by constant and point-wise varying metrics. Furthermore, we illustrate that the distortion can be minimized to curve (deform) the elements of a given high-order (linear) mesh and try to match with curved (linear) elements the point-wise alignment and stretching of an analytic target metric tensor. Specifically, this approach can be used to improve, by curving (deforming) the elements, the alignment and stretching of a mesh obtained with a straight-sided anisotropic mesher.

B. Examples

In this section, we present an example to illustrate the main features of the proposed quality measure. It features an analytic and point-wise varying smooth metric. We generate an initial mesh and we measure its quality according to the new measure for anisotropic elements. Next, we optimize the location of the nodes to minimize the element distortion using the framework presented in [5], [10].

The mesh resulting after the optimization is composed by elements as aligned and stretched as possible to match the target metric tensor. In both figures, the meshes are colored according to the elemental quality.

As a proof of concept, a mesh optimizer has been developed in MATLAB using the Optimization Toolbox, the PDE Toolbox and the Symbolic Math Toolbox. The MATLAB prototyping code is sequential (one execution thread), corresponds to the implementation of the method presented in this work. The optimization is reduced to find a minimum of a nonlinear unconstrained multi-variable function, where a trust-region algorithm is used. The stopping condition is set to reach a relative residual smaller than 10^{-8}.
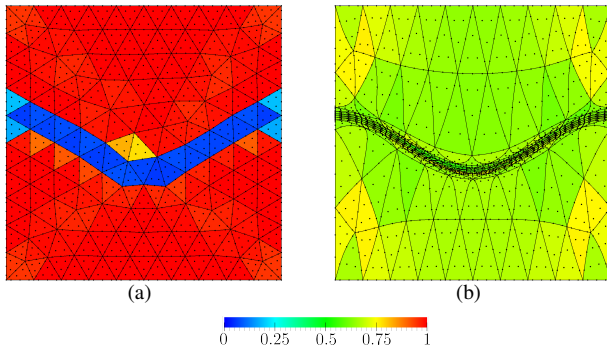
Fig. 1. Triangular meshes of polynomial degree 4 colored by quality: (a) initial straight-sided isotropic mesh, (b) optimized mesh from initial configuration.

Consider the quadrilateral domain $\Omega := [0,1]^2$ and the metric $\mathbf{M} := \nabla\boldsymbol{\varphi}^{\mathrm{T}} \cdot \nabla\boldsymbol{\varphi}$ induced by the surface $\boldsymbol{\varphi}$:

$$\boldsymbol{\varphi}(\mathbf{x}) := (x, y, z(x,y)), \quad z(x,y) := \tanh\left(f \cdot h(x,y)\right),$$

where

$$h(x,y) = 10y - \cos(2\pi x) - 5, \quad (x,y) \in \Omega, \ f = 5.$$

The metric is extracted from [11] and attains the highest level of anisotropy at the curve described by the points $(x,y) \in \Omega$ such that $h(x,y) = 0$. The anisotropy ratio of this metric is $1 : \sqrt{1 + |\nabla z|^2}$ and its maximum is approximately $1 : 60$. Note that the chosen φ surface is not related to a finite element solution. It is a standard analytical test to check anisotropic capabilities.

We illustrate this example for triangles in Figure 1. We generate an initial isotropic straight-sided high-order mesh 1(a) composed by 1905 nodes and 228 elements. The optimized mesh is illustrated in Figure 1(b). We observe that the elements away from the anisotropic region are enlarged vertically whereas the elements lying in the anisotropic region are compressed. In the optimized mesh the minimum is improved and the standard deviation of the element qualities is reduced when compared with the initial configuration.

C. Conclusion

In this work, we have presented a new definition of distortion (quality) measures for linear and high-order planar anisotropic meshes equipped with a point-wise metric. The proposed quality measures the alignment and stretching of the elements according to the given metric. In addition, it is valid for any interpolation degree and allow to detect the validity of a high-order element equipped with a metric. To assess the reliability of the technique, we have first analyzed the behavior of the measure for linear triangles equipped with a constant metric. The tests show that for a given metric the obtained quality measure detects invalid and low-quality configurations and, the alignments and stretching described by the metric.

The defined distortion measure is applied to curve linear meshes to improve the node configuration according to the desired metric. The numerical example show an optimized mesh with an improved alignment and stretching according to the metric. This improvement leads to an increase of the minimum elemental mesh quality and a reduction of the standard deviation between the different element qualities.

Our long term goal is to extend the quality measure for 3D anisotropic meshes. In addition, the quality measure developed in this work is devised to quantify the alignment and the stretching of the mesh according to the target metric. Thus, we would like to extend the proposed measure to also quantify the mesh sizing.

II. ACKNOWLEDGMENT

REFERENCES

[1] O. Sahni, X. Luo, K. Jansen, and M. Shephard, "Curved boundary layer meshing for adaptive viscous flow simulations," *Finite Elem. Anal. Des.*, vol. 46, no. 1, pp. 132–139, 2010.

[2] P.-O. Persson and J. Peraire, "Curved mesh generation and mesh refinement using lagrangian solid mechanics," in *Proc. 47th AIAA*, 2009.

[3] T. Toulorge, C. Geuzaine, J.-F. i. Remacle, and J. Lambrechts, "Robust untangling of curvilinear meshes," *Journal of Computational Physics*, vol. 254, pp. 8–26, 2013.

[4] A. Gargallo-Peiró, X. Roca, J. Peraire, and J. Sarrate, "Inserting curved boundary layers for viscous flow simulation with high-order tetrahedra," in *Research Notes, 22nd Int. Meshing Roundtable*. Springer International Publishing, 2013.

[5] ——, "Optimization of a regularized distortion measure to generate curved high-order unstructured tetrahedral meshes," *Int. J. Numer. Meth. Eng.*, vol. 103, pp. 342–363, 2015.

[6] D. Moxey, M. Green, S. Sherwin, and J. Peiró, "An isoparametric approach to high-order curvilinear boundary-layer meshing," *Computer Methods in Applied Mechanics and Engineering*, vol. 283, pp. 636–650, 2015.

[7] M. Fortunato and P.-O. Persson, "High-order unstructured curved mesh generation using the winslow equations," *Journal of Computational Physics*, vol. 307, pp. 1–14, 2016.

[8] A. Gargallo-Peiró, G. Houzeaux, and X. Roca, "Subdividing triangular and quadrilateral meshes in parallel to approximate curved geometries," *Procedia Engineering*, vol. 203, pp. 310–322, 2017.

[9] D. Moxey, D. Ekelschot, U. Keskin, S. J. Sherwin, and J. Peiró, "High-order curvilinear meshing using a thermo-elastic analogy," *Computer-Aided Design*, vol. 72, pp. 130–139, 2016.

[10] A. Gargallo-Peiró, "Validation and generation of curved meshes for high-order unstructured methods," Ph.D. dissertation, Universitat Politècnica de Catalunya, 2014.

[11] P. C. Caplan, R. Haimes, D. L. Darmofal, and M. C. Galbraith, "Anisotropic geometry-conforming d-simplicial meshing via isometric embeddings," *Procedia Engineering*, vol. 203, pp. 141–153, 2017, 26th International Meshing Roundtable.

Guillermo Aparicio-Estrems received his BSc degree in Mathematics from Universitat Politècnica de Catalunya (UPC), Spain in 2016. The following year, he completed his MSc degree in Applied Mathematics from Universitat Politècnica de Catalunya (UPC), Spain in 2017. Since 2017, he has been with the Geometry and Meshing for simulations group of Barcelona Supercomputing Center (BSC) as a PhD student of Universitat Politècnica de Catalunya (UPC), Spain.

Local bisection for conformal refinement of unstructured 4D simplicial meshes

Guillem Belda-Ferrín*, Abel Gargallo-Peiró*, Xevi Roca*

*Barcelona Supercomputing Center (BSC), Barcelona, Spain

E-mail: {guillem.belda, abel.gargallo, xevi.roca}@bsc.es

Abstract—**We present a conformal bisection procedure for local refinement of 4D unstructured simplicial meshes with bounded minimum shape quality. Specifically, we propose a recursive refine-to-conformity procedure in two stages, based on marking bisection edges on different priority levels and defining specific refinement templates. Two successive applications of the first stage ensure that any 4D unstructured mesh can be conformingly refined. In the second stage, the successive refinements lead to a cycle in the number of generated similarity classes and thus, we can ensure a bound over the minimum shape quality. In the examples, we check that after successive refinement the mesh quality does not degenerate. Moreover, we refine a 4D unstructured mesh and a space-time mesh (3D + 1D) representation of a moving object.**

Keywords—Meshing, Adaptive, 4D.

I. Extended Abstract

A. Introduction

In the last three decades refinement of 2D and 3D unstructured simplicial meshes [1]–[14], based on red/green refinement [1]–[7] and bisection [8]–[14], has been shown to be a key ingredient on efficient adaptive loops. Although one could expect the same in 4D, a case of special interest for space-time adaption, this line of research has not been extensively explored. For our space-time applications, we are interested in conformal bisection methods since they are really well suited to implement fast geometrical multi-grid conformal solvers. Moreover, bisection methods have ensured either a maximum number of generated similarity classes [11]–[13] or a minimum lower quality bound over the generated elements after successive refinements [8]–[10], [14]. Regarding 4D refinement, only a non-conformal local refinement method for pentatopic meshes has been proposed [15]. Unfortunately, existent conformal 4D (nD) bisection methods with a bound over the number of generated similarity classes [11], [12] cannot be applied to general unstructured meshes.

The main contribution of this work is to propose a local bisection procedure, with a bound over the number of generated similarity classes, for conformal refinement of 4D unstructured simplicial meshes. Specifically, we propose a recursive refine-to-conformity procedure, in two stages, based on marking bisection edges on different priority levels. The marking procedure allows classifying the pentatopes in different types and hence, determining different refinement templates, in an analogous manner to the 3D bisection method proposed in [13]. The refinement method is composed of two stages. Two successive applications of the initial stage of the bisection strategy , based on the proposed element classification, ensure that any initial 4D unstructured simplicial mesh can be conformingly refined. After the two initial refinements our recursive refine-to-conformity strategy switches to the second stage. This final stage is analogous to Maubach's algorithm, when it is successively applied to a single pentatope. Therefore, we can ensure a bound over the number of generated similarity classes. Thus, the minimum quality of the refined mesh is bounded, independently of the number of performed refinements. The main advantage and difference of our method when compared to Maubach's algorithm [11] is the first stage of the method, which allows the application of the method to any 4D unstructured simplicial mesh.

In the presented example we show that the proposed methodology leads to a periodic evolution of the minimum element quality illustrating the lower bound of the quality through successive refinement. We first illustrate how to check that an implementation of the proposed method is valid by successively refining a pentatope. With our implementation, we show that the proposed bisection technique can be used to refine general unstructured 4D meshes. Finally, we also illustrate our application of interest, the refinement of a 4D mesh corresponding to a space-time representation, with varying resolution, of the temporal evolution of a 3D moving object.

B. Examples

We illustrate our application of interest, the refinement of a 4D mesh corresponding to a space-time representation, with varying resolution, of the temporal evolution of a 3D moving object, see Figure 1. We generate an initial mesh on the hypercube $[0,1]^4$ composed by 24 pentatopes using Freudenthal-Kuhn algorithm [1]–[3]. Next, we apply 25 times the algorithm RefineToConformity to refine those elements that intersect the 4D sphere extrusion that represents the moving sphere. The final 4D mesh is composed by 5233296 pentatopes and 251457 nodes and it is illustrated in Figure 1. Figures 1(a)-1(c) show three slices of the mesh at $t = 0$, $t = 1/2$ and $t = 1$, respectively. We can observe that each one of the slices on t shows different positions of the moving sphere, from the initial point (0,0,0) at $t = 0$ to the final point $(0,0,1)$ at $t = 1$. In contrast with these three slices, in Figure 1(d) we show an slice of the mesh at $x = 0$. In the closest quadrilateral face of Fig. 1(d) we observe the path of the sphere on the surface of dimension 2 defined by the axis z and t at $x = y = 0$. In this quadrilateral face, we can see that the center of the sphere describes a straight line going from the lower left corner $(0,0,0,0)$ up to the top right corner $(0,0,1,1)$. This is so since the sphere goes from $z = 0$ to $z = 1$ with constant velocity starting at $t = 0$ and finalizing at $t = 1$. Specifically, the location on the z-axis of the sphere is $z = t$.
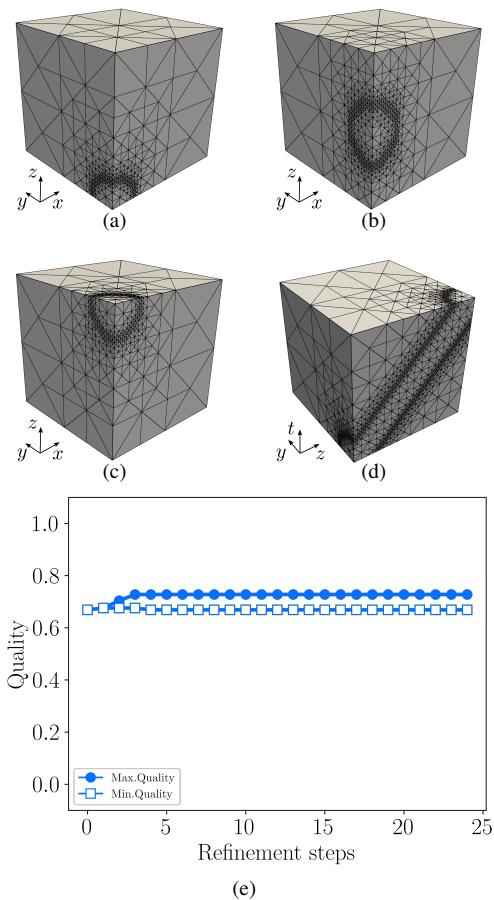
Figure 1. Slice of the 4D simplicial mesh of the hypercube with the hyperplane: (a) $t = 0$, (b) $t = 0.5$, (c) $t = 1$ and (d) $x = 0$. (e) Minimum (blue) and maximum (red) element quality for each refinement step.

C. Conclusion

In this work, we have presented a new refinement method via edge bisection for 4D pentatopic meshes. This method ensures that the mesh quality does not degenerate after successive refinements of a given element. To develop this method, we require to classify the elements of the mesh into different types in a similar fashion to [13]. Using the pentatope classification we provide four refinement templates to perform a cyclic bisection analogous to Maubach's method [11]. Combining two initializing refinements (Stage 1) with this templated refinement (Stage 2) we obtain a refinement strategy that can be applied to any given pentatopic mesh. Using this method a finite number of similarity classes are generated when a given element is refined. We apply the refinement scheme to different meshes to illustrate its features. First, we analyze that the mesh quality of the refinement of different element types does not degenerate. Second, we illustrate the applicability of the technique to refine unstructured 4D simplicial meshes. Finally, we analyze a space-time configuration of a sphere moving along an axis.

References

[1] H. Freudenthal, "Simplizialzerlegungen von beschrankter flachheit," *Ann. Math.*, pp. 580–582, 1942.

[2] H. Kuhn, "Some combinatorial lemmas in topology," *IBM Journal of research and development*, vol. 4, no. 5, pp. 518–524, 1960.

[3] J. Bey, "Simplicial grid refinement: on freudenthal's algorithm and the optimal number of congruence classes," *Numerische Mathematik*, vol. 85, no. 1, pp. 1–29, 2000.

[4] R. Bank *et al.*, "Some refinement algorithms and data structures for regular local mesh refinement," *Scientific Computing*, vol. 1, pp. 3–17, 1983.

[5] J. Bey, "Tetrahedral grid refinement," *Computing*, vol. 55, no. 4, pp. 355–378, 1995.

[6] A. Liu and B. Joe, "Quality local refinement of tetrahedral meshes based on 8-subtetrahedron subdivision," *Mathematics of Computation*, vol. 65, no. 215, pp. 1183–1200, 1996.

[7] S. Zhang, "Successive subdivisions of tetrahedra and multigrid methods on tetrahedral meshes," *Houston J. Math*, vol. 21, no. 3, pp. 541–556, 1995.

[8] M. Rivara, "Algorithms for refining triangular grids suitable for adaptive and multigrid techniques," *International journal for numerical methods in Engineering*, vol. 20, no. 4, pp. 745–756, 1984.

[9] E. Bänsch, "Local mesh refinement in 2 and 3 dimensions," *IMPACT of Computing in Science and Engineering*, vol. 3, no. 3, pp. 181–191, 1991.

[10] A. Liu and B. Joe, "Quality local refinement of tetrahedral meshes based on bisection," *SIAM Journal on Scientific Computing*, vol. 16, no. 6, pp. 1269–1291, 1995.

[11] J. Maubach, "Local bisection refinement for n-simplicial grids generated by reflection," *SIAM Journal on Scientific Computing*, vol. 16, no. 1, pp. 210–227, 1995.

[12] C. Traxler, "An algorithm for adaptive mesh refinement in n dimensions," *Computing*, vol. 59, no. 2, pp. 115–137, 1997.

[13] D. Arnold *et al.*, "Locally adapted tetrahedral meshes using bisection," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 431–448, 2000.

[14] A. Plaza and M. Rivara, "Mesh refinement based on the 8-tetrahedra longest-edge partition." in *IMR*, 2003, pp. 67–78.

[15] M. Neumüller and O. Steinbach, "A flexible space-time discontinuous galerkin method for parabolic initial boundary value problems," *Berichte aus dem Institut für Numerische Mathematik*, vol. 2, 2011.

Guillem Belda Ferrín received his BSc degree in Mathematics from Universitat Politècnica de Catalunya, Barcelona in 2014. He completed his MSc degree in Advanced Mathematics and Mathematical Engineering from Universitat Politècnica de Catalunya in 2015. Since 2017, he has been with the Geometry and Meshing for simulations group of Barcelona Supercomputing Center (BSC) as a PhD student of the Applied Mathematics doctorate program of Universitat Politècnica de Catalunya (UPC), Spain.

Supporting task creation inside FPGA devices

Jaume Bosch*†, Carlos Álvarez*†, Daniel Jiménez-González*†

*Barcelona Supercomputing Center, Barcelona, Spain
†Universitat Politècnica de Catalunya, Barcelona, Spain
E-mail: jbosch@bsc.es, {calvarez, djimenez}@ac.upc.edu

Keywords—Heterogeneous computing, Device offloading, Task Based Parallel Programming Models, High-performance computing.

I. Introduction

The most common model to use co-processors/accelerators is the master-slave model where the slaves (co-processors/accelerators) are driven by a general purpose cpu. This simplifies the management of the accelerators because they cannot actively interact with the runtime and they are just passive slaves that operate over the memory under demand. However, the master-slave model limits system possibilities and introduces synchronization overheads that could be avoided.

To overcome those limitations and increase the possibilities of accelerators, we propose extending task based programming models (like OpenMP [1] or OmpSs) to support some runtime APIs inside the FPGA co-processor. As a proof-of-concept, we implemented our proposal over the OmpSs@FPGA environment [2] adding the needed infrastructure in the FPGA bitstream and modifying the existing tools to support creation of children tasks inside a task offloaded to an FPGA accelerator. In addition, we added support to synchronize the children tasks created by a FPGA task regardless they are executed in a SMP host thread or they also target another FPGA accelerator in the same co-processor.

II. Design

The main design goal is to allow interaction of the FPGA with the runtime to make both parts cooperate in the application execution beyond the current offload model. The new interaction capabilities include the creation of tasks inside a device task and child task synchronization.

Listing 1 shows an example code where just replacing the line 6 (with clause `device(smp)`) by line 7 (`device(fpga)`), we will have a FPGA accelerator that creates and then synchronizes tasks inside the FPGA. All interactions are based on queues (memory regions) where the FPGA accelerators write requests to the runtime. Then, the runtime reads these requests and makes the needed actions. Some interactions are optimized and directly read and handled inside the FPGA, avoiding the latency between the host and the device. This way there is runtime support in both parts of the machine and both are coordinated when needed to correctly execute the application.

Queues and new IPs to manage them are placed inside the Task Manager as shown in Fig. 1. The proposed design

```
1  #pragma omp target device(fpga) num_instances(3)
2  #pragma omp task copy_inout([BS]a)
3  void update_fpga(int *a, int val, size_t BS) {
4      for (size_t i=0; i<BS; ++i) a[i] += val;
5  }
6  #pragma omp target device(smp)        //< Task exec. in SMP
7  //#pragma omp target device(fpga)  //< Task exec. in FPGA dev.
8  #pragma omp task copy_inout([LEN]a)
9  void update_blocked(int *a, int val, size_t LEN, size_t BS) {
10     for (size_t i=0; i<LEN; i+=BS)
11         update_fpga(a+i, val, BS);
12     #pragma omp taskwait
13 }
14 int main(...) {
15     int *a = (int *)malloc(NUM_ELEMENTS*sizeof(int));
16     update_blocked(a, 2019, NUM_ELEMENTS, NUM_ELEMENTS_BLOCK);
17     #pragma omp taskwait
18 }
```

Listing 1: OmpSs example of FPGA nested tasks

includes three new memory modules: Internal Ready Queue (Int. Ready Q), Remote New Queue (Remote New Q) and Remote Finished Queue (Remote Fini. Q); four new IPs: Remote Finished Task Manager (Rem. Fini. TM), New Task Manager (New TM), Taskwait Task Manager (Taskwait TM) and Scheduler Task Manager (Scheduler TM); and an additional interconnection network between the accelerators with creation capabilities and the Task Manager. The new Task Manager elements are automatically added to the design by autoVivado when needed by any of the accelerators. Moreover, Mercurium adapts the wrapper around the accelerator to support the runtime API calls from user code. Those calls may be explicitly invoked by the user or automatically inserted by the compiler during the translation phase where the OpenMP directives are translated into API calls.
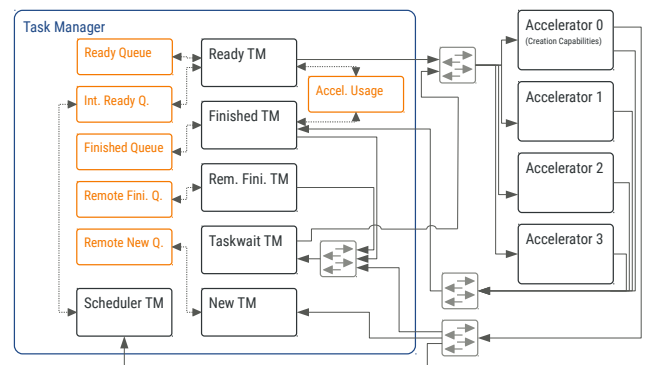

Fig. 1. OmpSs@FPGA bitstream organization with the proposed design

III. Evaluation

To analyze the performance, we used the synthetic benchmark in listing 1 that updates all elements of an array. Therefore, we parameterize the benchmark with the length of the array and the chunk length. The tools used to generate the application bitstream and binary are: Vivado Design Suite 2016.3, GNU C/C++ Compiler 6.2.0, PetaLinux Tools 2016.3 and all modified OmpSs@FPGA tools of release 1.2.1. The applications are run in a Zynq Ultrascale+ MPSoC Chip XCZU9EG-FFVC900 [3].

Fig. 2 shows the average execution time (y-axis) when decreasing the task size (chunk length) and increasing the number of tasks (x-axis). Note that the total amount of work remains constant for all the executions, as the array size equals the chunk length times the number of tasks. The same result is shown for different configurations, all of them executed within the OmpSs@FPGA environment. The label of each configuration defines the location where the tasks are created and the number of accelerators used to execute them.
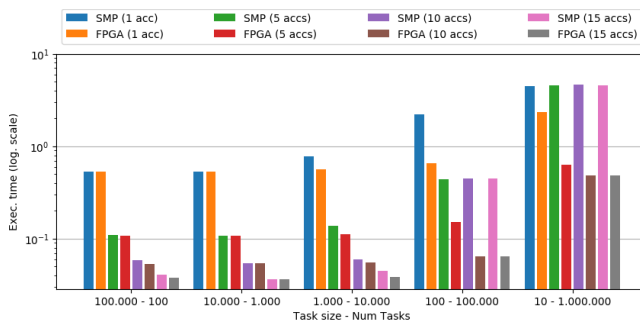
Fig. 2. Synthetic benchmark execution time with different configurations

The results show that the creation and management of FPGA tasks directly from the FPGA is faster than doing it from the host side. For any pair of benchmark arguments, the reduction of the execution time increases with the number of accelerators. This is because of the lower task creation overheads in our proposed implementation in comparison to the already implemented management in the host runtime. Consequently, the FPGA creation is able to discover more parallelism, increasing the accelerators utilization and reducing to overall execution time. In contrast, the host management and the FPGA management take a similar execution time when the task size is large enough to hide the runtime overheads with the tasks execution.
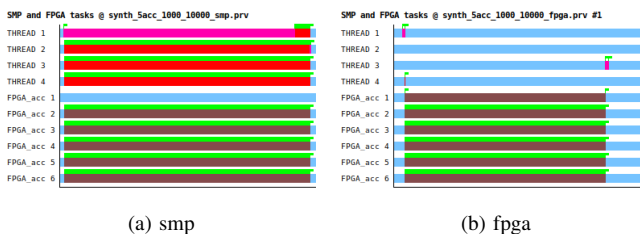

(a) smp   (b) fpga

Fig. 3. Execution traces of synthetic benchmark with 5 accelerators, 1.000 task size and 10.000 tasks

The execution traces in Fig. 3 show that the FPGA tasks execution take less time when they are directly created on the FPGA than when they are created and sent by the SMP threads. In addition, the FPGA_acc 1 accelerator does not execute any task when the FPGA tasks are created by the SMP threads and it runs the update_blocked task until all children tasks finish. The FPGA accelerators does not have the possibility to block a task and pick another one for execution, in contrast to SMP threads.

As Fig. 3b shows, the SMP threads are under-utilized when the FPGA creates the tasks directly. In contrast, they are needed in the SMP creation to create the tasks, offload them to the FPGA and retrieve the finalization messages that allow the host synchronize the remote tasks. With the proposed extension, we are able to replace the general purpose cores by an small FPGA accelerator that consumes a small portion of power when compared but does the same work.

IV. Conclusion

This paper presents an extension of the OmpSs@FPGA ecosystem to support task creation and synchronization operations in the FPGA co-processors. This extension enables a new dimension of possibilities for application programmers as they can mix tasks for different devices and nest them without restrictions. Finally, the initial performance results show that the creation of FPGA tasks inside the same FPGA device allows the usage of very fine-grain tasks thanks to the significant overheads reduction.

V. Acknowledgments

References

[1] L. Dagum and R. Menon, "OpenMP: an Industry Standard API for Shared-Memory Programming," *Computational Science Engineering, IEEE*, vol. 5, no. 1, pp. 46–55, Jan 1998.

[2] J. Bosch and et al., "Application Acceleration on FPGAs with OmpSs@FPGAS," *2018 International Conference on Field-Programmable Technology (FPT)*, 2018.

[3] Xilinx, Inc. (2019, April) ZYNQ Ultra-Scale+ MPSoC Overview. [Online]. Available: www.xilinx.com/support/documentation/data_sheets/ds891-zynq-ultrascale-plus-overview.pdf

Jaume Bosch received the B.S. and M.S. degrees in Computer Science from the Technical University of Catalunya (UPC) in 2015 and 2017, respectively. Currently, he is is a PhD student in the department of Computer Architecture of UPC and in the Programming Models Group of Barcelona Supercomputing Center (BSC). His research interest lies in parallel, distributed and heterogeneous runtime systems for High Performance Computing, specially systems with hardware accelerators like FPGAs or many-core architectures.

Maximal Entanglement in Quantum Computation

Alba Cervera-Lierta[*][†],

[*]Barcelona Supercomputing Center, Barcelona, Spain
[†]Institut de Cincies del Cosmos, Universitat de Barcelona, Barcelona, Spain
E-mail: {alba.cervera}@bsc.es

Keywords—Quantum computation, Quantum Information, Entanglement.

I. Extended Abstract

Entanglement is a quantum phenomenon that occurs when two or more quantum systems cannot be described independently from the others. Several experiments have highlighted the fact that this is a *genuine* quantum property that goes beyond any classical description [1]. One may expect that such a distinctive trait of quantum mechanics have several physical applications. Indeed, entanglement can also be understood as the resource that enables genuine quantum protocols such as cryptography based on Bell inequalities [2] and teleportation [3]. In addition, large entanglement is expected to be present in quantum registers when a quantum algorithm produces a relevant advantage in performance over a classical computer such as Shor's algorithm [4].

In recent years the quantum computing has dived fully into the experimental realm. Control of quantum systems has improved so much that quantum computing devices have become a near term reality. Private companies have also joined the field. Since 2016, IBM offers cloud based quantum computation platform [5]. It is not the only company that has launched this kind of service: Rigetti Computing also allows the use of its 19-qubits device on the cloud [6]. Although both companies are betting for superconducting qubits, their respective device characterization is not the same. As more quantum devices are appearing, it is important to find some methods to test their quality when running sophisticated quantum algorithms. In addition, there is a need to set up a thorough benchmarking strategy for quantum computers that prove the usefulness of these devices to perform tasks that classical computation can not achieve.

The main goal of this work is to present quantum algorithms to test current quantum computers. First, we introduce the exact circuit that simulates exactly the XY model Hamiltonian. This model can be solved analytically. Then, the results extracted after the experiment can be compared with the theoretical values and the possible discrepancies can be used to test the quality of the quantum device. Next, we present a hard but necessary test for a quantum computer: the simulation of absolutely maximally entangled (AME) states. Since entanglement is behind quantum advantage, quantum devices must be able to generate and hold highly entangled states.

A. Exact simulation of XY Hamiltonian on a Quantum Computer

Let's consider the existence of a quantum circuit that *disentangles* a given Hamiltonian and transforms its entangled eigenstates into product states. This circuit will be represented by an unitary transformation U_{dis} such that [7]

$$\widetilde{\mathcal{H}} = U_{dis}^\dagger \mathcal{H} U_{dis}, \tag{1}$$

where $\mathcal{H}$ is the model Hamiltonian and $\widetilde{\mathcal{H}}$ is a noninteracting Hamiltonian that can be written as $\widetilde{\mathcal{H}} = \sum_i \epsilon_i \sigma_i^z$. This diagonal Hamiltonian contains the energy spectrum ϵ_i of the original one and its eigenstates correspond to the computational basis states. Then, we will have access to the whole spectrum of the model by just preparing a product state and applying U_{dis}.

In general, to find this operation will be hard. However, for the case of XY Hamiltonian, the steps to obtain the U_{dis} quantum gate can be based on the analytical solution of the model: *i)* Implement the Jordan-Wigner transformation to map the spins into fermionic modes. *ii)* Perform the Fourier transform to get fermions to momentum space. *iii)* Perform a Bogoliubov transformation to decouple the modes with opposite momentum. Thus, the construction of the disentangling gate can be done by pieces:

$$U_{dis} = U_{JW} U_{FT} U_{Bog}. \tag{2}$$

Figure 1 shows the result of the exact simulation of an $n = 4$ spin chain with an Ising model interaction in three quantum computers. The Ising model is a particular case of the XY model. The quantum circuit used for this experiment is detailed in Ref. [10]. Left figure plots the expected value of the ground state transverse magnetization as a function of the external field. The right part of Fig. 1 shows the simulation of time evolution of the $|\uparrow\uparrow\uparrow\uparrow\rangle$ state. The results show a clear disagreement between the theoretical value and the experimental values. However, the error sources are systematic, as indicates the simulation of the time evolution.

B. Absolute Maximal Entanglement on a Quantum Computer

AME states are n qudit quantum states with local dimension d such that every reduction to $\lfloor n/2 \rfloor$ parties is maximally mixed. Such states are maximally entangled when considering the entropy of reductions as a measure of multipartite entanglement, that is, the average entropy of these states is $S = \lfloor n/2 \rfloor$ when taking the logarithm in d basis.

The existence of AME states for n qudit systems, denoted as $\text{AME}(n, d)$, is a hard open problem. Only for the case of
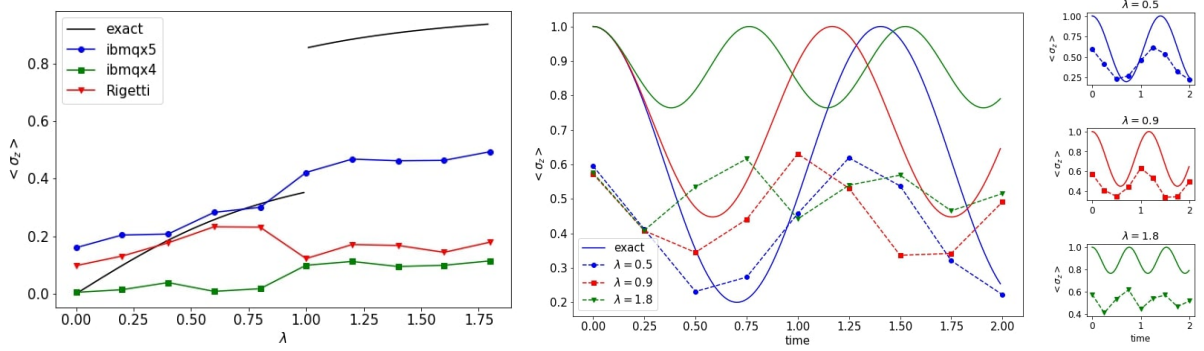
Fig. 1. *Left*. Expected value of $\langle \sigma_z \rangle$ of the ground state of a $n = 4$ Ising spin chain as a function of transverse field strength λ. Solid line represents the exact result in comparison with the experimental simulations represented by scatter points. The best simulation comes from ibmqx5 device, which is an expected result since the number of gates used is lesser than with the other devices because of qubits connectivity. *Right*. Time evolution simulation of transverse magnetization, $\langle \sigma_z \rangle$, for the state $| \uparrow\uparrow\uparrow\uparrow \rangle$ of a $n = 4$ Ising spin chain. Left plot compares the exact result with the experimental run in the ibmqx5 chip for different values of λ. Right plots detailed the results for each λ to compare them with the theoretical values. Although the magnetization is lesser than expected, the oscillations follow the same theoretical pattern.

qubit systems, i.e. $d = 2$, the problem is fully solved for any n: an AME$(n, 2)$ exists only for $n = 2, 3, 5, 6$. [11], [12]. For instance, Bell states and GHZ states are AME for bipartite and three partite systems of any d, respectively.

We propose to construct AME states quantum circuits using graph states. As an example, the AME state of five qubits can be implemented in a quantum computer using the following circuit:
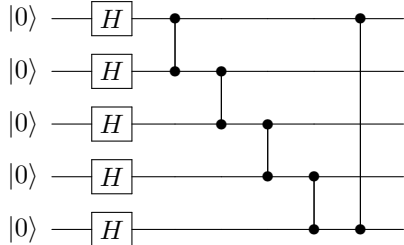
where H gates are Hadamard gates and vertical lines correspond with Controlled-Z (CZ) gates. In addition, some graph states work for any local dimension d. Then, the above circuit can be used to obtain AME states of five qudits of any dimension d by replacing Hadamard gates by Fourier gates and CZ gates by its generalized version.

II. CONCLUSIONS

We have presented two methods that can be used to test and benchmark quantum computers. The experimental results show that the current devices have still many error sources. However, they are capable to perform sophisticated algorithms that have interest, for instance, in condense matter physics and error correcting codes.

III. ACKNOWLEDGMENT

This work is part of the Thesis project title "Maximal Entanglement. Applications in Quantum Information and Particle Physics" to be published in July 2019. The program used to simulate the Ising model on a quantum computer was awarded with the IBM "Teach Me QISKit prize".

REFERENCES

[1] A. Aspect *et al.*, "Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities," *Physical Review Letters*, vol. 49, pp. 91–94, 1982.

[2] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, vol. 67, no. 6, pp. 661–663, 1991.

[3] C. H. Bennett *et al.*, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, pp. 1895–1899, 1993.

[4] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, 1997.

[5] IBM Q team and collaborators, "QISKit tutorials," 2018. [Online]. Available: https://github.com/Qiskit/qiskit-tutorials

[6] R. S. Smith *et al.*, "A practical quantum instruction set architecture," *arXiv:1608.03355 [quant-ph]*, 2016.

[7] F. Verstraete *et al.*, "Quantum circuits for strongly correlated quantum systems," *Physical Review A*, vol. 79, no. 3, p. 032316, 2009.

[8] E. Lieb *et al.*, "Two soluble models of an antiferromagnetic chain," *Annals of Physics*, 1961.

[9] S. Katsura, "Statistical mechanics of the anisotropic linear Heisenberg model," *Physical Review*, 1962.

[10] A. Cervera-Lierta, "Exact Ising model simulation on a quantum computer," *Quantum*, vol. 2, p. 114, 2018.

[11] W. Helwig, "Absolutely Maximally Entangled Qudit Graph States," 2013.

[12] F. Huber *et al.*, "Absolutely maximally entangled states of seven qubits do not exist," *Physical Review Letters*, vol. 118.

Alba Cervera-Lierta recieved her BSc degree in Physics from Universitat de Barcelona (UB), Spain in 2014. The following year, she completed her MSc degree in Astrophysics, Particle Physics and Cosmology from UB. She started her PhD in 2015 in Institut de Ciències del Cosmos and since 2017 she is also a technical researcher in Quantic group in BSC.

Solving Parametric Wave Propagation Models with Domain Decomposed Reduced Order Methods

Prattya Datta*†, David Modesto*†, Josep De La Puente*, José María Cela*†

*Barcelona Supercomputing Center, Barcelona (BSC), Spain
†Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
E-mail: {prattya.datta, david.modesto, josep.delapuente, josem.cela}@bsc.es

Keywords—PGD, PGD-DDM, Helmholtz model.

I. EXTENDED ABSTRACT

In the field of geophysics seismic imaging is performed to understand the sub-surface layers of the earth for commercial and research purposes. A typical seismic image is developed by solving the Helmholtz model iteratively over a wide range of frequencies. The seismic image consists of solving the Helmholtz model in a 10-15 km long spatial dimensions for a wide range of parameters as a forward modelling. Thus, extensive computations are needed. Reduced order techniques especially proper generalized decomposition (PGD) could be employed to accelerate this iterative models using surrogate models with real-time online computation phase. However, PGD computations require an offline phase which is highly penalized for large frequencies. For high frequencies this results in difficulty in implementing realistic geophysical models by the PGD strategy. To circumvent this issue a domain decomposition (DDM) technique combined with the PGD is proposed. In the present work DDM-PGD strategy is imposed on the spatial dimension of the model. PGD-DDM strategy computes local spatial surrogate models which result in convergences locally. Global convergence is ensured by choice of the DDM strategy.

The model is defined in a 2D heterogeneous velocity medium. The DDM-PGD has transmission conditions on the boundaries of the sub-domains based on perfectly matched layers (PML) which avoid artificial reflections and match incoming and outgoing waves. Specifically, transmission conditions are defined by overlapping layers of PML. Overlapping PML type transmission conditions are chosen because they are known to have better convergences in the global residual updates for full order models, which are also independent of the number of sub-domains used. PML transmission conditions absorb artificial reflections completely and have better conformity with the incoming and the outgoing waves. Global surrogate model is built by adding all the spatial local surrogate models. This process can be ensembled either in serial or parallel approach if HPC resources are available.

A. Problem Statement

The parametric Helmholtz model (2D in space) is defined as;

$$\nabla \cdot (\mathbf{P} \cdot \nabla u) + k^2 S_x S_y u = f \tag{1}$$

where,

$$u = u(\mathbf{x}, \omega, s, c), f = f(\mathbf{x}, \omega, s, c)$$

Here, s is the source position, c is the speed of propagation in the medium and ω is the frequency. f is the source term and k is the wave number defined as $k = \frac{\omega}{c}$. $\mathbf{P}$ is the anisotropy coefficient matrix which controls the absorption of the outgoing rays and signify the PML layer. In the present work, the computational domain is assumed to be a rectangle which is truncated by the PML. The sub-domains are of the identical shape of the full domain. The boundary conditions of the sub-domains are of PML type as well which are overlapped and act as transmission conditions.

B. PGD-DDM

To generate the PGD, the Helmholtz potential function is defined in terms of product of explicit separable functions wherein the parameters are also considered as a separable function. The separated representation form for the potential function u is defined as follows;

$$u(\mathbf{x}, \omega) \approx \sum_{i=1}^{N} F_1^i(\mathbf{x}) F_2^i(\omega) \tag{2}$$

The above equation has been reduced to multiplication of lower dimensional functions. It is noted that adding extra parameters do not increase the dimensions to the powers of N, but the further dimensions are added to the final tensor in the form of extended product of low dimensional functions. Solving (1) with the approximation of (2) requires a greedy strategy, which is;

$$u \approx u_{PGD} = \underbrace{F_1(\mathbf{x}) F_2(\omega)}_{\text{Nth mode}} + \sum_{j}^{N-1} \left(F_1^j(\mathbf{x}) F_2^j(\omega) \right)$$

It is assumed that the function u is defined by N modes out of which N-1 modes are known and our objective is to find the Nth to completely define u_{PGD}. To solve the above equation, previous terms are assumed to be known and the rest of the new terms are calculated iteratively enriching the final solution (this method is also called as fixed-point or alternating directions algorithm) until a global convergence is obtained, which is defined as:

$$\text{Convergence} = \frac{\left\| \prod_{i=1}^{2} F_i^m - \prod_{i=1}^{2} F_i^{m-1} \right\|_{\mathcal{L}^2(\Omega \times I_\omega)}^2}{\left\| \prod_{i=1}^{2} F_i^{m-1} \right\|_{\mathcal{L}^2(\Omega \times I_\omega)}^2},$$

in an appropriate norm. With the above PGD strategy we incorporate the DDM overlapping strategy. In the case of a Helmholtz model with first order absorbing boundary conditions, following scheme is used as DDM. Assuming appropriate overlap, the sub-domain equations are solved until the resulting full domain solution have the desired error wrt. original solution,

$$-\Delta u^{j,i} - k_i^2 u^{j,i} = f(\mathbf{x}), \quad \forall(\mathbf{x}) \in \Omega_\mathbf{x}^i$$
$$(\nabla \cdot \vec{n}) = iku^{j,i}, \quad \forall(\mathbf{x}) \in \partial\Omega_x^i/\Gamma^i$$

$u^{j,i} = u^{j-1,i-1} \in \partial\Omega_\mathbf{x}^{i-1} \bigcap \partial\Omega_\mathbf{x}^i$ compatibility condition

$u^{j,i} = 0 \in \Gamma^{j,i} \bigcap \partial\Omega_\mathbf{x}^i$ a homogenous dirichlet b.c is imposed
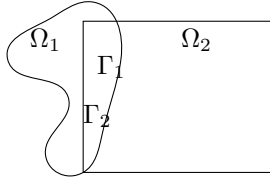
To increase convergences optimized Schwarz methods are used in place of regular Schwarz methods. [1][2] have shown much improved convergences when using optimized Schwarz. The following figures show the convergence for the multiplicative optimized Schwarz. It shows its utility to be used for DDM-PGD

(a) optimized Schwarz DDM for 200 waves

(b) Convergence of optimized Schwarz ddm for 200 waves
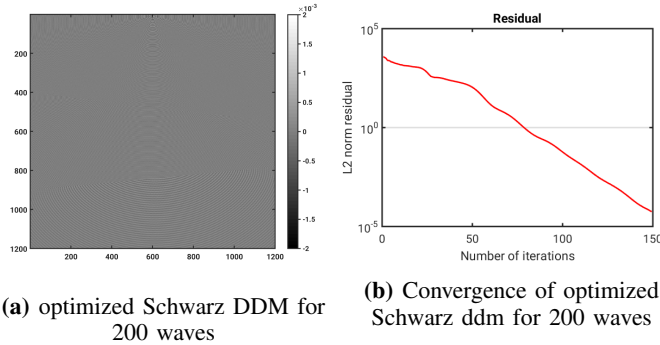
Fig. 1: Full order Helmholtz solved by Schwarz DDM

C. Numerical Experiments

In this section, examples of 2D spatial Helmholtz model is shown. The discrete formulation is solved using 9-point rotated finite difference scheme. The results are shown for a with the proposed strategy.For the PGD formulation the domain $\mathbf{x} \in [0,1]^2$ and $\omega \in [6,20]$. The model shows good convergence of PGD algorithm in each sub-domains. Fig(2) shows the expected DDM-PGD result is good representation of true solution. The model has transverse layers of heterogeneity.
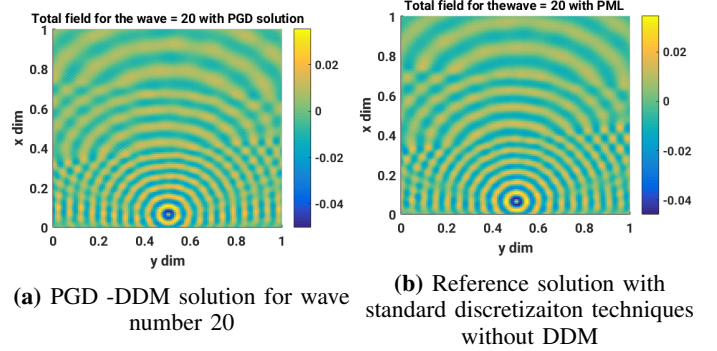

(a) PGD -DDM solution for wave number 20

(b) Reference solution with standard discretizaiton techniques without DDM

Fig. 2: Reference and PGD-DDM figures

(a) Convergence of domain numbered 1
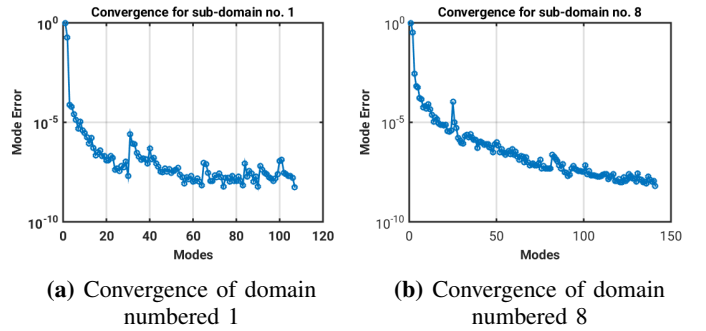
(b) Convergence of domain numbered 8

Fig. 3: Convergences for sub-domains chosen at random out of 16

The convergences are good and in line with the expectations as shown in fig(3). The convergences of each sub-domain takes less outer iteration of modes of PGD than the PGD over the complete domain.

D. Conclusion

In this study,we have presented a novel method to incorporate a DDM strategy to PGD algorithm for a wave problem like Helmholtz model. We believe this work will help in developing faster solution for efficient scalable algorithms for PGD of large scale wave problems

II. ACKNOWLEDGMENT

REFERENCES

[1] A. St-Cyr *et al.*, "Optimized multiplicative, additive and restricted additive schwarz preconditioning," *SIAM journal on scientific computing*, vol. 29, no. 6, pp. 2402–2425, 2007.

[2] M. Gander, "Schwarz methods over the course of time*," *Electronic transactions on numerical analysis*, vol. 31, pp. 228–255, 2008.
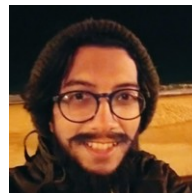
Prattya Datta received his Bachelor of Engineering degree from Jadavpur University, India. He subsequently received his Master in Technology from Indian Institute of Technology Madras and Msc from University Paris Diderot 7. He is currently employed as a PhD in Barcelona Supercomputing Center with Department of CASE.

Towards a low dissipation FE scheme for scale resolving turbulent compressible flows

Lucas Gasparino F. da Silva*†, Oriol Lehmkuhl*, Daniel Mira Martinez*

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {lucas.gasparino, oriol.lehmkuhl, daniel.mira}@bsc.es

Keywords—Compressible flow, CFD, LES, FEM, Numerical Simulation, Hyperbolic problems.

I. Extended Abstract

This work focuses on developing a highly accurate Finite Element numerical scheme for Large-Eddy simulation of turbulent compressible flows, where low numerical dissipation is required. The proposition is to develop a low-dissipation stabilizing term into an algorithm, implement it on a readily-available CFD code and test results to assess performance and suitability.

A. Discretisation basics

In general, numerical solution of partial differential equations like the Navier-Stokes set governing fluid flow involves some form of space-time discretisation procedure, where the physical domain is divided into a set of points called grid, or a set of connected subdomains referred to as *elements* or *volumes*. If the problem is transient, then the continuous time must also be subdivided and discretised to allow for a local temporal solution. In finite elements, the simplest form of discretisation procedure is the Continuous Galerkin method, which yields a central-like, 2nd order discrete solution to the problem.

For convection dominated problems (those containing a transport-like term), direct application of a standard FE technique often results in oscillatory, or fully unstable, solutions, and therefore some form of stabilization is required, generally referred to as artificial diffusion, as the added terms have this exact behavior. Common stabilized methods are the Streamline Upwind Petrov Galerkin method (SUPG) and its variations, which achieve a stable solution by effectively reducing discretization order, ensuring the scheme remains monotonicity preserving.

Compared to a standard Galerkin discretization procedure, classical stabilized methods such as SUPG and nth Taylor-Galerkin are known to introduce strong diffusion terms into the discrete Navier-Stokes equations, compatible with their low-order nature. Although this has minor influence over RANS and URANS models, Large-Eddy and Direct-Numerical simulations (LES and DNS) are heavily affected by it, and one is forced to move towards high-order central schemes, which have narrow stability ranges. In the compressible range, the hyperbolic nature of the problem exacerbates this stability issues, and if shock waves form, or any other form of discontinuity exists, the method fails completely. As Finite Volume schemes solve this issue by means of Riemann solvers on the shock and use of flux limiters that reduce the scheme's order locally, these are quite complex to implement in Finite Elements, and thus a different approach is necessary.

B. Entropy Viscosity and uniqueness of solutions

Research on stabilization of general hyperbolic problems has led to the development of a low-dissipation nonlinear term named *Entropy Viscosity*, which, when used in a Finite Elements context, behaves in a similar way to a FV flux-limiter, i.e., ensuring that 1st order solutions are employed only at high gradient regions, such as a shock or interface.

This method is based on the answer of a purely mathematical question: to find a unique solution for a general hyperbolic equation. As the strong form of this problem is not approachable, a *viscous solution* is sought that satisfies a weak form of the reformulated conservation law. The true solution is then found by assuming the limiting case where viscosity vanishes. Such viscous solution is referred to as the *entropy solution* of the hyperbolic problem, as it must satisfy *Kružkov's entropy condition*.

The mathematical theory described above lends itself well to a Finite Element discretisation procedure, as Kružkov's entropy condition is already stated in an ideal weak form. Moreover, it was found that the theory developed for 1D scalar cases can be directly extended to multi-dimensional hyperbolic (and quasi-hyperbolic) systems like the Euler and Navier Stokes equations, even if a nonlinear source term exists. This indicates a possible candidate for a low-dissipation algorithm theory.

C. Research steps

The first part of this research project has focused on adapting the generic concept developed on hyperbolic equations to the 1D Euler system describing inviscid compressible flows, as well as studying in practice how well it worked when compared to other approaches. Following previous work already developed by Guermond *et. al.* led to a unique implementation in the FE context, which has shown remarkably good performance overall. It is worth mentioning that this method lends itself well to explicit temporal discretization, and was only explored in this context. One interesting, and quite useful, property of the scheme is its capability for handling nonlinear source terms, such as the ones appearing in combustion problems.

D. Conclusion

The initial success of this study encourages extending the method to 2D and 3D applications of viscous flows, where turbulence is allowed to develop in the latter case. Applications to be tested range from thermally affected incompressible laminar flows to LES in subsonic regime and, should these work as expected, chemically reacting and transonic cases with weak shock formation. This later part is to be coupled with the in-house code Alya, a Finite-Elements multiphysics software. Should this research prove successful, the result will be a big step forward to its CFD capabilities in compressible turbulent ranges.

II. ACKNOWLEDGMENT
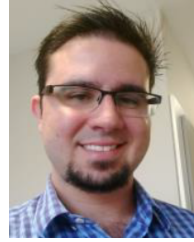
Lucas Gasparino received his BSc degree in Mechanical Engineering from North Universidade Paulista (UNIP), Sao Paulo, Brazil, in January 2016. He completed his MSc degree in Numerical Methods and Finite Elements in Engineering Mechanics from Swansea University, UK in September 2017. Since 2018, he has been with the Computational Fluid Dynamics group of Barcelona Supercomputing Center (BSC) as a PhD student in the CASE department. His experiences include Finite Element modeling and analysis of structures and structural components under static and dynamic conditions, as well as computational analysis of complex flows.

Exploration of architectural parameters for future HPC systems

Constantino Gomez*†, Francesc Martinez*†, Adria Armejach*†, Marc Casas*†, Filippo Mantovani*†, Miquel Moreto*

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {constantino.gomez, francesc.martinez, adria.armejach, marc.casas, filippo.mantovani, miquel.moreto}@bsc.es

Keywords—Co-Design, Large-scale simulations, High-performance computing.

I. EXTENDED ABSTRACT

Trends in High Performance Computing (HPC) systems are shifting. The use of commodity server processors as the main option to design these systems is moving towards a more specialized landscape. Processor trends are evolving in several directions, such as, leaner core designs [1], larger core counts per socket [2], wide vector units [3], or with integrated memory like high-bandwidth memory (HBM) modules via silicon interposer technologies [4].

In our work, we undertake a design space exploration study that considers the most relevant design trends we are observing today in HPC systems. To perform this study, we follow a recently introduced multi-level simulation methodology (MUSA) [5]. MUSA enables fast and accurate performance estimations and takes into account inter-node communication, node-level architecture, and system software interactions. Through our extensive design space exploration, we provide hardware and software co-design recommendations for next-generation large-scale HPC systems.

A. Co-Design opportunities

The design space for next-generation HPC machines is expanding. First, the trend to use commodity server processors as the common choice is changing towards processors with leaner core designs that feature different microarchitectural characteristics. For example, Cray has already deployed Isambard [6], a system with 10,000+ Armv8 cores; and now supports ARM-based processors (including the Cavium ThunderX2) across their main product line. Second, vector architectures with larger lengths than the ones employed in recent years are starting to be considered again. In this regard, Arm recently introduced the Scalable Vector Extensions (SVE) that support up to 2,048 bit vectors and per-lane predication. Third, several memory technologies are starting to appear in the HPC domain, for example: die-stacked DRAM like the one employed in Knights Landing [7], or High-Bandwidth Memory (HBM) already used in a number of GPUs.

The advent of these trends and technologies leads to a large design space for next-generation HPC machines that needs to be carefully considered. There is a clear opportunity to co-design hardware and software by mapping application requirements to the available hardware ecosystem that these trends are opening. In addition, the ability to predict and fine-tune application performance for selected hardware designs that are deemed of interest is of paramount importance to system architects.

B. Parameter exploration

After reviewing the HPC systems landscape, we select a set of important compute node features in current and upcoming HPC architectures. These features expose relevant energy and performance trade-offs when considering different HPC workloads. We focus our exploration on six features: number of cores in a socket, out-of-order (OoO) capabilities of the core, memory technology, floating-point unit (FPU) vector width, CPU frequency and cache size. Additionally, to do our simulations we select five relevant hybrid (MPI+OpenMP) HPC applications: HYDRO, SP-MZ, BT-MZ, SPECFEM3D and LULESH.

Table I shows a detailed list of all the parameters and values we explore and the names (labels) we will use to refer to them.
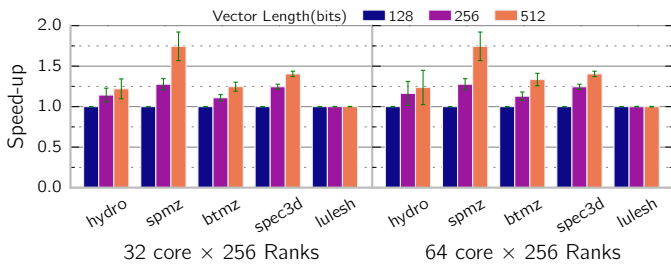
L3:L2-caches	Size / associativity / latency	
Label	**L3**	**L2**
32M:256KB	32MB / 16 / 68	256kB / 8 / 9
64M:512KB	64MB / 16 / 70	512kB / 16 / 11
96M:1MB	96MB / 16 / 72	1MB / 16 / 13

Core OoO Label	ROB	Issue& commit	Store buffer	#ALU/ #FPU	IRF/ FRF
low-end	40	2	20	1 / 3	30/ 50
medium	180	4	100	3 / 3	130 / 70
high	224	6	120	4 / 3	180 / 100
aggressive	300	8	150	5 / 4	210 / 120

Other param.	Values
Frequency [GHz]	1.5, 2.0, 2.5, 3.0
Vector width [bits]	128, 256, 512
Memory [DDR4-2333]	4-channel, 8-channel
Number of Cores	1, 32, 64

TABLE I. SIMULATION ARCHITECTURAL PARAMETERS AND VALUES USED IN OUR DESIGN SPACE EXPLORATION INCLUDING: CACHE SIZE, ASSOCIATIVITY AND LATENCY; AND OoO DETAILS LIKE REORDER BUFFER (ROB) AND INTEGER/FLOAT REGISTER FILE (RF).

Fig. 1. Average performance speedup increasing FPU width up to 512-bits. Normalized to 128-bit configurations.

C. Results

Figure 1 summarizes the performance-energy trade-off when we increase the vector Floating Point (FP) registers used for SIMD operations in each core. Results for 32 and 64 core configurations are very similar. Excluding LULESH, wider 512-bit FP units yield 20% (HYDRO) to 75% (SP-MZ) application performance speed-up; 40% on average.
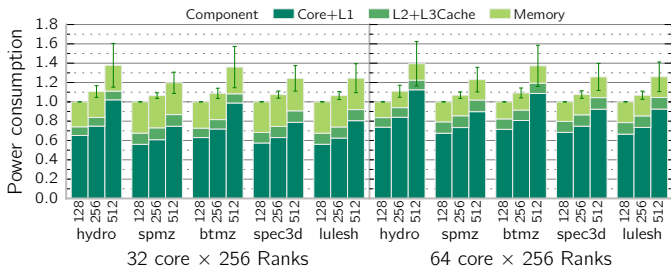

Fig. 2. Average power consumption increasing FPU width up to 512-bits. Normalized to 128-bit configurations.

In Figure 2, we see that using 512-bit vector width translates into an average power increment across applications of 60% with respect to 128-bit units in each core. As expected, the core power consumption is relatively larger in compute-intensive applications like HYDRO and BTMZ than in memory bound counterparts.
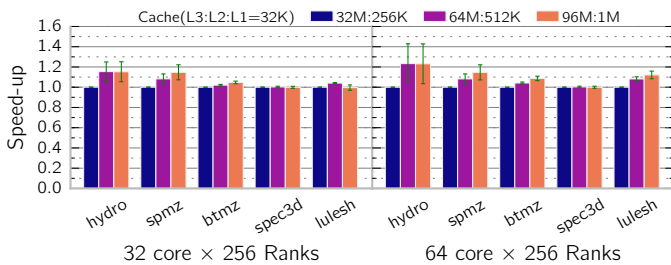
Fig. 3. Average performance speedup varying L3- and L2-cache parameters. Normalized to 32MB:256KB cache configs.

Figure 3 shows how only modifying L2- and L3-cache sizes affects performance in our simulations; at 64 cores, upgrading to a cache configuration with 96MB:1MB (1.5MB:1MB per core) results in an 11% average speedup across applications.

Taking into account these observations, we simulate parallel executions of SPMZ considering architectures with increasing SIMD widths of 1024- (*Vector+* configuration) and
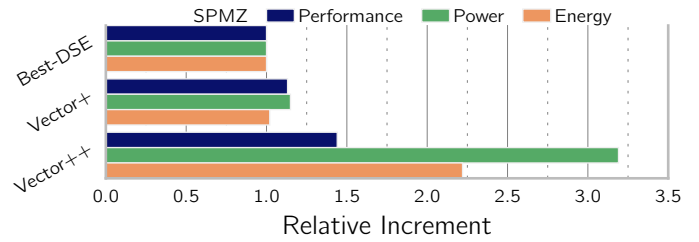

Fig. 4. Performance, power and energy-to-solution of a specific configurations targeting spmz.

2048-bits (*Vector++* configuration) while keeping the rest of the architectural features settings that give the best possible performance-power tradeoff.

D. Conclusion

In this study, we look at speedup and energy consumption exploring the design space (i.e., changing SIMD width, number of cores, and type of cores), and we provide architectural recommendations that can be used as hardware and software co-design guidelines targeting specific applications.

II. Acknowledgment

References

[1] F. Haohuan *et al.*, "The sunway taihulight supercomputer: system and applications," *SCIENCE CHINA Information Sciences*, vol. 59, no. 7, 2016.

[2] "Thunderx2 arm processors." [Online]. Available: cavium.com/product-thunderx2-arm-processors.html

[3] "SX-Aurora TSUBASA Architecture." [Online]. Available: nec.com/en/global/solutions/hpc/sx/architecture.html?

[4] "Fujitsu reveals details of processor that will power Post-K supercomputer." [Online]. Available: top500.org/news/fujitsu-reveals-details-of-processor-that-will-power-post-k-supercomputer/

[5] T. Grass *et al.*, "MUSA: a multi-level simulation approach for next-generation HPC machines," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2016*, pp. 526–537.

[6] "GW4 Isambard." [Online]. Available: gw4.ac.uk/isambard/

[7] A. Sodani, "Knights landing: 2nd generation intel xeon phi processor," in *2015 IEEE Hot Chips 27 Symposium (HCS)*, 2015, pp. 1–24.
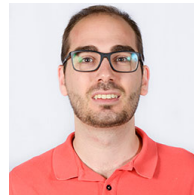
Constantino Gómez is a third year Ph.D student at the Barcelona Supercomputing Center. He received the BSc and MSc degrees in Computer Science from the Universitat Politcnica de Catalunya (UPC) in 2014 and 2016. He has been involved as a researcher in the Mont-Blanc european project series since 2014. His research interests include simulation tools, emerging memory technologies and co-design for future massively parallel systems.

Structural and Dynamics Analysis of Pyruvate Kinase from Erythrocytes: Implications in Pathology

Luis Jordà[#1], Josep Lluís Gelpí[#*2]

#*Barcelona Supercomputing Center (BSC)*

University of Barcelona (UB), Department of Biochemistry and Molecular Biology, 08028-Barcelona, Spain

[1]luis.jorda@bsc.es, [2]gelpi@ub.edu

Keywords —— **Pathogenicity prediction, pathogenic variant, protein dynamics, molecular dynamics, pyruvate kinase.**

EXTENDED ABSTRACT

Our current study revolves around the dynamic characterization of the human erythrocyte pyruvate kinase (R-PYK). The deficiency of this protein is a common cause of nonspherocytic hemolytic anemia, a rare, autosomal recessive disease. We plan to perform a comprehensive set of molecular dynamics simulations of both the wildtype (WT) and mutant variants of R-PYK in different conditions, in order to explore the dynamic behavior of the enzyme, describe the function and allosteric mechanism in terms of its dynamics fingerprint and identify altered dynamic behavior of the known pathogenic variants of the enzyme.

A. Introduction

Amino acid substitutions are implicated in a wide range of human diseases, as they are often the direct cause for a protein to partially or totally lose its function [1]. With the rise of high-throughput sequencing techniques, a very frequent clinical procedure consists in analyzing the genotype of patients who suffer from a disease to confirm or deny the presence of a mutation in the involved genes. Distinguishing such mutations from polymorphisms without significant effect on human health is a necessary step in understanding the etiology of mutation-related diseases [1]. Amino acidic replacements can affect to a different extent the stability, catalytic efficiency and regulatory properties of proteins. However, in the majority of the cases there is no apparent relationship between the nature and location of the replaced amino acid and the type of molecular perturbation [2]. This fact makes predicting the functional effects of missense variants in proteins a challenging task.

This field of study has been addressed in the recent years, and a considerable amount of bioinformatics tools to tackle the problem have been developed. However, while most state-of-the-art pathogenicity predictors rely on sequence-based features of proteins to identify damaging variants [3], structural and dynamic features have been largely neglected, mainly due to the difficulties and computational cost of their obtainment. Although molecular dynamics (MD) simulations have been traditionally employed to study the impact of mutations on the dynamic behavior of proteins, these studies are generally based on qualitative terms, and we still lack techniques to better quantify these dynamic effects. Our research line aims to improve pathogenicity prediction algorithms by establishing a protocol or guidelines that consider the role of dynamics features in the functional behavior of protein variants.

B. Use case

To this end, we are studying the particular case of the human erythrocyte pyruvate kinase (R-PYK). This protein is one of the most widely studied enzymes throughout the history of biochemistry, due to its major role in the regulation of glycolysis. It catalyzes the irreversible conversion of phosphoenolpyruvate (PEP) to pyruvate, generating an ATP molecule in the process. The enzyme needs cofactors (K^+ and Mg^{2+}) for proper catalytic activity, and is allosterically regulated by fructose-1,6-biphosphate (FBP), an activator of its catalytic efficiency. Pyruvate kinase has a highly conserved architecture throughout evolution (Fig. 1) [2].
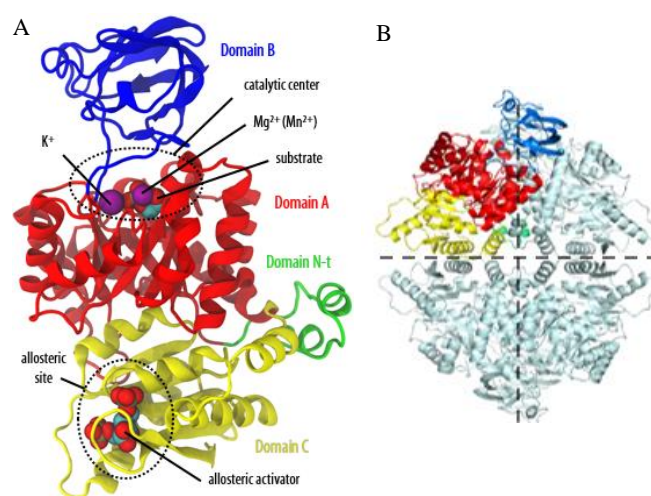

Fig. 1 Tertiary and quaternary structure of R-PYK. The different domains of the protein: N-terminal, A, B and C are colored green, red, blue and yellow respectively. A) View of a monomer of R-PYK and bound cofactors, substrate and allosteric activator. B) View of the tetramer of R-PYK. The two symmetry axes are shown with a dashed line. Only one subunit is colored as indicated above (the rest are colored pale cyan). Images generated with the software VMD.

Hundreds of missense mutations in various sites of R-PYK have been clinically associated to a disease called hereditary nonspherocytic hemolytic anemia (5 cases per 10,000 individuals in Europe) [4]. The main aim is to go deeper into the coupling between structure and functions of this protein, as well as to provide information about its dynamic properties for both the WT form and some of its mutant variants of known pathological consequences.

C. Preliminary results

So far, we have started to characterize the main motions occurring in the WT enzyme, as well as some potential alterations of these patterns in the mutated forms of the protein. These preliminary descriptions are the result of the analysis of 100 ns long MD simulations run on both the WT (5 replicas) and 62 relevant pathogenic and neutral variants extracted from literature and clinical sources. The amino acidic substitutions affect various regions of the protein, such as the active and allosteric sites, subunit interfaces and hydrophobic cores.

A thorough examination of the generated trajectories allowed us to understand the main concerted motions that govern the dynamics of the enzyme, in accordance with the

previously described features of its biological function and mechanism.

D. Future development

In the current stage of the project we are expanding the amount of MD simulations to cover a more comprehensive representation of the biologically meaningful conditions of the enzyme. Until now, the simulations consisted of the apoenzyme (only protein residues), therefore, we will establish several additional conditions of the enzyme bound to its cofactors (K^+ and Mg^{2+}), substrates (PEP and ADP) and the allosteric activator (FBP), in different combinations, and both for the WT and the mutant variants. To achieve this, we will use quantum mechanics (QM) approaches to model the bound ligands, prior to the MD simulations. The intended production MD simulation length of each variant will be of 200 ns. Through the analysis of all the generated data, we aim to establish a model to better explain the potential dynamic and/or structural alterations of each amino acid replacement in contrast with the WT behavior.

Later stages of the project will include, at least, generating a set of metrics (structural and derived from essential dynamics analyses) with discriminative power, training an R-PYK specific machine learning model to predict pathogenicity, validating it with more R-PYK variants and testing it with more proteins of similar features.

ACKNOWLEDGEMENTS

References

[1] C. Feinauer and M. Weigt, "Context-Aware Prediction of Pathogenicity of Missense Mutations Involved in Human Disease", 1–21, 2017.

[2] G. Valentini *et al.*, "Structure and function of human erythrocyte pyruvate kinase: Molecular basis of nonspherocytic hemolytic anemia", J. Biol. Chem., 277, 23807–23814, 2002.

[3] S.E. Flanagan *et al.*, "Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations", Genet. Test. Mol. Biomarkers, 14, 533–537, 2010.

[4] L. Montllor *et al.*, "Red cell pyruvate kinase deficiency in Spain: A study of 15 cases", Med. Clínica (English Ed.), 148, 23-27, 2017.

Author biography

Luis Jordà was born in Barcelona, Spain, in 1993. He received the BSc in Biochemistry from the University of Barcelona (UB), in 2016, and later the MSc in Bioinformatics for Health Sciences, from the University Pompeu Fabra (UPF), in 2018, in Barcelona, Spain.

He is currently pursuing his PhD in Biomedicine (UB) and performing his research at the facilities of the Barcelona Supercomputing Center (BSC), under the supervision of Prof. Josep Lluís Gelpí. He is also an active collaborator of the Molecular Modeling and Bioinformatics (MMB) group from the Institute for Research in Biomedicine (IRB) of Barcelona. His current research interests include structural bioinformatics, protein dynamics and modeling of biomolecules.

How much soil dust aerosol is man-made?

Martina Klose[#1], Carlos Perez García-Pando[#2]

#*Earth Sciences Department, Barcelona Supercomputing Center (BSC), Barcelona, Spain*
{[1]martina.klose, [2]carlos.perez}@bsc.es

Keywords── **mineral dust, global modeling, anthropogenic dust**

EXTENDED ABSTRACT

A. Introduction

Large parts of the Earth's land surface have undergone significant modification by humans due to, for example, urbanization and agriculture. Anthropogenic changes in land use due to cultivation and grazing can enhance the emission of soil mineral dust, the most abundant aerosol in mass originating from land sources, and thereby affect weather and climate.

The contribution of anthropogenic sources to the global soil dust load has been under debate over more than two decades with estimates ranging from 10 – 50% [e.g. 1,2,3,4]. Main reasons for this large uncertainty are (1) deficits in the representation of small-scale anthropogenic dust sources (cropland and pasture); (2) a lack of data available to constrain the global dust load; (3) deficits in the model representation of parameters affecting dust emission as well as of the dust emission process itself. Using a high-resolution ($0.1° \times 0.1°$) satellite estimate of atmospheric column dust load for dust source identification and land use maps for source attribution, a recent estimate suggests that 25% of global dust emissions originate from anthropogenic sources [5,6].

B. Objectives and Methods

Here we hypothesize that a combination of the recent advances on source identification and attribution with state-of-the-art integrated numerical modeling and a diverse set of global dust observations will help to better address the following core questions:

- What are the relative contributions of natural and anthropogenic sources to global dust emissions depending on land use classification?
- How large is the uncertainty of natural and anthropogenic dust emissions?
- What are the key processes affecting this uncertainty?

We use NMMB-MONARCH, the Multiscale Online Nonhydrostatic AtmospheRe CHemistry model [7,8], to conduct multiple global model simulations. We thoroughly evaluate and constrain the model results based on measurements of dust concentration, deposition, and optical depth to obtain a model best estimate and to quantify the global natural and anthropogenic emission and deposition along with their uncertainty. We discuss the challenges of constraining the anthropogenic fraction of dust and identify model deficits that we are going to address in the future.

C. Modeling framework

NMMB-MONARCH has the capability to simulate the atmosphere including atmospheric constituents such as aerosols or trace gases. Parallelization of NMMB-MONARCH follows a subdomain basis approach, i.e. the model domain into horizontal tiles. Since the global latitude-longitude grid deforms toward the pole region, the nodes closer to the poles perform additional filtering using Fast Fourier Transform (FFT) to maintain stable integration using a time step of reasonable size. MONARCH is a coupled model constructed over the Earth System Modelling Framework (ESMF) coupling library. This implies that in between the execution of each module (dynamics, physics, chemistry, aerosol), the model executes a coupling step to exchange information. The numerical methods employed within the model are: the Adams-Bashford scheme for horizontal advection, the Crank-Nicholson scheme for vertical advection tendencies, the forward-backward scheme for horizontally propagating fast waves, and an implicit scheme for vertically propagating sound waves. The I/O of the system uses dedicated writing nodes, resulting in a partitioning between computational and I/O nodes.

To run the model, we use the workflow manager *autosubmit* [9] together with the MONARCH-specific extension *auto-monarch* (BSC-ES CES, with acknowledgment), which facilitates easy management of the model runs as well as pre- and post-processing on high-performance computing platforms, such as MareNostrum 4.

D. Preliminary results

Numerical experiments examining the role of vegetation cover, land-use (cropland, pasture, and rangeland), and dust emission parameterization were conducted to investigate the anthropogenic emission fraction and to quantify its uncertainty. Our preliminary results suggest that the global anthropogenic emission fraction is about 10%, i.e. smaller than previously thought when using similar land-use criteria (cropland and pasture) to define anthropogenic sources, but an updated data set. When including rangeland as anthropogenic dust source, the anthropogenic contribution to the global dust cycle increases to about 40% and show a much stronger seasonal and spatial variability (not shown).
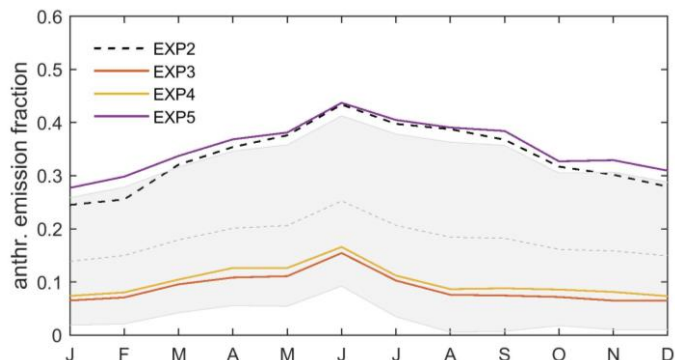

Fig. 1 Global anthropogenic soil dust emission fraction per month in 2012 based on different numerical experiments (colored solid lines) compared to a reference scenario (black dashed line) obtained using NMMB-MONARCH.

E. Summary and Conclusions

Our results demonstrate that anthropogenic sources contribute to the global dust loading and that soil mineral dust zcannot be considered only as a pure natural aerosol. Using

different dust emission parameterizations, land-use scenarios, and land-surface representations, we quantify the uncertainty related to the fraction of anthropogenic emissions. We will further refine our estimate of natural and anthropogenic dust emissions in the future, by (a) conducting higher-resolution global model runs, (b) implement a representation of moist-convective dust storms, and (c) expand the use of observational constraints to calibrate and evaluate modeled dust loadings.

F. ACKNOWLEDGEMENTS

References

[1] I. Tegen and I. Fung (1995), Modeling of mineral dust in the atmosphere: Sources, transport, and optical thickness, *J. Geophys. Res.,* 100(D9), 18707-18726, doi:10.1029/95JD02051.

[2] I. N. Sokolik and O. B. Toon (1996), Direct radiative forcing by anthropogenic airborne mineral aerosols, *Nature*, 381, 681-683, doi:10.1038/381681a0.

[3] I. M. Tegen, Werner, S. P. Harrison, and K. E. Kohfeld (2004), Relative importance of climate and land use in determining present and future global soil dust emission, *Geophys. Res. Lett.*, 31, L05105, doi:10.1029/2003GL019216.

[4] N. M. Mahowald, G. D. Rivera Rivera, and C. Luo (2004), Comment on "Relative importance of climate and land use in determining present and future global soil dust emission", *Geophys. Res. Lett.*, 31, L24105, doi:10.1029/2004GL021272.

[5] P. Ginoux, D. Garbuzov, and N. C. Hsu (2010), Identification of anthropogenic and natural dust sources using Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue level 2 data, *J. Geophys. Res.*, 115, D05204, oi:10.1029/2009JD012398.

[6] P. Ginoux, J. M. Prospero, T. E. Gill, N. C. Hsu, and M. Zhao (2012), Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products, *Rev. Geophys.*, 50, RG3005, doi:10.1029/2012RG000388.

[7] C. Pérez, et al. (2011), Atmospheric dust modeling from meso to global scales with the online NMMB/BSC-Dust model – Part 1: Model description, annual simulations and evaluation, *Atmos. Chem. Phys.*, 11, 13001-13027, doi:10.5194/acp-11-13001-2011.

[8] A. Badia, et al. (2017), Description and evaluation of the Multiscale Online Nonhydrostatic AtmospheRe Chemistry model (NMMB-MONARCH) version 1.0: gas-phase chemistry at global scale, *Geosci. Model Dev.*, 10, 609-638, doi:10.5194/gmd-10-609-2017.

[9] D. Manubens-Gil, J. Vegas-Regidor, C. Prodhomme, O. Mula-Valls and F. J. Doblas-Reyes, 2016, Seamless management of ensemble climate prediction experiments on HPC platforms, 2016 International Conference on High Performance Computing & Simulation (HPCS), Innsbruck, pp.895-900, doi:10.1109/HPCSim.2016.7568429

Author biography

Dr. Martina Klose received her diploma (2010) and doctoral degree (2014) in Meteorology at the University of Cologne, Germany. She obtained a two-year postdoctoral fellowship by the German Research Foundation to conduct research in the USA. In 2017, she joined BSC as a Beatriu-de-Pinós Postdoctoral Research Fellow. She was awarded a Marie-Skłodowska-Curie Individual Fellowship, which started in 2018. Her research interests include aeolian processes, their frequency, and impacts on Earth and other planets. Her research focus is to understand the physics of dust emission and to advance its parameterization in models. Her work resulted in 18 publications in peer-reviewed international journals (7 as lead author; 4 reprinted as book chapters), the participation as presenter at 21 and (co-)convener at 5 international conferences/workshops, the organization of 1 international workshop, several invited visits at renowned research institutions around the world, and an invited presentation at the American-Geophysical-Union Fall Meeting. Since 01/2018, she is Associate Editor of the international journal *Aeolian Research*.

Cyberinfrastructure programming with COMPSs

Francesc Lordan*, Daniele Lezzi*, Rosa M. Badia*

*Barcelona Supercomputing Center (BSC)

E-mail: {francesc.lordan, daniele.lezzi,rosa.m.badia}@bsc.es

Keywords—Distributed Computing, Workflow Managers, Programming Models, Edge Computing, Internet of Things, Cloud Computing

I. EXTENDED ABSTRACT

Cyberinfrastructures are compositions of computing systems, data repositories and storage systems, data collection instruments and visualization environments all linked together to improve research productivity and enable breakthroughs not otherwise possible.

As for previous computing infrastructure transformations, developers had to adopt new execution paradigms to exploit them. Distributed systems deprecated the traditional monolithic model in favor of service-oriented applications. To exploit the Cloud and offer SaaS, developers embraced the microservices model so that services could adjust the number of instances of each microservice to the current workload. Bringing down the computation from the Cloud to the Edge mitigates the network issues - latency and bandwidth - and enables new opportunities. On Fog infrastructures, computing devices can join in or leave at their own will. For dealing with such dynamicity, microservices became serverless and stateless.

IoT devices have sensors that permanently produce data. These devices can monitor this data themselves and, when a certain condition is met, trigger a response that may require heavy computation capabilities, or they can provide other devices with this information so they process it on a real-time basis. To support the former scenario, developers can turn to Function as a Service (FaaS): functions executed on the underlying platform in a serverless, stateless approach. To support the latter, developers need to code using stream processing frameworks such as Kafka streams or Spark Streams.

A. The COMPSs Programming Model

COMPSs [1], [2] is a framework that aims to ease the development of distributed applications. Its core it is a task-based programming model [3], [4] and a runtime toolkit executed along with the application which automatically detects its parallelism and orchestrates the execution of its tasks on the available computing nodes. Thus, it is able to exploit distributed, heterogeneous and highly-dynamic infrastructures while keeping the code totally unaware of the infrastructure and parallelism details.

To develop a distributed application with COMPSs, programmers code the logic of the application in a sequential, infrastructure-unaware fashion with no API invocations as if the code was to be run on a single-core computer. For the runtime system to detect the tasks composing the application, developers select a set of methods whose invocations create new asynchronous tasks by annotating them with the @task tag. In order to guarantee the sequential consistency of the code, the runtime system monitors the data values accessed by each tasks to find dependencies among them. So the runtime can better exploit the application parallelism, developers need to describe how the method operates (reads, generates or updates) on each data value by indicating its directionality (IN, OUT, INOUT, respectively).

The code snippet in Figure 1 shows a sample application with three tasks: two of them (generateReport) process their respective input files generating two Report objects that are merged in a third task – merge_report.

```
@task(file=FILE_IN, returns=Report)
def generate_report(file):
    ...
    return report;

@task(r1=INOUT, r2=IN)
def merge_report(r1, r2):
    ...

def main(files):
    final_report = None
    for file in files:
        partial_report = generateReport(file)
        if final_report:
            merge_report(final_report, partial_report)
        else:
            final_report = partial_report
```

Fig. 1. Sample COMPSs application using Python

B. Envisaged Computing Patterns

In cyberinfrastructure environments, authors envisage three possibilities that require computation. On the first one, known as *sense-process-actuate*, the infrastructure is supposed to give a proper response to an event detected on one of its sensors; for instance, turning on the AC system when temperature in a room reaches more than 27 degrees. A second approach that might require computation is stream processing. Sensors may continuosly produce data that need to be processed in real-time to produce a response or update the data visualization components. For instance, a camera could video-stream the inside of the room, and the AC system could be turned when the processing of such stream identifies typical reactions to heat. Finally, the third scenario that might require computation is batch processing. Cyberinfrastructures collect, generate and store big amounts of data. All this information can be analyzed to produce new knowledge; for instance, machine learning techniques could be applied to improve the models detecting people reactions and change the behavior of the infrastructure.

C. Runtime System Infrastructure

Unlike previous environments with a single entry point to start an application, on cyberinfrastructures any device can

detect an event and trigger a computation. Thus, the each device must be able to work in a standalone manner and run the computation in an efficient way: detecting the parallelism inherent in the computation and run the detected tasks on the computing elements embedded on the device. For doing so, each node of the infrastructure will host the execution of a COMPSs agent, a deamon process waiting for execution requests. When a the node triggers a new execution, it contacts this agent which runs the corresponding code detecting the tasks composing it and orchestrates their execution on the available computing resources.

On the one hand, the computing resources locally embedded on the same device might not be enough to host the expected computation. On the other hand, not all the devices of the infrastructure will trigger executions; indeed, even those that raise events may not compute anything temporarily. Thus, if COMPSs agents were able to only exploit the local processors, most of the computing power would remain idle. In order to better exploit the computing power of the infrastructure and shorten the execution time of the application, COMPSs agents must interact with each other and offload part of their computation. For that purpose, one COMPSs agent can call the same method that started a computation on another agent to execute a task in the remote agent. This agent will run the task code, analyze it detecting new tasks on the task code and orchestrate the execution of these tasks on other agents of the infrastructure.

For infrastructures with a small number of devices, a COMPSs agent can individually manage the execution of the tasks it has detected on any of the other COMPSs agents composing the infrastructure. However, the bigger the infrastructure is, the higher the complexity of the scheduling problem becomes. Assigning tasks to resources may become a computational load bigger than the actual computation to perform, or for very large infrastructures, it may simply not fit in resource-scarce devices.

To solve that problem, we propose to organize all the agents of the infrastructure in a hierarchic way. The topology of such organization can be represented as an acyclic graph where each node of the graph represents an agent of the infrastructure, and edges illustrate the possibility of one agent to submit tasks onto the agent represented by the other end of the edge. Thus, one agent can decide to run tasks on its local devices or offload them onto any of the other agents with whom it is directly connected on the hierarchy. In turn, the receiving agent could decide to run the task on its local computing devices or onto any of its direct neighbors. Thus, a computation triggered by any of the agents could use as many resources of the infrastructure as the computation needs to ensure that the application runs and it does so in the shortest execution time possible without adding a significant overhead due to the scheduling.

An important aspect to bear in mind when dealing with cyberinfrastructures is the potential mobility of its components. Smartphones, tablets and more complex devices attached to a battery with wireless network connectivity may be part of the system. These kind of devices may join in the infrastructure or leave it at their own will. When a device joins in, it is included in the system topology to add its resources into the pool so other agents can use them and, on the other hand, the mobile device can exploit the rest of the infrastructure to enhance its performance. When a device gets disconnected from the infrastructure, a communication link is broken and an edge of the graph is removed. At that point, the agents at both ends of the edge need to re-schedule the execution of the tasks already offloaded onto the other agent and assign them to resources to which it remains connected. Thus, any computation already started can go on and finish its execution.

Good criteria to build the topology are the stability and latency of the network. Establishing the topology according to that ensures that agents will always try to offload tasks onto nearby resources, on the fog, rather than submitting them to the Cloud achieving a higher performance. Besides, in the case of network disruption, the system could remain usable even without the cloud.

References

[1] R. M. Badia and et al., "COMP superscalar, an interoperable programming framework," *SoftwareX*, vol. 3, pp. 32–36, 12 2015. [Online]. Available: http://dx.doi.org/10.1016/j.softx.2015.10.004

[2] Workflows and Distributed Computing - Barcelona Supercomputing Center (BSC). (2018) COMP Superscalar. [Online]. Available: http://compss.bsc.es

[3] F. Lordan *et al.*, "ServiceSs: an interoperable programming framework for the Cloud," *Journal of Grid Computing*, vol. 12, no. 1, pp. 67–91, 3 2014. [Online]. Available: https://digital.csic.es/handle/10261/132141

[4] E. Tejedor *et al.*, "PyCOMPSs: Parallel computational workflows in Python," *The International Journal of High Performance Computing Applications (IJHPCA)*, vol. 31, pp. 66–82, 2017. [Online]. Available: http://dx.doi.org/10.1177/1094342015594678

Francesc Lordan obtained the Ph.D. in Computer Architecture from the Universitat Politecnica de Catalunya in 2018 after defending his Ph.D. thesis: "Programming Models for Mobile Environments". Since 2010, Francesc is part of the Workflows and Distributed Computing group of the Barcelona Supercomputing Center. His efforts have focused on the COMPSs programming model: a task-based model for developing parallel applications running on large distributed infrastructures such as clusters, supercomputers, grids and clouds. Francesc has published more than 20 articles in International conferences and journals, and he has been directly involved in the European projects mF2C, ASCETIC and OPTIMIS. He has also provided support to other collaborative projects such as Venus-C, EU-Brasil OpenBio, Transplant and the Human Brain Project. His research focuses on programming models that aim to ease the development of parallel applications by hiding the technical concerns of heterogeneous and distributed infrastructures.

TauRieL: Targeting Traveling Salesman Problem with deep reinforcement learning

Gorker Alp Malazgirt, Osman Unsal, Adrian Cristal

Barcelona Supercomputing Center, Barcelona, Spain

E-mail: {gorker.alp.malazgirt, osman.unsal, adrian.cristal}@bsc.es

Keywords—TSP, deep reinforcement learning, algorithms

I. EXTENDED ABSTRACT

We propose TauRieL [1], a novel deep reinforcement learning (DRL) based method and target Traveling Salesman Problem (TSP) since it has broad applicability in theoretical and applied sciences. TauRieL utilizes an actor-critic inspired DRL architecture that adopts ordinary feedforward nets to obtain an update vector v. Then, we use v to improve the state transition matrix from which we generate the policy. Also, the state transition matrix allows the solver to initialize from precomputed solutions such as nearest neighbors. In an online learning setting, TauRieL unifies the training and the search where it can generate near-optimal results in seconds. The input to the neural nets in the actor-critic architecture are raw 2-D inputs, and the design idea behind this decision is to keep neural nets relatively smaller than the architectures with wide embeddings with the tradeoff of omitting any distributed representations of the embeddings.

TSP has been formulated and studied widely as a combinatorial optimization problem, and approximate heuristics based methods have been proposed. However, the design of heuristics requires detailed domain-specific knowledge and tuning. In this paper, we attack this aspect and develop a general and neural net (NN)-based TSP method that can produce feasible results two orders of magnitude faster than current NN-based methods with competitive solution quality. Current methods that target TSP use recurrent neural net (RNN) based methods. Figure 1(b) presents the RNN method. The RNN based algorithms stochastically explore the design space and learn the latent space using the RNN architectures. The work proposed by Bello et al. [1] is an example of RNN based method.

A. TauRieL

We represent the problem as a Markov Decision Process (MDP). Given a graph G that consists of cities $G = \{x\}_i^n$, the objective is to find shortest tour by visiting each city. We represent the environment as a Markov Decision Process (MDP), which is a tuple $\langle S, A, P, R \rangle$. S defines a state space where each state is a city in a given TSP instance. P is a state transition probability matrix such as the probability of reaching state s_j at time $t + 1$ from s_i at time t is $P_{i,j}$. We define R as reward when agent transition from state s_i to s_j with action a. We are look for finding tours that generate higher

---

[1] A wood-elf character from Hobbit the Movie who possesses *superior pathfinding abilities*
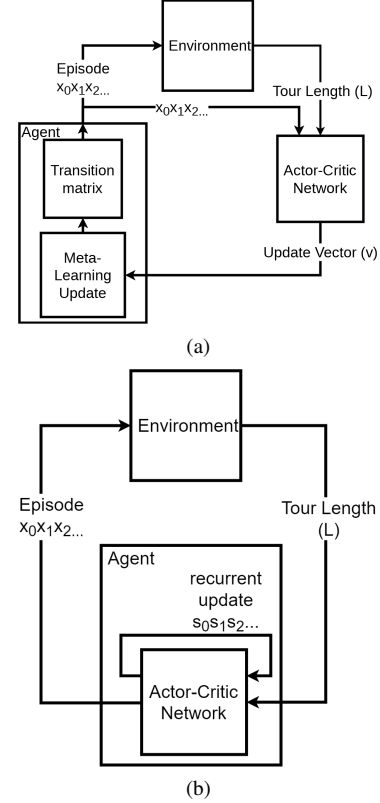

Fig. 1. High level schemas of TauRieL (a) and a state-of-the-art Actor-Critic based TSP solver using RNN [1](b)

rewards (where tour length is minimal) and assign them higher probabilities.

Our goal is to optimize the parameters $\theta_{act,cri}$ of the Actor-Critic neural nets that yield the best update vector v. We optimize the parameters of the neural net with respect to the objective i.e. the expected tour length:
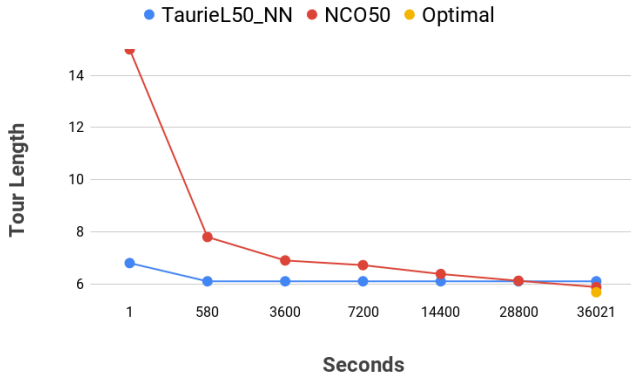
$$J(\theta_{act}|G) = \mathbb{E}_{\phi \sim p(.|s)} L(\phi \mid G) \qquad (1)$$

$L(\phi \mid G)$ is defined as the tour length of a given tour. The gradient $\nabla_{\theta_{act}}$ of the expected tour length is calculated using the REINFORCE algorithm [2] that we tailored for TSP and it is estimated by stochastic batch gradient method as:

$$\nabla_{\theta_{act}} J(\theta_{act}|G) = \frac{1}{B} \sum_{i=1}^{B} [(L(\phi \mid G) - b(G)) \\ \nabla_{\theta_{act}} log \ p(\phi \mid G)] \qquad (2)$$
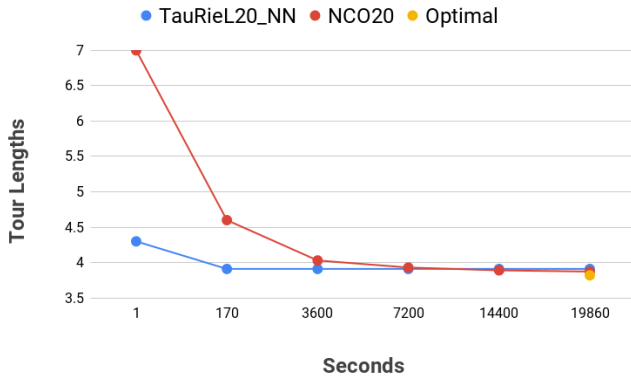
(a)

(b)

Fig. 2.   The average tour length vs training duration for 20 and 50-city instances of TauRieL and NCO [1]

N	OPTIMAL	A3	NCO	SPG	TAURIEL
10	2.87	3.07	2.88	NA	2.88
20	3.83	4.24	3.88	4.62	3.91
50	N/A	6.46	6.09	N/A	6.37

The $b(G)$ is called the baseline, which is a parametric metric that all generated tour lengths can be compared against. Hence, we use $b(G)$ to reduce large fluctuations of the tour lengths observed by the agent during the search. In this work, we select the baseline as the estimated tour length value obtained from the second neural network which is called the critic.

We update the transition matrix P with the update vector v using the update in Equation 3:

$$P_{i,j} = P_{i,j} + \epsilon \, (v_i - P_{i,j})$$
$$\forall i \; [i \in 1, \dots, n] \; and \; \exists j \; [j \in 1, \dots, n] \tag{3}$$

The actor net in the actor-critic architecture in Figure 1(a) is responsible for producing the update vector v and after each K episode the transition matrix is updated with v which aims to redirect the exploration towards shorter tour lengths.

B. Experimental Environment

We compare our results with the tour lengths obtained from the Google OR-Tools, Ptr-Net, NCO and Sinkhorn Policy Gradient [3], [1], [4]. We obtain 2.4% and 6.1% from the optimal compared to 1.4% and 3.5% which are reported by NCO[1] for 20-city and 50-city instances respectively. TauRieL outperforms A3 [5] for both 20 and 50 city instances and obtains tour lengths within 0.007 % and 2% of Ptr-Net for 20 and 50-city instances. Moreover, TauRieL outperforms an actor-critic based Sinkhorn Policy Gradient [4] in 20-cities case which is the only reported TSP size by the authors.

In Figures 2(a) and 2(b), we introduce the improvements in tour lengths with respect to the training. Specifically, we measure the training times for 20-city and 50-city instances of NCO as 19860 and 36021 seconds respectively. On the other hand, without needing any data sets, TauRieL runs in 170 seconds for 20-city and 580 seconds for 50-city instances when sample and episode step size is 250. For 20-city TSP, NCO [1] necessitates more than two hours of training in order to outperform TauRieL which can solve a 20-city instance in less than three minutes. Similarly, for 50-city, NCO needs to train at least eight hours to reach TauRieL's performance whereas TauRieL can obtain a solution in less than ten minutes. Hence, on average, training NCO below the specified duration yields poorer results than TauRieL.

C. Conclusion

In this study, we show that TauRieL generates TSP solutions two orders of magnitude faster per TSP instance as compared to state-of-the-art offline techniques with a performance impact of 3% in the worst case.

REFERENCES

[1] I. Bello, H. Pham, Q. V. Le, M. Norouzi, and S. Bengio, "Neural combinatorial optimization with reinforcement learning," *arXiv preprint arXiv:1611.09940*, 2016.
[2] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine learning*, vol. 8, no. 3-4, pp. 229–256, 1992.
[3] O. Vinyals, M. Fortunato, and N. Jaitly, "Pointer networks," in *Advances in Neural Information Processing Systems*, 2015, pp. 2692–2700.
[4] P. Emami and S. Ranka, "Learning permutations with sinkhorn policy gradient," *arXiv preprint arXiv:1805.07010*, 2018.
[5] "C++ implementation of traveling salesman problem using christofides and 2-opt," https://github.com/beckysag/traveling-salesman, 2017.

Gorker Alp Malazgirt received his B.S. in Microelectronics Engineering from Sabanci University, Turkey, his M.S. in Electrical Engineering from Lund University, Sweden and his Ph.D. from Universitat Politècnica de Catalunya, Spain. He worked at Ericsson and ST-Ericsson in Lund, Sweden as an R&D engineer. His research interests are algorithms, computer architectures and reconfigurable computing.

Accurately modeling IR spectra of astronomical interesting nanoclusters

Joan Mariñoso Guiu[#1], Stefan T. Bromley[#*2]

#Departament de Ciència de Materials i Química Física & Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

`1jmaringu11@alumnes.ub.edu`

*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.

`2s.bromley@ub.edu`

***Keywords*— Silicate nanoclusters, Interstellar dust, IR spectra**

EXTENDED ABSTRACT

Silicates are ubiquitous both terrestrially and throughout the universe, where they are often present as small particles. Nanosized silicate particles are likely to be particularly important for understanding the formation, processing and properties of cosmic dust grains. Although astronomical infrared (IR) observations and laboratory studies have revealed much about silicate dust, our knowledge of this hugely important class of nanosolids largely rests on top down comparisons with the properties of bulk silicates.

Herein, we assess the accuracy of various computational methods for obtaining IR spectra of silicate nanosized dust grains of astronomical interest, directly from the atomistic structure and their atomic motions.

A. Obtaining IR spectra

To compute the IR spectra in this work we have used molecular dynamics (MD) as well as the well-known harmonic approximation.

In order to obtain IR spectra from MD simulations, one needs to compute the dipole moment of the studied system at each step of the dynamics. With that information is just necessary to compute the dipole autocorrelation function to check which are the dipoles that repeat the most during the dynamics, and therefore the ones that will produce actives modes to the IR spectra.

Finally, is just necessary to Fourier transform it to obtain IR spectra directly based in the atomic motions.

B. Small clusters

IR spectra for a selection of small nanosilicate clusters with a range of sizes and chemical compositions are obtained using density functional theory (DFT) within the harmonic oscillator approximation. To check if anharmonicity effects play a significant role in the IR spectra of these nanoclusters, we further obtain IR spectra from finite temperature DFT-based ab initio molecular dynamics (AIMD) (**Figure 1**).

The obtained results show a generally good agreement between harmonic and MD-based IR spectra, with just a small shifting in the position in which the peaks appear that can be assigned to the fact that under AIMD is possible to capture anharmonic effects.
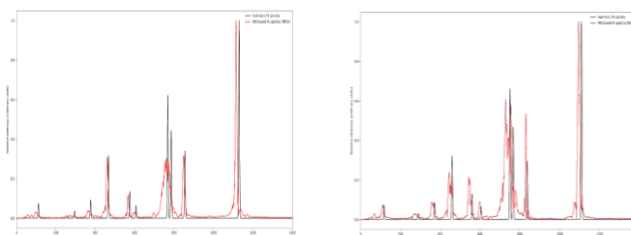

Fig. 1 Harmonic (black), and MD-based (red) IR spectra for $(MgSiO_3)_2$ using PBE0 (Left) and B3LYP (Right) functionals.

C. Larger silicate grains

In order to study the effect of temperature on the broadening of the obtained IR spectra peaks, we compute the IR spectra of a range of larger nanosilicate. In this case, less computationally costly classical MD simulations are necessary due to the large number of atoms involved. Results show once again a good agreement between harmonic and MD-based spectra. We present the harmonic broadening that better fit the MD based IR spectra for those large nanosilicate grains IR spectra at certain temperature (**Figure 2**).
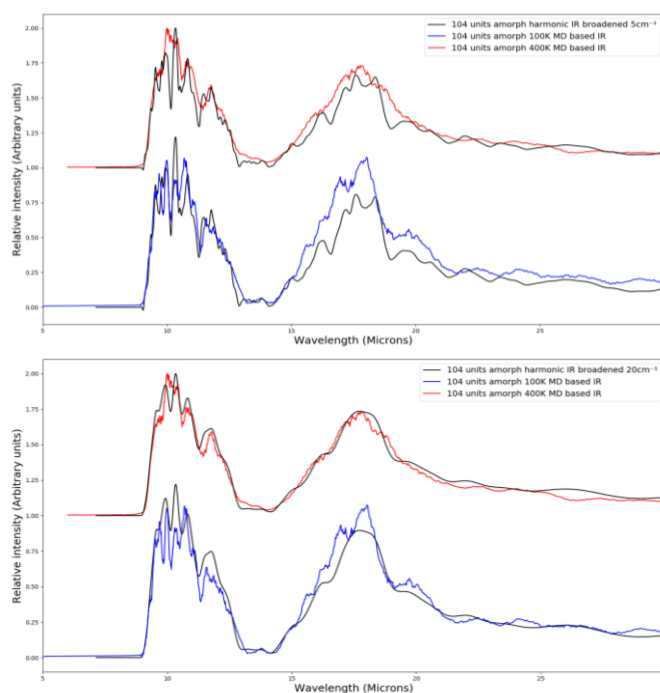

Fig. 2 Harmonic (black), and MD-based (blue and red) IR spectra for $(MgSiO_3)_{104}$

D. Conclusions

Generally, we find that although DFT-based methods are more accurate, surprisingly good IR spectra can also be obtained from classical MD calculations.

We also are able to observe that increasing the temperature produces an increase in the broadening of the peaks obtained in the MD-based IR spectra.

Overall, our work should provide a new platform for an accurate and detailed understanding of the IR spectra of nanoscale silicates which will thus assist the interpretation of experiment and observation.

Author biography
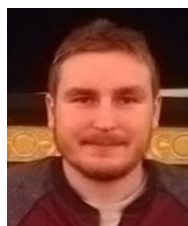
Joan Mariñoso was born in Barcelona, Spain, in 1995. He received its bachelor's degree in chemistry from Universitat de Barcelona, Spain in 2017.

Since September 2017, he has been enrolled in the Theoretical Chemistry and Computational Modelling (TCCM) master degree. Currently, he is carrying out his master thesis in the Computational Material Science Laboratory (CMSL) research group in Universitat de Barcelona, under the supervision of Stefan T. Bromley.

Multifractal characterization of seismicity:
the case of Carterbury region (New Zealand), 2000 -2018

Marisol Monterrubio-Velasco[(1)], Xavier Lana[(2)], Ramón Zúñiga[(3)], Dolors Martínez[(2)], and Josep de la Puente[(1)]

(1) Barcelona Supercomputing Center -CASE; (2) Polytechnic University of Catalonia, Physics Department ; (3) National Autonomous University of Mexico, Geosciences Center

email: marisol.monterrubio@bsc.es

Abstract: The Multifractal Detrended Fluctuation (MF-DF) algorithm is applied to measure the complexity of two time series, the inter-event hypocentral distance Δδ(t), and the inter-event time series Δτ(t) In particular, we apply this methodology to the seismic sequences produced in the Carterbury region during 18 years (2000-2018). Results indicate a clear multifractal behavior of Δδ(t) and Δτ(t). Moreover, an increses in the complexity is observed when a large event occurs . These results suggest that the MF-DF algorithm could be useful as a seismic precursor index.

I.   INTRODUCTION

The concept of multifractal modeling has been used intensively in various fields of science for characterizing measures with self- similarity. Since [1], numerous studies have characterized the self-similar (or scale-invariant) properties of a wide variety of natural phenomena by using the concepts of fractal geometry and fractal dimension [2].In the case of earthquake phenomena, several studies found that the temporal, epicentral and hypocentral, and also the energy distribution of earthquakes, have multifractal characteristics [3-6].

In this work the Multifractal Detrended Fluctuation (MF-DF) algorithm is applied to characterize multifractal scaling properties of two time series, the inter-event hypocentral distance $\Delta\delta(t)$, and the inter-event time series $\Delta\tau(t)$ from a particular seismic catalogue. These series give information of the seismic distribution in space and time.

II.   SEISMIC CATALOGUE

New Zealand is located at the border of two major tectonic plates, the Australian and Pacific plates. The slowly driving movement of these plates breaks the Earth's crust into separated blocks producing fractures, also known as faults. One example is the Alpine Fault, which it is over 600 km long and is responsible for some of the largest earthquakes in New Zealand's history.

We analize in the Carterbury region, located in the South Island, close to Christchurch city. From 2000 to 2018 the seismic network registered 15889 events with magnitude larger than 1. However, we only consider the events with magnitudes larger than 2.5, because from this value a completeness in the seismic catalogue is assured to follow the Gutenberg-Richter relation [7].

Fig. 1. Frequency-Magnitude distribution of 15889 events registered in Carterbury (New Zealand) region from 2010 to 2018. Red line indicates the magnitude of completeness 2.5
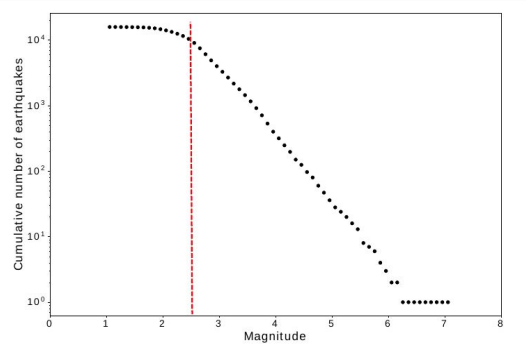
III.   METHODOLOGY

A.   Multifractal Detrended Fluctuation algorithm

The multifractal properties of nonstationary series is analysed by means of the multifractal detrended fluctuation MF-DF, analysis [8-11]. After appling the MF-DF algorithm, one can compute the singularity spectrum f(α) as:

$$f(\alpha) = q\alpha - \tau(q) \tag{1}$$

where q is the q-th order fluctuation moment computed from the MF-DF algorithm, α is the singularity strength or Hölder exponent, and $\tau(q)$ is the global scaling exponent.

Fig. 2 shows an example of the singularity spectrum for an empirical time series. Eq. (1) can be well fitted by a polynomial of second order around the position α_0 which is Hölder exponent with maximum singularity spectrum, being

$$f(\alpha) = A(\alpha - \alpha_0)^2 + B(\alpha - \alpha_0) + C \tag{2}$$

Fig. 2 Example of the singularity spectrum f(α) for an empirical time series (dotted line). Blue line is the fitting of a polynomial of second order.
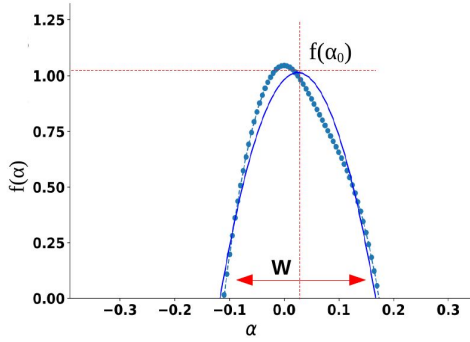
We use the complexity index, CI, proposed in [12] to quantify the degree of complexity in a time series. This value is defined as,

$$CI(Z) = \{Z - <Z>\} / \sigma(Z) \qquad (3)$$

where $Z = z(\alpha_0) + z(B') + z(W)$, W is the spectral amplitude, $B' = -B/2A$ is the asymmetry, and $<Z>$ and $\sigma(Z)$ are the mean and the standard deviation of the standardized Z respectively

B. Seismic data

We divided the total series using a moving window, in a sub-series of 1000 events length and shifted by 25 events each one. In total we consider 355 windows (or sub-series), and for each we compute their singularity spectrum and measure its multifractal behavior. Fig. 3 shows the length in time of each sub-series. The larger events are marked by dotted lines

III. RESULTS AND CONCLUSIONS

To summarize the results Fig. 4 shows the CI parameter because it is a combination of the other parameters (Eq. 3) and gives a measure of the multifractal behavior of $\Delta\delta(t)$ and $\Delta\tau(t)$ series. The results indicate evident multifractal behavior for the analyzed series $\Delta\delta(t)$ and $\Delta\tau(t)$. Moreover the maximum of the parameters captures a large event in the analyzed window. We interpreted that not only a large event produces a complexity increase in the seismicity but this could also be generated by seismic swarm. We explore the MF-DFA as a possible precursor index measure. However it is not obvious to define a precursor threshold value. To confirm this hypothesis would require to repeat this procedure in other series to obtain a wide perspective about the MF-DFA as a precursor tool.

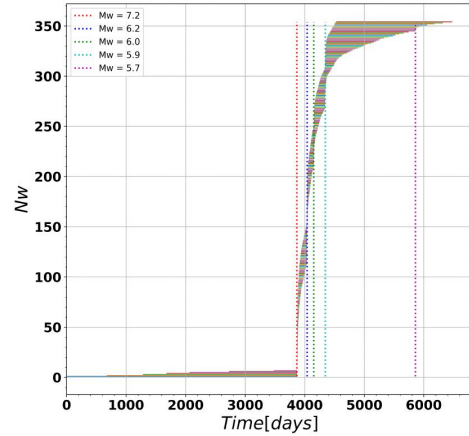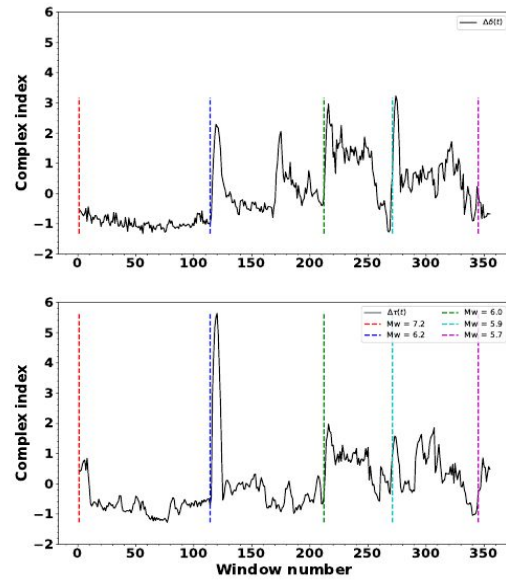Fig. 3 Evolution of each window length considering the same number of events per each one.

Fig. 4 Complexity Index as function of each window number. In dotted color lines the largest events are indicated.

REFERENCES

[1]   B.B. Mandelbrot. The fractal geometry of nature/revised and enlarged edition. New York, WH Freeman and Co., 1983, 495 p., 1983.

[2]   B. Enescu, K. Ito, M. Radulian, E Popescu, and O. Bazacliu. Multifractal and chaotic analysis of vrancea (romania) intermediate-depth earthquakes: investigation of the temporal distribution of events. Pure Aappl. Geo., 162(2):249–271, 2005.

[3]   T. Hirabayashi, K. Ito, and T. Yoshii. Multifractal analysis of earthquakes. In Fractals and Chaos in the Earth Sciences, pages 591–610. Springer, 1992.

[4]   C. Godano, ML Alonzo, and A Bottari. Multifractal analysis of the spatial distribution of earthquakes in southern italy. Geoph.l J. Int.l, 125(3):901–911, 1996.

[5]   P.N.S. Roy and S.K. Mondal. Multifractal analysis of earthquakes in kumaun himalaya and its surrounding region. J. Earth Syst. Sci., 121(4): 1033–1047, 2012.

[6]   A. Zamani and M. Agh-Atabai. Multifractal analysis of the spatial distribution of earthquake epicenters in the zagros and alborz-kopeh

dagh regions of iran. Iranian Journal of Science and Technology (Sciences), 35(1):39–51, 2011.

[7]   B. Gutenberg and C. F Richter. Frequency of earthquakes in california. Bul. Seismol. Soc. Am., 34(4):185–188, 1944.

[8]   E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, and H.J. Schellnhuber. Indication of a universal persistence law governing atmospheric variability. Phys. Rev. Lett., 81(3): 729, 1998.

[9]   P. Talkner and Rudolf O Weber. Power spectrum and detrended fluctuation analysis: Application to daily temperatures. Phys. Rev. E, 62(1):150, 2000.

[10]   J. W Kantelhardt, S.A Zschiegner, E. Koscielny-Bunde, Sh. Havlin, A. Bunde, and H.E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. Physica A: Statistical Mechanics and its Applications, 316(1-4):87–114, 2002

[11]   S. Shadkhoo and GR Jafari. Multifractal detrended cross-correlation analysis of temporal and spatial seismic data. The European Physical Journal B, 72(4): 679, 2009.

[12]   Shimizu et al. (2002), Fractals 10, 103-116

Marisol Monterrubio-Velasco received her PhD in computational physics by University Politécnica de Catalunya. She was postdoctoral researcher in the Geosciences center at Universidad Nacional Autónoma de México. At present she is postdoctoral researcher in the Barcelona Supercomputing Center Her research interests focus on computational physics, statistical analysis and numerical simulation applied to the Earthquakes.

Enhancing Scheduling through Monitoring and Prediction Techniques

Antoni Navarro Muñoz*†, Vicenç Beltran Querol*, Eduard Ayguadé Parra*†

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {antoni.navarro, vbeltran, eduard.ayguade}@bsc.es

Keywords—High-Performance Computing, OmpSs-2, Scheduling, Monitoring, Predictions, Cost

I. Extended Abstract

Modern applications become larger and more complex with each passing day. To name a few, weather forecasting or particle simulations are examples of how applications may have significant differences in features, constraints, and limitations.

Most runtimes supply users with functionalities to tune their executions. However, many aspects have to be taken into consideration when optimizing applications. Input sizes, recursive depths, system workloads, or the underlying architecture onto which apps are running, are just a few. Users often try different configurations until they stumble upon one which seems to yield the most performance. This proves to be nonportable, as a slight change in any of the aspects mentioned before might yield undesirable negative effects in performance.

In this work, our primary goal is to add several monitoring modules to runtimes. These modules introduce precise information about the units of work these libraries must schedule. The extension of these libraries allows for accurate real-time predictions for present and future executions. Such predictions can be used to obtain better scheduling of future units of work automatically and, therefore, improve the overall performance of executions or the utilization of resources. All this, while being unnoticed by users, thus giving more power to the runtimes.

Through the evaluation provided, we demonstrate the precision of our predictions and how they can be used to optimize resource utilization among others. We integrate all the extensions mentioned above on an already existing runtime maintaining the vision of the integration being capable of any similar runtime or library.

A. Monitoring Techniques

For the purpose of improving scheduling techniques adaptively and automatically, we propose a monitoring infrastructure. The primary objective of the infrastructure is to gather metrics and use them in real time. Our approach consists of an API that couples with an existent runtime. To exemplify this, we use OmpSs-2, the second generation of OmpSs [1], a task-based programming model. More specifically we integrate this infrastructure mentioned above in Nanos6 [2], a runtime library that implements OmpSs-2.
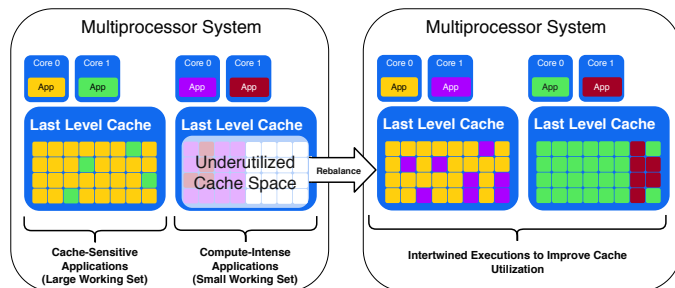

Fig. 1. Rebalancing applications across processors for optimal resource utilization

With this infrastructure, scenarios that are optimizable come to surface. One of these is to detect when resources can be exploited more efficiently. To serve as an example, depending on the internal static scheduling policies of a runtime, resources such as CPUs can be underutilized. Another scenario includes detecting when units of work can be scheduled more efficiently to improve execution time and overall performance.

Our primary goal is to detect these scenarios. Figure 1 exemplifies this. In this figure, we can see two multiprocessor systems, each with a last level cache and two cores. In each of these cores is running an application. On the first pair of cores we observe cache-intensive applications. On the second pair, compute-intensive ones. As shown, the applications on the left utilize the whole last level cache. This can lead to inefficient usage of the last level cache, as both applications will fight over the resources. On the other hand, the ones on the right underutilize the last level cache. This scenario can be optimized by rebalancing the workloads. Compute-intensive applications can be interleaved with cache-intensive ones. This is shown in the right part of the figure, where applications are mixed up. Rebalancing in these scenarios is bound to improve cache utilization.

Two APIs form our monitoring infrastructure. The first monitors timing metrics for elements such as tasks (units of work), threads, and CPUs. The latter monitors hardware events for the same parts. Both are generic APIs, and independent from each other. Next we show how researchers may benefit from using our infrastructure by creating smarter scheduling policies or mechanisms that take advantage of the collected metrics and predictions.
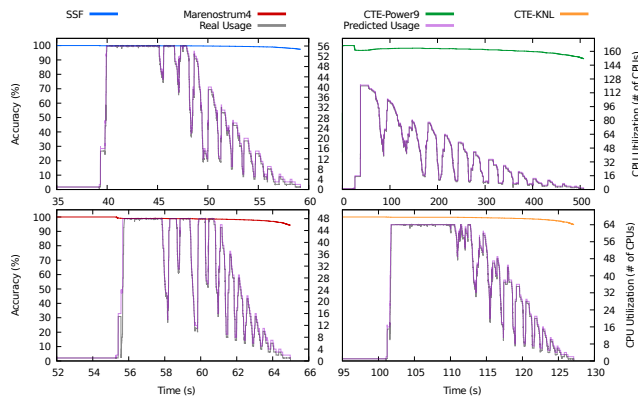
Fig. 2. Accuracy of the CPU predictor across different architectures

B. Current Contributions

So far, our contributions have targeted both of the previously mentioned scenarios. In our first contribution [3] we created a mechanism that automatically detects when an excessive amount of parallelism is being generated in recursive applications. Upon detecting such an event, the mechanism adapts the execution through the use of timing metrics and predictions. Through these predictions, it is capable of automatically ceasing the generation of recursive tasks. Instead, these are inlined in parent tasks.

In other recent contributions, we created a predictor that uses real-time data from the same monitoring infrastructure. This predictor is able to infer the amount of CPU needed to execute the current workload of the system. Next, we present the evaluation of the predictor.

C. Evaluation & Results

Our predictor was evaluated using a set of six benchmarks with varying features and granularities. Moreover, it was tested in different architectures, as our approaches are application and architecture independent. Some of the architectures tested are IBM's Power9 8335-GTG processors, Intel's Xeon Phi Knights Landing processors and Intel's Xeon Phi E5-2690v4 processors.

In figure 2 we demonstrate the effectiveness of our predictor in the Cholesky factorization application. We showcase four figures, each with the real CPU usage, the predicted CPU Usage and the overall accuracy of the prediction at each timestep. The four figures represent the accuracy of the predictor in the four architectures we tested.

Our results show that in all the architectures that were tested, and for any application or parameter, our predictor was able to precisely predict the CPU usage for the most part of all executions. This, as shown in that same figure, is not limited to sudden drops or peaks of workload. These last scenarios are tackled by using timing predictions and the size of internal queues.

D. Future Work

After demonstrating the accuracy of our timing predictions in previous works, we aim to assess the accuracy of our hardware-event based predictions.

Ensuring a consistent and precise infrastructure for metrics will lead to creating other mechanisms to ensure efficient executions – this, taking into account both resource usage and application performance. Thus, our roadmap includes detecting scenarios like those shown in figure 1 and introducing countermeasures to tackle them.

E. Conclusion

In our research, we have surfaced the need for adaptive scheduling policies that are independent of application and architecture features. We have also proven our predictions to be both accurate and useful for task-based programming models. In this study, we have focused on two of our previous works. The first targets application performance. The second resource efficiency optimization.

By having established a consistent ground for our research with a monitoring infrastructure, we are confident that our next contributions will enable researchers to realize the potential of adaptive techniques in their programming models and runtimes.

II. ACKNOWLEDGMENTS

REFERENCES

[1] Alejandro Duran, Eduard Ayguadé, Rosa M Badia, Jesús Labarta, Luis Martinell, Xavier Martorell, and Judit Planas. Ompss: a proposal for programming heterogeneous multi-core architectures. *Parallel processing letters*, 21 (02):173–193, 2011.

[2] Barcelona Supercomputing Center Programming Models Group. The nanos6 runtime repository, 2018. URL https://github.com/bsc-pm/nanos6. Accessed: 29-12-2018.

[3] Antoni Navarro, Sergi Mateo, Josep Maria Perez, Vicenç Beltran, and Eduard Ayguadé. Adaptive and architecture-independent task granularity for recursive applications. In Bronis R. de Supinski, Stephen L. Olivier, Christian Terboven, Barbara M. Chapman, and Matthias S. Müller, editors, *Scaling OpenMP for Exascale Performance and Portability*, pages 169–182, Cham, 2017. Springer International Publishing. ISBN 978-3-319-65578-9.

[4] Antoni Navarro, Vicenç Beltran, and Eduard Ayguadé. *Enhanced scheduling techniques through lightweight monitoring for OmpSs-2*. Master's thesis, Universitat Politècnica de Catalunya (UPC), 2019.

Antoni Navarro received his BSc degree in Computer Engineering and his MSc degree in the Master in Innovation and Research in Informatics with High-Performance Computing specialization from Universitat Politècnica de Catalunya (UPC), Barcelona, in 2016 and 2018 respectively. Since 2016 he has been in the Programming Models group of the Computer Science Department of Barcelona Supercomputing Center (BSC-CNS). In 2018, he started as a Ph.D. student at the department of Computer Architecture of Universitat Politècnica de Catalunya (UPC), Spain.

A One-Dimensional Finite Element Model for Human Circulatory Systems

David Oks*† and Mariano Vázquez*‡

*Barcelona Supercomputing Center (BSC)
†Universitat Politècnica de Catalunya (UPC)
‡ELEM Biotech
E-mail: david.oks@bsc.es, mariano@elem.bio

Index Terms—**Hemodynamics, Computational Biomechanics, Finite Element Method, Cardiovascular Mechanics, Respiratory Mechanics**

Extended Abstract

Cardiovascular disease and particle deposition in the human airways are some of the most critical issues in public health. Biomechanic simulations of the human cardiovascular and respiratory circulation, and electrical conduction systems address these concerns, enabling a further understanding of the physiological mechanics involved. These tools aid clinical diagnosis and have implications for the design of interventions such as surgical procedures, devices and aerosol drug delivery. The complexity of these systems motivate simplifying models capable of reproducing the main characteristics of interest, i.e.: wave propagation in the cardiovascular circulatory system, and particle-deposition in the respiratory system. Drastically reducing the computational cost, one-dimensional mathematical models are an option which has been extensively studied in application to blood circulation in arteries [1], and to a lesser extent to respiratory [2] and electrical conduction along fibres such as those contained in the Purkinje network [3].

Following the work of Formaggia et al. [4] we have developed a one-dimensional finite element model capable of reproducing the main wave-propagation characteristics in the human arterial system. With the objective of being able to run patient-specific simulations in the limited time frames required by practitioners, the algorithm is fully parallelizable between tubular segments so that it will provide HPC-grade efficient simulations. Our code allows the prescription of time-dependent boundary conditions making possible to couple with three-dimensional models. In particular, we plan to couple the one-dimensional model to the fully coupled fluid-electro-mechanical model of the human heart developed by Santiago et al. [5] and the large-scale Computational Fluid Dynamics model of human airways developed by Calmet et al. [6], both contained in Alya, the HPC multi-physics code developed at CASE, BSC [7].

We are currently in the process of parametrization and validation of the cardiovascular arterial model as well as optimizing the code to meet HPC efficiency requirements. In addition, we plan to develop the one-dimensional respiratory airways and electric fibre conduction models which will be coupled to the aforementioned codes.

Acknowledgments

References

[1] Shi Y., Lawford P. and Hose R. (2011) *Review of Zero-D and 1-D Models of Blood Flow in the Cardiovascular System*, BioMedical Engineering OnLine, Vol 10, 1, p. 33. doi: 10.1186/1475-925X-10-33

[2] Clavica F., Alastruey J., Sherwin S. J. and Khir A.W. (2009) *One-dimensional modelling of pulse wave propagation in human airway bifurcations in space–time variables*, 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, pp. 5482-5485. doi: 10.1109/IEMBS.2009.5334564

[3] Ramirez E., Saiz J., Trenor B., Ferrero J.M., Molto G. and Hernandez V. (2007), *Influence of 1B ischemic ventricular tissue on the automaticity of Purkinje fibers: A simulation study*, 2007 Computers in Cardiology, Durham, NC, pp. 617-620. doi: 10.1109/CIC.2007.4745561

[4] Formaggia L., Lamponi D. and Quarteroni A. (2003), *One-dimensional models for blood flow in arteries*, Journal of Engineering Mathematics, Vol. 47, Issue 3-4, pp. 251-276. https://doi.org/10.1023/B:ENGI.0000007980.01347.29

[5] Santiago A., Aguado-Sierra J., Zavala-Aké M., Doste-Beltran R., Gómez S., Arís, R., Cajas, J.C., Casoni, E. and Vázquez, M. (2018) *Fully coupled fluid-electro-mechanical model of the human heart for supercomputers*, Int J Numer Meth Biomed Engng, 34:e3140.

[6] Calmet H., Yamamoto T., Eguzkitza B., Lehmkuhl O., Olivares E., Kobayashi Y., Tomoda K., Houzeaux G. and Vázquez M. (2019) *Numerical evaluation of aerosol exhalation through nose treatment*, Journal of Aerosol Science, Vol. 128, pp. 1-13, ISSN 0021-8502.

[7] Vázquez M., Houzeaux G., Koric S., Artigues A., Aguado-Sierra J., Arís R., Mira D., Calmet H., Cucchietti F., Owen H., Taha A., Dering Burness E., Cela J.M. and Valero M. (2016). *Alya: Multiphysics engineering simulation toward exascale.* Journal of Computational Science, Vol. 14, pp. 15-27.

David Oks received his Licenciatura degree (equivalent to MSc degree) in Physics from the Universidad de Buenos Aires in 2017 and a MSc degree in Physics from the École Normale Supérieure de Lyon in 2018, specializing in experimental fluid dynamics. In 2018 he was awarded the la Caixa INPhINIT Fellowship Grant for Doctoral Studies at Spanish Research Centres of Excellence. He is now a PhD student at the Barcelona Supercomputing Center where his research focuses on simulations of the human arterial system and heart valves.

Training CNNs using high-resolution images of variable shape

Ferran Parés*, Dario Garcia-Gasulla*, Jesús Labarta*†

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {ferran.pares, dario.garcia, jesus.labarta}@bsc.es

Keywords—Neural Network, CNN, Medium dataset, High-memory.

I. EXTENDED ABSTRACT

Neural Networks, concretely Convolutional Neural Networks, have proven empirically to be capable of successfully solve image recognition problems. Usually, image recognition datasets are composed by a set of images that have been pre-processed to have exactly the same width and height, since consistency between input shapes is a mandatory condition for typical Neural Network architectures and a really helpful condition for training them successfully. Handling images of varying size and varying aspect ratio during a CNN training process have not been tackled in the literature and presents several challenges like adapting its architecture to avoid the equal image shape limitation, how to group images in batches, how the CNN is going to learn patterns at this wide range of sizes, how to generalize this new unprocessed problem, among others.

Additionally, image recognition images are usually down-sampled to a size of 256x256 or similar, which is a really low resolution that allows to fit in memory all the activations produced in the typical CNN architectures. In fact, using images of greater size arises memory consumption issues and new challenges from the point of view of learning.

So, the main objective of this work is to successfully train a CNN using high resolution images of varying size and aspect ratio.

A. Dataset

In this project we used the Medium dataset, a subset of the recently available Met dataset [1]. This dataset consist of several museum pieces like sculptures, paintings, photographs and its main material they are composed by. As you may expect, the problem is to predict the main material of the museum piece based on its image. This dataset contains images of varying size and aspect ratios, so it is suitable for this project. Smaller images have 0.25MP with shape 500x500 (width x height) and larger images have 16MP with shape 4000x4000 (Figure 1a), while aspect ratios vary from landscape images like 2000x1173 (Figure 1b) to portrait images with 572x1026 dimensions (Figure 1c).

B. CNN architecture

As said before, traditional CNN architectures are unable to work with images of varying shape. That's mainly due to fully connected layers, since these layers require a fixed input shape. In contrast, convolutional layers can work with different input shapes. So, there exists in the literature a specific type of CNN architecture called fully-convolutional that allows to work using variable input shapes. The main characteristic of this architecture is that it aggregates all the spatial information to a single value before feeding fully-connected layers, effectively fixing its shape allowing fully-connected usage.

The fully-convolutional layers may solve the ability to feed images of different shapes, but it will be really hard for the CNN to learn patterns at this wide range of scales. For this purpose we propose to use the Spatial Pyramid Pooling (SPP) [2] to handle several pattern scales before discriminating classes.

C. Input pipeline

The architecture proposed in the previous subsection allows to use input images of different shapes on each step, but still needs to join images together in batches. It is possible to join several images of different shapes by padding small images to match biggest ones in the same batch but, it is interesting to

(a) Vase made of terracotta

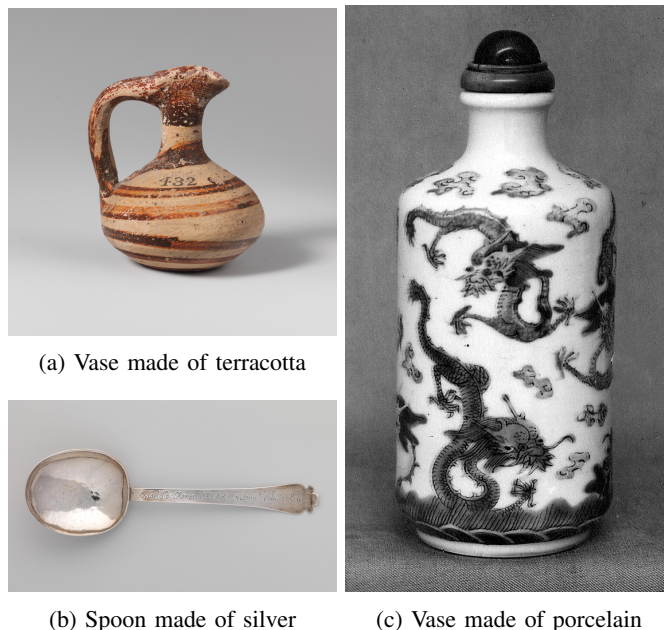(b) Spoon made of silver

(c) Vase made of porcelain

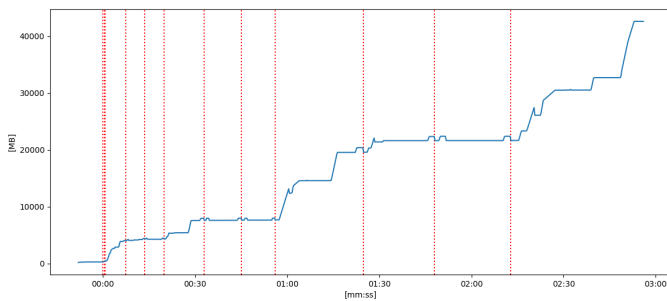Fig. 1: Medium dataset: Subset of images

Fig. 2: Memory profile during training process using images from a single bucket. In this case, the training process iteratively doubles the batch size every 3 steps. Initial batch size is 8 and it crashes when trying to use batch size of 64.

group similar images in the same batch to keep in memory as less non-informative padding as possible. Another reason to minimize the padding is the difficulties it generates in terms of learning. It is unknown in the literature how much padding affects training when it gets increased too much.

In order to group images by its shape similarity, we have already designed an input pipeline that we call Buckets to Buckets (B2B). B2B pipeline consists of 2 steps. The first one groups images in superbuckets based on the number of pixels and then, the second one groups images from superbuckets into smaller buckets based on their aspect ratio. However, the performance of this method resides in the right choice of the most appropriate size and aspect ratio boundaries used to separate superbuckets and buckets.

Size and aspect ratio boundaries are chosen based on an heuristic technique. This technique tries to find the set of boundaries that groups images producing as less padding as possible. So, the heuristic orders all the images based on their size or their aspect ratio, places arbitrary boundaries to split them in groups and, finally, iteratively move these boundaries back and forth to reduce the padding of all groups.

D. Dynamic batch size

Since we have grouped images in superbuckets based on its image size, there will be groups with small images and groups with larger ones. So, when using large images the pipeline will batch as much as possible while fitting in memory, but when using smaller ones the pipeline will be capable of batching more images and still fit in memory. So, on each step, the batch size used will vary depending on the bucket images are coming from.

However, the dynamic batch size also brings some drawbacks. From the learning point of view the system has to adapt its learning rate depending on the batch size in order to get training stability [3], and from the machine point of view, the dynamic batch size brings memory usage variability.

E. Memory profiling

Memory profiling has been a great tool in this project to detect memory instability, memory reduction when applying different B2B boundaries, batch size limits, etc. Memory profiles are build from tracking the process resident memory

during training and then visualized in plots like the one in Figure 2.

F. Current limitations to overcome

Currently this project uses Tensorflow and Marenostrum4 nodes. The main reason to use Marenostrum4 nodes its the amount of memory that nodes offer (96GB on regular nodes and even more on the high-memory nodes). Regardless of having 96GB available on the MN4 node, we are only capable of using around 40GB of this memory due to a Tensorflow limitation on its implementation. Tensorflow indexes some of its tensors with an int32 variable, so the library is unable to use tensors bigger than 2.147.483.648 positions.

G. Conclusions

This project offers a lot of challenges from the point of view of Artificial Intelligence and High-Performance Computing. There is no standard procedure to train CNN using images of varying size and aspect ratio in the AI literature, while the use of high-resolution images introduces a challenge in terms of memory consumption.

Some of the problems are already tackled and explained in this extended abstract and some of them remain to be solved. The main goal of this project is to successfully train a CNN under these dataset conditions and, train it as fast as possible using greater memory consumption.

REFERENCES

[1] M. Museum, "Open access initiative," https://github.com/metmuseum/openaccesse, 2017.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[3] P. Goyal, P. Dollár, R. Girshick, P. Noordhuis, L. Wesolowski, A. Kyrola, A. Tulloch, Y. Jia, and K. He, "Accurate, large minibatch sgd: Training imagenet in 1 hour," *arXiv preprint arXiv:1706.02677*, 2017.

Ferran Parés received his BSc degree in Electronics in the Universitat Politècnica de Catalunya (UPC), Spain in 2014. Afterward, he completed his MSc degree in Artificial Intelligence in 2016 from a MSc program with three participating universities: Universitat Politècnica de Catalunya (UPC), Universitat de Barcelona (UB) and Universitat Rovira i Virgili (URV). The same year, he joined Barcelona Supercomputing Center (BSC) in the High-Performance Artificial Intelligence group and started his PhD in Artificial Intelligence from Universitat Politècnica de Catalunya (UPC) in the following year (2017), Spain.

Correcting Air Quality Forecasts with Machine Learning Algorithms

Hervé Petetin[*1], Albert Soret[*2], Carlos Pérez Garcia-Pando[*3]

[*] *Barcelona Supercomputing Center (BSC)*

{[1]herve.petetin, [2]albert.soret, [3]carlos.perez}@bsc.es

Keywords— **air quality, forecast, machine learning**

EXTENDED ABSTRACT

Air pollution is a major environmental problem affecting human health and ecosystems. Mitigating the effect of pollution episodes requires reliable air quality forecasting (AQF) systems for both warning vulnerable populations and taking short-term measures of emission control (e.g. traffic reduction, interruption of some industrial facilities). The rise of chemistry-transport models (CTMs) over the last decades have allowed major improvements in AQF, supplanting the use of purely statistical forecasting systems. However, AQF systems remain affected by numerous sources of uncertainty (e.g. emissions, meteorology, initialization). In order to improve the forecasts, so-called model output statistics (MOS) methods are commonly applied to correct CTM outputs where continuous observations are available, typically at surface AQ monitoring stations. Various methods have been proposed in the literature, including moving averages (MA), Kalman filter (KF) and, more recently, analogs (AN) (e.g. Delle Monache et al., 2006; Kang et al., 2010; Djalalova et al., 2010, 2015; Huang et al., 2017). Although these methods allow removing a large part of the bias, they still suffer from some limitations, in particular when it comes to the detection of fast changes in atmospheric conditions.

In this work, we explore the use of machine learning (ML) techniques for correcting the raw AQF produced by CTMs based on various types of ancillary information, including meteorological features, past forecast errors, or temporal features. In ML terms, this corresponds to a problem of supervised regression in which the past observations along with their associated features are used to train a model that estimates the future pollutant concentrations on the basis of new (unseen) features. Several popular ML algorithms are tested, including gradient boosting machine (GBM), random forest (RF) and support vector machine (SVM). Results are compared with other MOS methods proposed in the literature.

The analysis is applied on the AIRE-CDMX operational AQF system of Mexico City. Recently built by the Earth Science Department, this AQF system is based on the combine use of the Community Multiscale Air Quality (CMAQ) chemistry-transport model and the Weather Research and Forecasting (WRF) meteorological model. The study focuses on the correction of the coarse particulate matter (PM_{10}, particles with aerodynamical diameter below 10 µm) forecasts over the period ranging from August 2017 to December 2018. We mimic an operational AQF system in which new forecasts and observations are obtained continuously. The first ML model is trained at the 30th day and then updated every 30 days. Different sets of features are tested. Both the training and tuning of the hyper-parameters are performed only based on the past available data. This is done independently at all surface stations of the local AQ monitoring network.

Among the traditional MOS methods, the AN method is found to give the best results, closely followed by the KF method. Concerning the MOS-ML methods, results show that they are able to compete with these other MOS methods when an appropriate set of features is taken into account. Relatively poor during the first months due a very short training period, their performance shows substantial improvement with time as the size of the training dataset increases (see Fig. 1). Among the different ML algorithms tested, best results are obtained with GBM. Compared to the raw forecast, the best GBM configuration gives a reduction of the root mean square error from 71 to 48%, and an increase of the correlation from 0.48 to 0.62.
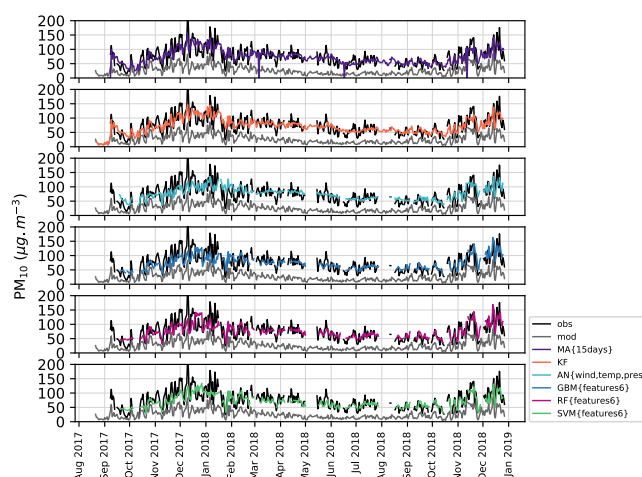

Fig. 1 Daily PM_{10} concentrations at a surface station of the AQ monitoring network. The lines show the observed concentrations ("obs"), the concentrations simulated by the AQF system without any correction ("mod"), and the concentrations corrected by the different MOS methods : moving averages ("MA"), Kalman filter ("KF"), Analogs ("AN"), gradient boosting machine ("GBM"), random forest ("RF") and support vector machine ("SVM").

One important objective of the AQF systems is to be able to predict in advance the occurrence of pollution episodes (here defined as PM_{10} concentrations exceeding a given threshold concentration) with a good reliability. The different MOS methods have been compared on a set of metrics related to episode detection skills. The raw forecasts given by the AQF system show a poor skill for detecting PM_{10} episodes due to a strong negative bias. The KF method shows a good ability to detect episodes but with numerous false detections. In other terms, many of the observed episodes are predicted by the AQF system but a forecasted episode has relatively low chance to actually happen. Conversely, most ML methods detect a lower number of pollution episodes but with a better confidence. Such a characteristic can be interesting when the short-term measures of emission control have a high financial cost and/or a low social acceptance.

A. Conclusion and Perspectives

These preliminary results highlight the potential of ML algorithms for improving the AQF provided by geophysical models. Considering the small size of the training datasets, these results are considered promising. In order to assess more precisely the ability of ML to correct AQF better than more established methods, longer datasets are required. Such an analysis is currently ongoing with a multi-year simulation from the NMMB-MONARCH model developed at BSC. This new dataset will allow to monitor how the performance of the MOS-ML methods evolves in time.

B. ACKNOWLEDGEMENTS

References

[1] L. Delle Monache, T. Nipen, X. Deng, Y. Zhou, and R. stull, "Ozone ensemble forecasts:2. A Kalman filter predictor bias correction", J. Geophys. Res., Vol. 111(D5), D05308, 2006.

[2] D. Kang, R. Methur, S. T. Rao, and S. Yu, "Bias adjustment techniques for improving ozone air quality forecasts", J. Geophys. Res., Vol. 111(D23), D23308, 2008.

[3] I. Djalalova, J. Wilczak, S. McKeen, G. Grell, S. Peckman, M. Pagowski, L. Delle Monache, J. McQueen, Y. Tang, and P. Lee, "Ensemble and bias-correction techniques for air quality model forecasts of surface O3 and PM2.5 during the TEXAQS-II experiment of 2006", Atmos. Environ., Vol. 44(4), Pp. 455-467, 2010.

[4] I. Djalalova, L. Delle Monache , and J. Wilczak, "PM2.5 analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model", Atmos. Environ., Vol. 108, Pp. 76-87, 2015.

[5] C. Borrego, A. Monteiro, M. T. Pay, I. Ribeiro, A. I. Miranda, S. Basart, and J. M. Baldasano, "How bias-correcction can improve air quality forecasts over Portugal, Atmos. Environ., Vol. 45(37), Pp. 6629-6641, 2011

[6] R. Zeng, W. Dietzel, R. Zettler, J. Chen, and K. U. Kainer, "Microstructure evolution and tensile properties of friction-stir-welded AM50 magnesium alloy," Trans. Nonferrous Met. Soc. China, Vol. 18, Pp. s76–s80, Dec. 2008.

[7] J. Huang, J. McQueen, J. Wilczak, I. Djalalova, I. Stajner, P. Shafran, D. Allured, P. Lee, L. Pan, D. Tong, H.-C. Huang, G. DiMego, S. Upadhayay, and L. Delle Monache, "Improving NOAA NAQFC PM2.5 predictions with a bias correction approach", Weather Forecast, Vol. 32(2), Pp. 407-421, 2017.

Author biography

Hervé Petetin was born in Caen, France in 1986. He holds an engineering diploma from the Ecole Centrale de Lille, France, a M.Sc. in Mechanics and fluid dynamics from the University of Science and Technology Lille, a M.Sc. in Atmospheric physics and chemistry from the University of Paris Est Creteil, France, and a Ph.D. in Atmospheric physics and chemistry from the University of Paris Diderot, France.

His research has first focused on the study of air quality and more specifically the aerosol pollution in large megacities, using regional chemistry-transport models and in-situ observations. As a member of the IAGOS European Research Infrastructure, he then investigated during several years the variability and trends of two important intermediate-lifetime gaseous compounds, namely ozone and carbon monoxide, in the troposphere based on airborne in-situ observations provided by the IAGOS fleet.

In 2018, he obtained a postdoctoral funding at the BSC from the STARS program (Marie-Sklodowska-Curie Action COFUND program) for working on the improvement of air quality forecasts with machine learning techniques.

Orchestration of Software Packages in Data Science Workflows

Cristian Ramon-Cortes*, Jorge Ejarque*, Rosa M. Badia*

*Barcelona Supercomputing Center (BSC)

E-mail: {cristian.ramoncortes, jorge.ejarque, rosa.m.badia}@bsc.es

Keywords—Distributed Computing, Workflow Managers, Programming Models, Data Science pipelines, Dust Forecast

I. EXTENDED ABSTRACT

Nowadays, Data Science applications are complex workflows composed of binary executions, MPI simulations, multi-threaded applications, and user-defined analysis (possibly written in Java, Python, C/C++, or R). Our proposal integrates this heterogeneity into a single task-based programming model capable of orchestrating the execution of the different frameworks in a transparent way and without modifying nor its behaviour, nor its syntax. Thus, our prototype is designed for non-expert users that want to build complex workflows where some steps require a highly optimised state of the art software packages.

A. Integration with Software Packages

To integrate the execution of other frameworks transparently into a single programming model, our prototype extends the COMPSs framework [1], [2] to act as an orchestrator rather than a regular application executor. Next, we detail the modifications of the programming model annotations, although the Runtime master, and worker executors have also been modified to schedule multi-node tasks and executing the different software packages.

The COMPSs Programming Model [3], [4] defines annotations that must be added to the sequential code in order to run the applications in parallel. These annotations can be split into Method Annotations and Parameter Annotations. Our prototype extends the programming model by providing a new set of Method Annotations and Parameter annotations to support the execution of binaries, multi-threaded applications (OmpSs [5]), MPI simulations, nested COMPSs applications, and multi-node tasks inside a workflow. From now on, tasks that must execute software packages are called *Non-Native Tasks*.

On the one hand, a new Method Annotation is defined for each supported non-native task. Notice that the Method Annotation must contain framework related parameters and, thus, its content varies depending on the target framework. Next, we list the currently supported frameworks and their specific parameters.

- **Binaries:** Execution of regular binaries (e.g., BASH, SH, C, C++, FORTRAN)
  - Binary: Binary name or path to an executable

- Working Directory: Working directory for the final binary execution

- **OmpSs:** Execution of OmpSs binaries
  - Binary: Path to the execution binary
  - Working Directory: Working directory for the final binary execution

- **MPI:** Execution of MPI binaries
  - Binary: Path to the execution binary
  - MPI Runner: Path to the MPI command to run
  - Computing Nodes: Number of required computing nodes
  - Working Directory: Working directory for the final binary execution

- **COMPSs:** Execution of nested COMPSs workflows
  - Application Name
  - Runcompss: Path to the runcompss command
  - Flags: Extra flags for the nested runcompss command
  - Computing Nodes: Number of required computing nodes
  - Working Directory: Working directory for the nested COMPSs application

- **Multi-node:** Execution of native Java/Python tasks that require more than one node
  - Computing Nodes: Number of required computing nodes

On the other hand, in order to input and output data from the execution of non-native tasks, the Parameter Annotation also needs to be enhanced. When using multi-node tasks the parameters and the return value of the task are the same than when using a regular method task. However, when executing standalone binaries, OmpSs processes, MPI processes, or COMPSs applications the exit value of the processes is used as the return value. Thus, we have decided that the COMPSs non-native tasks must use the exit value of their internal binary as the return value of the task. In this sense, our prototype allows the users to capture this value by defining the return type of the non-native task as an *int* (for implicit synchronisation), as an *Integer* (for post-access synchronisation) or to forget it (declaring the function as *void*).

However, the users do not only need the process exit value to work with this kind of applications but need to set the Standard Input (*stdIn*) and capture the Standard Output (*stdOut*) and Error (*stdErr*). For this purpose, our prototype includes a new Parameter Annotation, `stream`, that allows the users to set some parameters as I/O streams for the non-native

tasks. Stream parameters are not passed directly to the binary command but rather they are set as *stdIn*, *stdOut* or *stdErr* of the binary process. Since this kind of redirection is restricted to files in *LINUX* Operating Systems, we have decided to keep the same restrictions to the annotation. Consequently, all *stream* parameters must be files.

B. Evaluation

1) NMMB-MONARCH: The NMMB-MONARCH is a fully online multiscale chemical weather prediction system for regional and global-scale applications. The system is based on the meteorological Nonhydrostatic Multiscale Model on the B-grid [6], developed and widely verified at the National Centers for Environmental Prediction (NCEP). The model couples online the NMMB with the gas-phase and aerosol continuity equations to solve the atmospheric chemistry processes in detail. It is also designed to account for the feedbacks among gases, aerosol particles, and meteorology.

The NMMB-MONARCH workflow is composed by five main steps, namely *Initialization*, *Fixed*, *Variable*, *UMO Model*, and *Postprocess*. Although all the phases spawn several binaries, the NEMS binary soars above the rest. It is called during *UMO Model* step and its a Fortran 90 application parallelised with MPI. Therefore, NEMS can be executed in multiple cores and multiple nodes, relying on the MPI paradigm. The model uses the Earth System Modelling Framework library as the main framework [7].

2) Parallelisation design: The original NMMB-MONARCH application consists in a BASH script defining the main workflow and a set of Fortran binaries. Since our prototype supports Java and Python workflows, we have ported two different versions of the NMMB-MONARCH. Both of them parallelise the main workflow while keeping the Fortran binaries.

Regarding the parallelisation, all the binaries have been considered as tasks. For instance, Figure 1 shows the annotation of the `deeptemperature` binary in Python. We do not show the equivalent Java version due to space constraints.

```
@binary(binary='/path/to/deeptemperature.x')
@task(returns=int,
      seamask=FILE_IN,
      deep_temperature=FILE_OUT)
def deeptemperature():
    pass
```

Fig. 1.   Annotation of the `deeptemperature` binary in Python

As previously explained, the NEMS simulator invoked within the *UMO Model* step is implemented using MPI. Thus, in contrast to the rest of binaries, we have annotated it as an MPI task. Figure 2 shows the annotation of the `NEMS` binary in Python. Notice that, considering that the NEMS execution can be performed with a different number of nodes, the constraint decorator attached to the MPI task may vary between executions. Hence, to ease this management to users, the constraint has been defined using an environment variable.

To conclude, the final workflow structure for both the Java and the Python versions is depicted in Figure 3. Readers can identify the *Fixed*, *Variable*, *UMO Model*, *Postprocess* and *Figures and animations creation* steps, where the *Variable*, *UMO Model* and *Postprocess* are executed thrice.

```
@constraint(computingUnits='$NEMS_CUS_PER_NODE')
@mpi(runner='mpirun',
     binary='/path/to/NEMS.x',
     workingDir='/path/to/nems/out',
     computingNodes='$NEMS_NODES')
@task(returns=int,
      stdOutFile=FILE_OUT_STDOUT,
      stdErrFile=FILE_OUT_STDERR)
def nems():
    pass
```

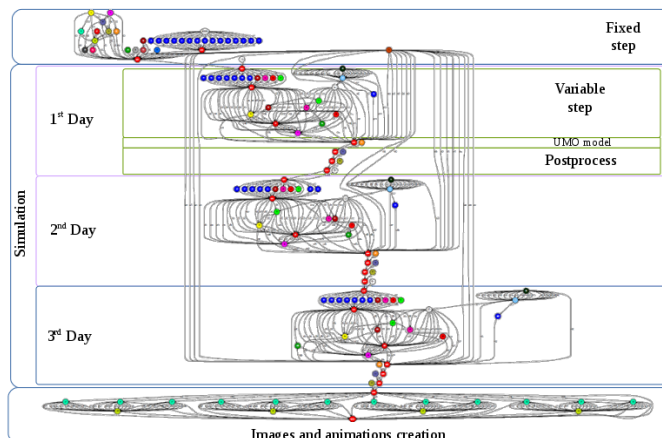Fig. 2.   Annotation of the `nems` MPI binary in Python

Fig. 3.   Dependency graph of three days simulation

3) Computing infrastructure: The infrastructure used in this comparison is the Nord 3 cluster (a subset of the MareNostrum III Supercomputer [8]), located at the Barcelona Supercomputing Center (BSC). This supercomputer is composed of 84 nodes, each with two Intel SandyBridge-EP E5-2670 (8 cores at 2.6 GHz with 20 MB cache each), a main memory of 128 GB, FDR-10 Infiniband and Gigabit Ethernet network interconnections, and 1.9 PB of disk storage.

The Java version used is 1.8.0_u112 64 bits, the Python version used is 2.7.13, and the MPI used is OpenMPI 1.8.1.

4) Performance: Table I compares each step of the previous version (BASH) against the Java and Python ones using 4 worker processes (64 cores).

Step	Execution Time (s)			Speed-up (u)	
	BASH	**Java**	**Python**	**Java**	**Python**
Fixed	290	117	119	2.48	2.43
Variable	26	19	22	1.37	1.18
Model Simulation	244	242	233	1.01	1.04
Post process	38	34	33	1.12	1.15
Total	**601**	**413**	**415**	**1.45**	**1.45**

TABLE I.   PERFORMANCE PER STEP WITH 4 WORKERS (64 CORES)

Notice that both Java and Python versions improve the performance in the *Fixed* and *Variable* steps due to the possibility of performing multiple tasks at the same time during these steps. In opposition, the *Model Simulation* and *Post Process* do not improve because they are composed, respectively, by a single task, and two tasks with a sequential dependency. However, the performance of these steps does not degrade either.
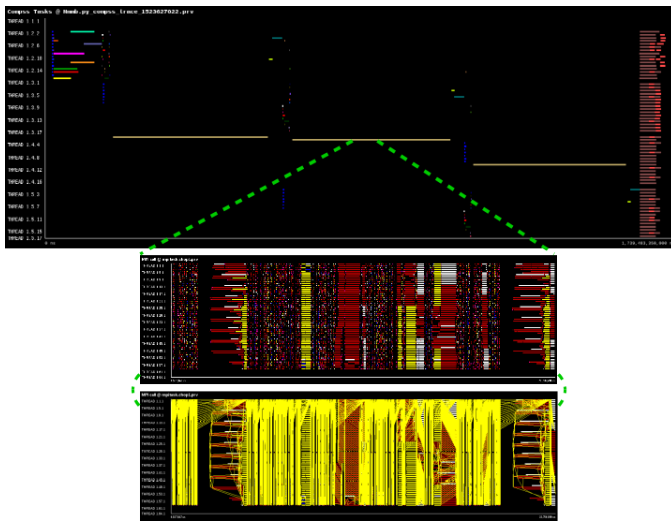
Fig. 4. Paraver trace of the Python version using 4 workers (64 cores)

Figure 4 shows Paraver [9] traces of a three-day simulation with 4 workers (64 cores) of the Python workflow. There are three timelines where each row corresponds to a thread in the worker. The top view shows a task view, where different coloured segments represent tasks, and each colour identifies a different task type. The middle and bottom views are the internals of the MPI NEMS task, where the yellow lines represent communications between MPI ranks.

The trace detail in the middle and bottom views show, respectively, the MPI events and the MPI communications of the *Model simulation* task. Notice that, during this phase, both frameworks, COMPSs and MPI, are working together and sharing the computing nodes. Although in this case MPI is using the four nodes, other applications may reserve some nodes for MPI and use the others to compute remaining sequential PyCOMPSs tasks.

C. Conclusions and Future Work

Our prototype enables non-expert users to develop complex Data Science workflows where some steps require a highly optimised state of the art software package. It provides a single task-based framework with a homogeneous annotation to orchestrate the execution of native Java and Python tasks along with binaries, MPI simulations, multi-threaded applications, and nested applications. From the users' point of view, the annotation is transparent and does not require to modify nor the behaviour, nor the syntax of the underlying software packages.

Moreover, the evaluation demonstrates that our prototype eases the development of complex workflows such as NMMB-MONARCH. We have developed two new implementations in Java and Python that provide better configuration management and object-oriented structure and improve the debugging, maintenance, and extension of the code. Also, the performance analysis demonstrates that our prototype is capable of extracting the parallelism of the NMMB-MONARCH application achieving 1.45 overall speed-up and sharing resources with multi-node frameworks such as MPI.

II. ACKNOWLEDGMENT

REFERENCES

[1] R. M. Badia and et al., "COMP superscalar, an interoperable programming framework," *SoftwareX*, vol. 3, pp. 32–36, 12 2015. [Online]. Available: http://dx.doi.org/10.1016/j.softx.2015.10.004

[2] Workflows and Distributed Computing - Barcelona Supercomputing Center (BSC). (2018) COMP Superscalar. [Online]. Available: http://compss.bsc.es

[3] F. Lordan *et al.*, "ServiceSs: an interoperable programming framework for the Cloud," *Journal of Grid Computing*, vol. 12, no. 1, pp. 67–91, 3 2014. [Online]. Available: https://digital.csic.es/handle/10261/132141

[4] E. Tejedor *et al.*, "PyCOMPSs: Parallel computational workflows in Python," *The International Journal of High Performance Computing Applications (IJHPCA)*, vol. 31, pp. 66–82, 2017. [Online]. Available: http://dx.doi.org/10.1177/1094342015594678

[5] A. Duran *et al.*, "Ompss: a proposal for programming heterogeneous multi-core architectures," *Parallel Processing Letters*, vol. 21, no. 02, pp. 173–193, 2011.

[6] Z. Janjic and R. Gall, "Scientific Documentation of the NCEP Non-hydrostatic Multiscale Model on the B Grid (NMMB). Part 1 Dynamics, Technical Report," NCEP, BOULDER, COLORADO, Tech. Rep. 80307-3000, 4 2012.

[7] C. Hill *et al.*, "The architecture of the Earth System Modeling Framework," *Computing in Science Engineering*, vol. 6, no. 1, pp. 18–28, 1 2004.

[8] Barcelona Supercomputing Center (BSC). (2017) MareNostrum 3. [Online]. Available: https://www.bsc.es/marenostrum/marenostrum/mn3

[9] ——. (2017) Paraver Tool. [Online]. Available: https://tools.bsc.es/paraver

[10] J. Conejero *et al.*, "Boosting Atmospheric Dust Forecast with PyCOMPSs," in *2018 IEEE 14th International Conference on e-Science (e-Science)*, Oct 2018, pp. 464–474.

Cristian Ramon-Cortes is a PhD Student for the Computer Architecture Department (DAC - UPC) working in collaboration with the Workflows and Distributed Computing group (WDC) at the Barcelona Supercomputing Center (BSC). He holds a MSc. on Research and Innovation in Informatics - High Performance Computing (MIRI - HPC, 2017), an Engineering Degree on Computer Science (FIB, 2014), an Engineering Degree on Industrial Engineering (ETSEIB, 2014), and a Dual BSc. Diploma (CFIS, 2014); all of them from the Technical University of Catalonia (UPC). During his career at the BSC he has contributed in the design and development of COMPSs, PyCOMPSs, and PMES. His areas of interest include Programming Models for Distributed Platform, Workflow Managers, Task Flows, Data Flows, and Streaming Technologies.

Towards PET degradation engineering

Sergi Rodà*†, Gerard Santiago*, Víctor Guallar*

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Autònoma de Barcelona, Barcelona, Spain

E-mail: {sergi.rodallordes, gerard.santiago, victor.guallar}@bsc.es

Keywords—PET, PETase, Plurizymes, Monte Carlo.

I. EXTENDED ABSTRACT

Plastics and their massive production are a world threat that needs to be dealt as soon as possible. Polyethylene terephthalate (PET) is one of the most widely used plastic polymers in the globe and its stability leads to a long lifetime in nature. Thus, its biodegradation was not considered until recently when a PET hydrolyzing bacteria was found in a PET bottle recycling site on Japan [1]. From this bacteria, two novel enzymes were found called PETase and MHETase. PETase degraded the polymer into subpolymers, while the second enzyme fully cleaved the monomer into subunits that the bacteria used as carbon source.

The following years consisted in elucidating the structure of the enzyme to decipher the exact catalytic mechanism of action against the PET polymer. Several research groups obtained X-ray data to generate different models of the WT enzyme and even mutant engineered versions of it [2], [3], [4], [5].

Our research group has previous experience on enzyme engineering and had a recent paper regarding the addition of an extra active site to a functional enzyme, what we call plurizymes [6].

Therefore, the main goal of this research is to rationally design the current PETase enzyme to contain an extra active site for the polymer increasing its activity.

A. PELE simulations

To design a new active site in the enzyme, we have to explore the surface of the protein for unknown binding sites first. To accomplish this, we use PELE software, which is a Monte Carlo technique combined with side chain prediction [7].

The PDB codes for the used crystal structures in the simulations are 6EQE and 5YNS for the WT and the R280A mutant enzymes, respectively [2], [5]. The R280A mutant version is also studied, since it is more efficient than the WT protein.

Regarding the ligands, we have used the monomer (MHET), the dimer ((MHET)$_2$), the dimer with an extra tail of ethylene glycol (2-HE(MHET)$_2$) and the tetramer with an extra tail of ethylene glycol (2-HE(MHET)$_4$) due to their common use in the structural studies of the PETase (see Figure 1 to visualize both the enzyme and the ligands).

The analysis of the results consists in mainly plotting the different parameters/metrics stored during the simulation to
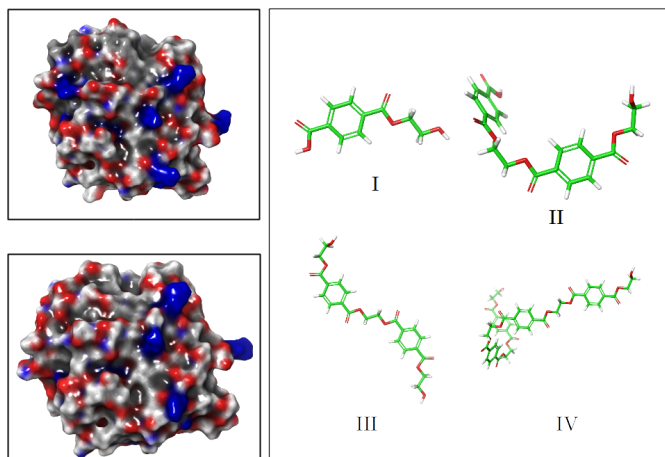

Fig. 1. Left: Surfaces of 6EQE (top) and 5YNS (bottom) crystal structures. Right: I; MHET, II; (MHET)$_2$, III; 2-HE(MHET)$_2$, IV; 2-HE(MHET)$_4$.

find the the local minima, which are the structures of the protein with the ligand bound to it.

Once new binding sites are found, they are functionalized according to different concepts related with chemistry and statistical mechanics.

B. Results

After running PELE simulations for all the mentioned ligands with specific parameters, we see different unknown binding sites that correspond to local minima or are metastable as it can be seen in Figure 2.

For instance, the local minima found 20 Å away from the serine residue (S) in the main active site shows nice properties that enable its functionalization. It shows a S near the binding of the ligand and two near residues that can be easily changed by histidine and an acidic residue. Besides, the binding site resembles a cavity and it presents a tryptophan residue that interacts with the benzene ring in the substrate.

Hence, we rationally mutated some of the near residues to create the catalytic triad and to still accommodate the substrate as well.

Once we have the selected mutations, we rerun PELE on those mutations to see if the binding site is explored with the inserted mutations. Once we have done this and we see that the ligand still explores the new potential active site, we run a MD simulation of some ns to see the stability of the complex with the mutations.
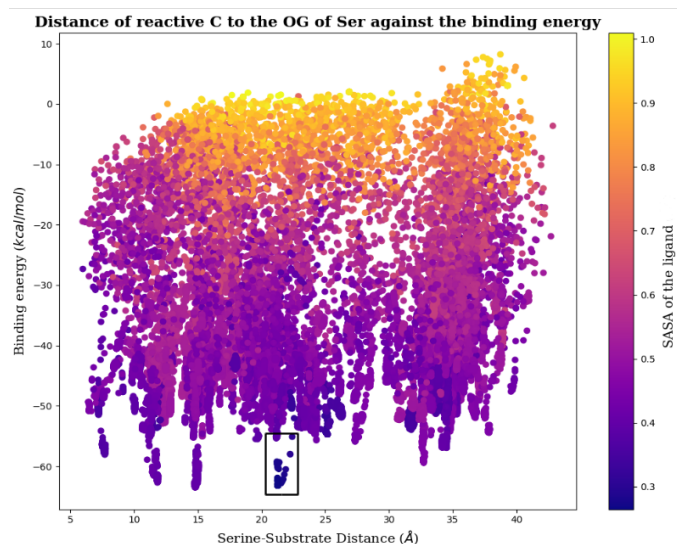
Fig. 2. Plot that represents the distance of the reactive C in the substrate to the nucleophile O_γ in the serine residue in the main active site against the binding energy of the substrate with the enzyme. The colorbar refers to the normalized solvent accessible surface area (SASA) of the ligand.

We have created 3 potential active sites but we are still currently working on the MD simulations and exploring all the PELE results.

REFERENCES

[1] S. Yoshida, K. Hiraga, T. Takehana, I. Taniguchi, H. Yamaji, Y. Maeda, K. Toyohara, K. Miyamoto, Y. Kimura, and K. Oda, "A bacterium that degrades and assimilates poly(ethylene terephthalate)," *Science*, vol. 351, no. 6278, pp. 1196–1199, 3 2016.

[2] S. Joo, I. J. Cho, H. Seo, H. F. Son, H.-Y. Sagong, T. J. Shin, S. Y. Choi, S. Y. Lee, and K.-J. Kim, "Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation," *Nature Communications*, vol. 9, no. 1, p. 382, 12 2018.

[3] X. Han, W. Liu, J.-W. Huang, J. Ma, Y. Zheng, T.-P. Ko, L. Xu, Y.-S. Cheng, C.-C. Chen, and R.-T. Guo, "Structural insight into catalytic mechanism of PET hydrolase," *Nature Communications*, vol. 8, no. 1, p. 2106, 12 2017.

[4] T. Fecker, P. Galaz-Davison, F. Engelberger, Y. Narui, M. Sotomayor, L. P. Parra, and C. A. Ramírez-Sarmiento, "Active Site Flexibility as a Hallmark for Efficient PET Degradation by I. sakaiensis PETase," *Biophysical Journal*, vol. 114, no. 6, pp. 1302–1312, 3 2018.

[5] H. P. Austin, M. D. Allen, B. S. Donohoe, N. A. Rorrer, F. L. Kearns, R. L. Silveira, B. C. Pollard, G. Dominick, R. Duman, K. E. Omari, V. Mykhaylyk, A. Wagner, W. E. Michener, A. Amore, M. S. Skaf, M. F. Crowley, A. W. Thorne, C. W. Johnson, H. L. Woodcock, J. E. McGeehan, and G. T. Beckham, "Characterization and engineering of a plastic-degrading aromatic polyesterase," *Proceedings of the National Academy of Sciences*, vol. 115, no. 19, pp. E4350–E4357, 5 2018.

[6] G. Santiago, M. Martínez-Martínez, S. Alonso, R. Bargiela, C. Coscolín, P. N. Golyshin, V. Guallar, and M. Ferrer, "Rational Engineering of Multiple Active Sites in an Ester Hydrolase," *Biochemistry*, vol. 57, no. 15, pp. 2245–2255, 4 2018.

[7] Kenneth W. Borrelli, Andreas Vitalis, R. Alcantara, , and V. Guallar*, "PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique," 2005.

Sergi Rodà is a native of Terrassa and is currently a MSc research student at the BSC in the EAPM group led by Víctor Guallar. He got his BSc degree in Biochemistry at the "Universitat Autònoma de Barcelona". He has worked in the group of molecular bases of disease at the "Institut Investigació Biomèdica - Sant Pau" as a researcher. He started his MSc studies in Bioinformatics on October of 2018 with the "Becas Másteres de Excelencia - Fundació Catalunya-La Pedrera" grant.

Containers in HPC: A Scalability and Portability Study in Production Biological Simulations

Oleksandr Rudyy*, Marta Garcia-Gasulla*, Filippo Mantovani*, Alfonso Santiago*, Raül Sirvent*, Mariano Vazquez*†

*Barcelona Supercomputing Center (BSC), Barcelona, Spain

†ELEM Biotech, Spain

E-mail: {oleksandr.rudyy, marta.garcia, filippo.mantovani, alfonso.santiago, raul.sirvent, mariano.vazquez}@bsc.es

Keywords—Container, HPC, Biological Simulation, Scalability, Portability

I. Extended Abstract

Container technology as lightweight virtualization has revolutionized during the last decade the IT business as well as drawn the attention of the High Performance-Computing (HPC) community. After the appearance of Docker [1] in 2013, there have emerged various container implementations aimed for HPC among which stand out Singularity [1] and Shifter [2]. However, the current literature lacks about deep evaluations of containers in HPC [3]. Even though we can find researches measuring and comparing containers performance with bare-metal, it is not enough representative for large HPC centers nor real scientific applications. Thus, we decided to take advantage of our resources conducting a prominent study about container viability in real HPC environments.

In this work we will not only study container performance using Alya [4], a in-production Computational Fluid Dynamics (CFD) code optimized for HPC, up to 256 computational nodes (around 12k cores), but also test container portability on three different state-of-the-art HPC architectures (Intel Skylake, IBM Power9, and Arm-v8) and compare three important container implementations. From the outcomes of all this, we hope to provide to system administrators, facility managers, HPC experts and field scientists a valuable research which to refer for guidelines and use-case examples.

A. Experimental environment

We will be using **Docker**, **Singularity** and **Shifter** container implementations. All they offer operating system virtualization, though, Docker paradigm differs from Singularity's and Shifter's. In addition that Docker relies on a root owned daemon, it also leverages both `cgroups` [2] and `namespaces` [3] capabilities causing full isolation of the container system from the host. On the contrary, Singularity and Shifter manage the SUID (Set owner User ID upon execution) method for running privileged system calls necessary to deploy the container's environment. Besides, they only handle Mount and PID `namespaces`, so their systems possess a more transparent interaction with the host, for instance, for MPI container-host communications.

About where to test our containers, in this research we leverage four high-end HPC clusters detailed below

Lenox Cluster is a four-nodes cluster, owned by Lenovo, where we have administrative rights. Each node contains a dual-sockets motherboard, housing $2\times$ Intel Xeon E5-2697v3, with 14 cores each (28 cores per node). It has installed Docker 1.11.1, Singularity 2.4.5 and Shifter 16.08.3. Compute nodes are interconnected via 1GbE network over TCP.

MareNostrum4 is a Tier-0 supercomputer in production at Barcelona Supercomputing Center (BSC) in Barcelona, Spain. Its nodes are based on Intel Xeon Platinum 8160 CPUs with 48 cores per node, with a total number of 3456 nodes available. The interconnection network is 100 Gbit/s Intel Omni-Path. Singularity 2.4.2 is deployed as container technology.

CTE-POWER is also hosted at Barcelona Supercomputing Center. This cluster is based on IBM Power9 8335-GTG processors of 20 cores where each compute node contains two CPUs providing 40 cores per node, with a total number of 52 nodes available. Nodes are interconnected via an Infiniband Mellanox EDR network. It has available Singularity 2.5.1.

ThunderX mini-cluster belongs to Mont-Blanc project [5]. The cluster houses four compute nodes, each containing 96 Armv8-a cores organized in two CN8890 sockets with 48 cores each. Nodes are interconnected using 40 GbE network over TCP. It runs Singularity 2.5.2.

B. Evaluation

1) **Containerization Solutions:** Where we compare the performance of the three container solutions under evaluation (Docker, Singularity and Shifter) regarding deployment overhead, image size and execution time.
2) **Portability:** Where we discuss container portability, executing the same containerized application with Singularity in three different architectures and using two techniques to build the container images.
3) **Scalability:** Where we compare the scalability of the Alya use-case comparing performance obtained running it at scale on MareNostrum4 in a Singularity container versus a bare-metal execution.

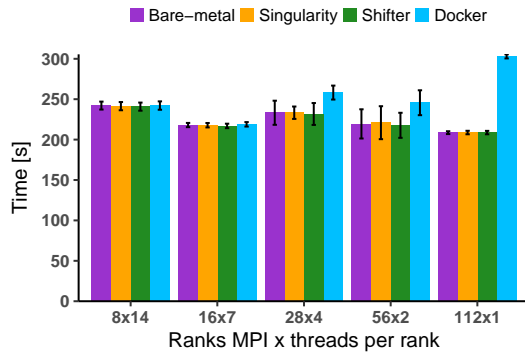For our evaluations we employ two biological use cases of Alya:

---

[1]More details about Docker in: https://www.docker.com/

[2]cgroups: http://man7.org/linux/man-pages/man7/cgroups.7.html

[3]namespaces: http://man7.org/linux/man-pages/man7/namespaces.7.html

Fig. 1: Average elapsed time of the artery CFD case in Lenox
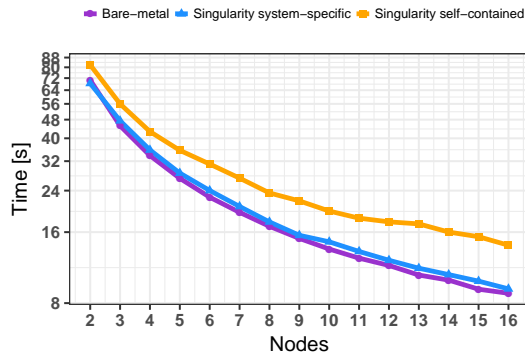
Fig. 2: Average elapsed time of artery CFD case in CTE-POWER

Fig. 3: Scalability plot of Alya artery FSI case in MareNostrum4

- **CFD:** The simulation of the fluid (blood) through the artery, which is a single code solving the Navier-Stokes equations for fluid dynamics.
- **FSI:** a fluid-structure artery simulation that requires two instances of different codes: the first code studying the fluid sub-domain and the second one simulating the solid sub-domain.

C. Results

In Fig. 1 the average time step duration for each version (Bare-metal, Docker, Singularity and Shifter) is shown. In the x-axis is displayed different configurations of MPI processes and OpenMP threads per rank. We can observe that HPC designed containers (i.e., Shifter and Singularity) can reach close to bare-metal performances whereas Docker degrades soon as we scale in MPI. During the portability test, we have realized that containers can be integrated with the host to leverage its specific features, for example the fast MPI interconnection. So, Fig. 2 shows the average time of three CFD artery versions. The integrated container (Singularity system-specific) can equal bare-metal performance, the opposite of the self-contained container, which is unable to use the Mellanox EDR network. Finally, Fig. 3 presents our scalability test of Alya's FSI use case in MareNostrum4 using up to 256 nodes (12.288 cores). As before, the integrated container can leverage the Intel Omni-Path network, unlike the self-contained which at 32 nodes stops scaling.
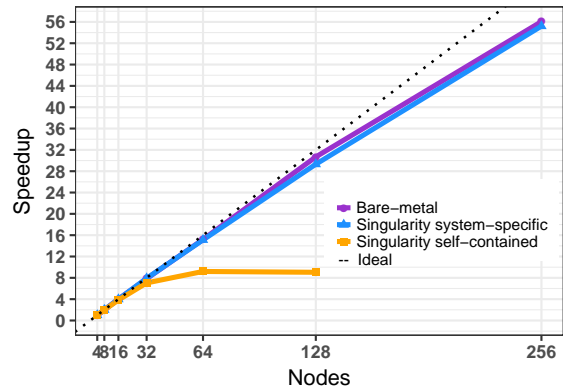
D. Conclusions and Future Work

In this paper we extensively investigated the deployment and performance of three container technologies under a production HPC workload. In summary, we found that containers are able to obtain bare-metal performances and can be tuned to leverage host-specific features in exchange of portability. Our study lacks a deeper evaluation of I/O and distributed storage performance using containers, which could be interesting future work.

II. ACKNOWLEDGMENT

REFERENCES

[1] G. M. Kurtzer, V. Sochat, and M. W. Bauer, "Singularity: Scientific containers for mobility of compute," *PloS one*, vol. 12, no. 5, p. e0177459, 2017.

[2] D. M. Jacobsen, "Contain this , unleashing docker for hpc," 2015.

[3] N. G. Bachiega, P. S. L. Souza, S. M. Bruschi, and S. d. R. S. de Souza, "Container-based performance evaluation: A survey and challenges," in *2018 IEEE International Conference on Cloud Engineering (IC2E)*, April 2018, pp. 398–403.

[4] E. Casoni, A. Jérusalem, C. Samaniego, B. Eguzkitza, P. Lafortune, D. D. Tjahjanto, X. Sáez, G. Houzeaux, and M. Vázquez, "Alya: computational solid mechanics for supercomputers," *Archives of Computational Methods in Engineering*, vol. 22, no. 4, pp. 557–576, 2015.

[5] N. Rajovic, A. Rico, F. Mantovani *et al.*, "The Mont-Blanc prototype: An alternative approach for HPC systems," in *SC '16: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov 2016, pp. 444–455.

Oleksandr Rudyy was born in Ivano-Frankivsk, Ukraine, in 1997. Currently he is finishing his Bachelor's Degree in Computer Engineering at UPC-FIB. Since 2018 he has been working in Barcelona Supercomputing Center as a junior research engineer, where he evaluates application container technologies within the Joint Research Activities of HPC-Europa3 project.

Experimental Study of Aggressive Undervolting in FPGAs

Behzad Salami, Osman S. Unsal, Adrian Cristal Kestelman

Barcelona Supercomputing Center (BSC), Barcelona, Spain.

{behzad.salami, Osman.unsal, adrian.cristal}@bsc.es

Keywords— **FPGAs, Voltage Underscaling, Energy, Reliability**

EXTENDED ABSTRACT

In this work, we evaluate aggressive undervolting, i.e., voltage scaling below the nominal level to reduce the energy consumption of Field Programmable Gate Arrays (FPGAs). Usually, voltage guardbands are added by chip vendors to ensure the worst-case process and environmental scenarios. Through experimenting on several FPGA architectures, we measure this voltage guardband to be on average 39% of the nominal level, which in turn, delivers more than an order of magnitude power savings. However, further undervolting below the voltage guardband may cause reliability issues as the result of the circuit delay increase, i.e., start to appear faults. We extensively characterize the behavior of these faults in terms of the rate, location, type, as well as sensitivity to environmental temperature, with a concentration of on-chip memories, or Block RAMs (BRAMs). Finally, we evaluate a typical FPGA-based Neural Network (NN) accelerator under low-voltage BRAM operations. In consequence, the substantial NN energy savings come with the cost of NN accuracy loss. To attain power savings without NN accuracy loss, we propose a novel technique that relies on the deterministic behavior of undervolting faults and can limit the accuracy loss to 0.1% without any timing-slack overhead.

A. Introduction

The power consumption of digital circuits, e.g., FPGAs, is directly related to their operating supply voltages. On the other hand, usually, chip vendors introduce a conservative voltage guardband below the standard nominal level to ensure the correct functionality of the design in the worst-case process and environmental scenarios. For instance, this voltage guardband is empirically measured to be 12%, 20%, and 16% of the nominal level in commercial CPUs [1], GPUs [2], and DRAMs [3], respectively. However, in many real-world applications, this guardband is extremely conservative and eliminating it can result in significant power savings without any overhead. Motivated by these studies, we extended the undervolting technique to commercial FPGAs, with a preliminary concentration on on-chip memories, or Block RAMs (BRAMs). Our experiments cover several representative FPGA platforms from Xilinx, a main vendor including a VC707, two identical samples of KC705, and a ZC705. The experimental results show the voltage guardband to be on average 39% of the nominal level (Vnom= 1V, Vmin = 0.61V), which in turn, directly delivers an order of magnitude BRAM power savings. Further undervolting below the minimum safe voltage, i.e., Vmin delivers more power savings up to 40%; however, causes faults occurrence in some locations of some of BRAMs. These faults are the consequence of timing violations since the circuit delay increases by further undervolting. Note that simultaneously downscaling the frequency is a promising approach to prevent the generation of these faults; however, it can limit the energy reduction achievement. Alternatively, our aim is to understand the behavior of these faults, through which customized and low-overhead fault mitigation techniques can be deployed to achieve power saving gains.

B. Experimental Methodology

The overall methodology is shown in Fig. 1. Our FPGA design includes raw Read/Write accesses to BRAMs, while their supply voltage, i.e., VCCBRAM is controlled in the host through the Power Management Bus (PMBus) interface. The onboard voltage regulator with the part number of UCD9248 has the responsibility to handle these PMBus commands and set the appropriate voltage to different components, e.g., BRAMs. Note that other FPGA components, e.g., LUTs, DSPs, operate at their default nominal voltage levels.
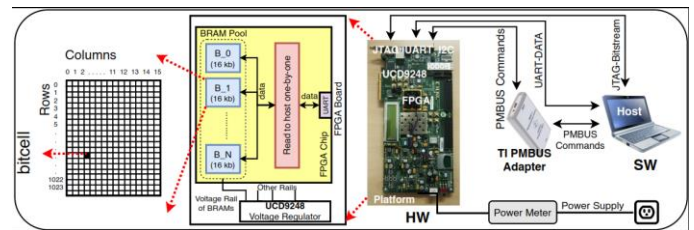

Fig. 1 The overall methodology for FPGA undervolting experimental study.

C. Selected Results

Through experiments on this setup, the overall power and reliability trade-off is summarized in Fig.2 for VC707, when VCCBRAM is downscaled from the nominal level Vnom = 1V to the minimum level that the FPGA practically operates, Vcrash = 0.54V. As can be seen, BRAMs start experiencing faults in regions below Vmin = 0.61V with an exponentially increasing behavior up to 653 faults per 1Mbit equals to 0.06%. Major observed properties of these faults are summarized as follows:

- There is significant variability of fault rate among different BRAMs, which is the consequence of the inherent process variation. Through our experiments, we observed that more than 38.9% of BRAMs never experience faults. Also, among BRAMs the maximum, minimum, and average fault rate are 2.84%, 0%, and 0.06%.

- The faults locations and rate do not change over time. Also, by experimentally evaluating different data patterns, we observed that the fault rate directly depends on the number of "1" bits since a vast majority of generated faults are '1' to '0' bit-flips.

- More than 90% of undervolting faults are single-bit, and a further 7% are double-bit faults. Due to this observation on the behavior of faults, the built-in ECC of BRAMs can be effective to mitigate these faults. Note that the built-in ECC of BRAMs has the type of Single-Error Correction and Double-Error Detection (SECDED) capability, potentially with good efficiency to mitigate BRAMs undervolting faults.

- Faulty bitcells in a certain voltage stay faulty in lower voltages, as well, and potentially, expand to other bitcells. This property is called Fault Inclusion Property (FIP). Our work experimentally confirms that FIP exists in

FPGAs, under aggressive low-voltage operations. FIP can be potentially used to build efficient fault mitigation techniques.
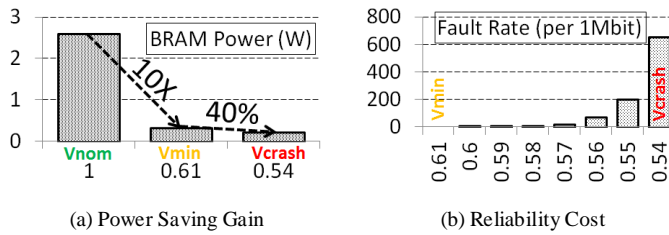

(a) Power Saving Gain   (b) Reliability Cost

Fig. 2 Power saving gain and reliability costs of FPGA BRAMs through undervolting.

D. CASE-STUDY: NEURAL NETWORK (NN)

In this section, we present and discuss results of our study on the impact of BRAMs undervolting in a typical FPGA-based NN accelerator. When VCCBRAM is underscaled in the critical region between Vmin= 0.61 V until Vcrash= 0.54 V, faults occurring in some of bitcells degrades the NN accuracy. Hence, the classification error is increased from 2.56% (inherent classification error without any fault) to 6.15% when VCCBRAM= Vcrash= 0.54 V, see Fig. 3. The NN classification error (left y-axis) increases exponentially, correlated directly with the fault rate increase in BRAMs (right y-axis), as expected.
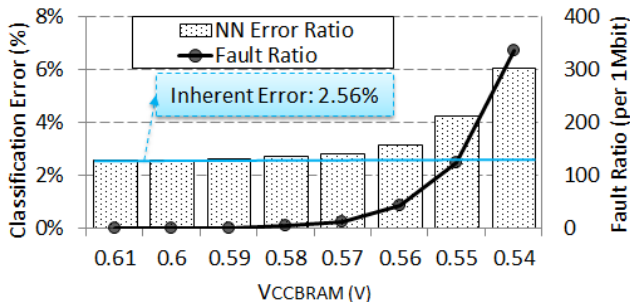

Fig. 3 Impact of BRAMs voltage scaling in the NN classification error, lowering VCCBRAM from Vmin = 0.61 V to Vcrash = 0.54V (VC707).

E. Conclusion and Future Enhancement

This paper experimentally evaluated the supply voltage underscaling below the nominal level in commercial FPGAs. We discovered that there is a significant voltage guardband gap, where data can be safely retrieved from BRAMs. However, by further undervolting observable faults occur, as a result of the timing delay increase. We extensively characterized the behavior of these faults, more specifically for on-chip memories of FPGAs. Finally, we evaluated the impact of the undervolting in the accuracy and power of an FPGA-based NN accelerator in the inference phase. To attain the power efficiency without NN accuracy loss, we proposed an efficient application-aware BRAM placement algorithm that relies on the behavior of undervolting faults. As an ongoing work, we are working on a more comprehensive voltage scaling in other components of commercial FPGAs and on different FPGA technologies of vendors.

References

[1] A. Bacha and R. Teodorescu. "Dynamic reduction of voltage margins by leveraging on-chip ecc in itanium ii processors," In ACM SIGARCH Computer Architecture News, volume 41, pages 297–307. ACM, 2013.

[2] J. Leng, A. Buyuktosunoglu, R. Bertran, P. Bose, and V. J. Reddi. "Safe limits on voltage reduction efficiency in gpus: a direct measurement approach", In 2015 48th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), pages 294–307. IEEE, 2015.

[3] K. K. Chang, G. Yaglıkcı, S. Ghose, A. Agrawal, N. Chatterjee, A. Kashyap, D. Lee, M. O'Connor, H. Hassan, and O. Mutlu. "Understanding reduced-voltage operation in modern dram devices: Experimental characterization, analysis, and mechanisms", In Proceedings of the ACM on Measurement and Analysis of Computing Systems, 1(1):10, 2017.

Author biography

Behzad Salami is a post-doc researcher in the Computer Science (CS) department of Barcelona Supercomputing Center (BSC). He received his Ph.D. in Computer Architecture from Universitat Politècnica de Catalunya (UPC) in 2018. During his Ph.D. studies, he visited the University of Manchester (UNIMAN) as a research internship student using a collaboration grant awarded from HiPEAC. Also, he obtained MS and BS degrees in Computer Engineering from Amirkabir University of Technology (AUT) and Iran University of Science and Technology (IUST), respectively. He was/is involved in several H2020/FP7 EU-funded research projects including AXLE (Advanced Analytics for Extremely Large European Databases), LEGaTO (Low Energy Toolset for Heterogeneous Computing), and EuroEXA. His research interests are heterogeneous computing and low-power & fault-resilient hardware accelerators.

Poster Abstracts

High-Integrity GPU Designs
for Critical Real-Time Automotive Systems

Sergi Alcaide*†, Leonidas Kosmidis*, Carles Hernandez*, Jaume Abella*

*Barcelona Supercomputing Center, Barcelona, Spain

†Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {sergi.alcaide, leonidas.kosmidis, carles.hernandez, jaume.abella}@bsc.es

Keywords—GPU, reliability, functional safety, diverse redundancy

I. EXTENDED ABSTRACT

The advent of autonomous driving (AD) makes automotive industry embrace high-performance hardware such as accelerators to execute performance-hungry tasks (e.g. object recognition and tracking) timely. However, while those accelerators have been widely deployed in the consumer electronics market, where performance within given power and thermal envelopes is the main concern, critical real-time systems, such as those in AD, pose a set of different challenges related to functional safety. In particular, safety-related automotive systems (e.g. braking, steering, etc), which include most AD functionalities, need to meet specific requirements described in the ISO26262 functional safety standard [1] to be deployed in cars. Those requirements, which mostly relate to the ability of the system to detect faults and prevent hazardous situations, impose strict verification and validation (V&V) requirements on the system and its components thereof. Hence, high-performance accelerators deployed in cars for AD must adhere to those requirements, which needs to be conveniently proven.

In the context of ISO26262, functionalities are classified in different Automotive Safety Integrity Levels (ASIL) based on the type of hazard they can cause, their severity, their exposure and the controllability upon a failure. In particular, safety-related functionalities are ranked from ASIL-D (the highest integrity level) to ASIL-A (the lowest), being ASIL-D components involved in ASIL-D functionalities those subject to the strictest V&V processes.

GPUs are becoming the most popular accelerator for AD, and they are already included in specific AD commercial platforms, such as RENESAS R-Car H3 [2] and NVIDIA Xavier [3] platforms. Those platforms include general purpose microcontrollers (e.g. ARM or Infineon cores) proven ASIL-D capable, as well as high-performance accelerators whose adherence to ASIL-D must also be proven so that they can perform AD-related activities. In particular, as detailed in ISO26262, ASIL-D systems must not allow a single fault lead the system to a hazardous situation. Appropriate safety measures include some form of independent redundancy, thus ensuring that a single fault will not lead redundant elements to identical erroneous outputs. For instance, Error Detection and/or Correction Codes are often used for storage and communication interfaces, whereas Dual Core LockStep (DCLS) with some source of diversity (e.g. staggered execution) is used for computation elements so that a single fault affecting all redundant components (e.g. a voltage droop) does not

cause them to fail identically (see example on Figure 1). In the case of GPUs, they have already been proven ASIL-B compliant, but, to our knowledge, ASIL-D compliance has only been achieved by implementing expensive independent redundancy means. In particular, either full system replication or heterogeneous implementations are used. The former, for instance, performs object recognition based on cameras and LIDAR, with different software implementations and, potentially, different hardware support. The latter, for instance, performs object recognition based only on cameras, but software is implemented and deployed for different accelerators (e.g. a GPU and a Deep Neural Network – DNN – accelerator) [4]. In both cases, design and V&V costs are duplicated, which is against efficiency and costs. For instance, these approaches clash with that followed for ASIL-D microcontrollers, which build upon diverse DCLS. Hence, it is critically important enabling some form of diverse DCLS on GPUs so that ASIL-D compliance can be achieved without needing to design and certify multiple heterogeneous software and hardware components.
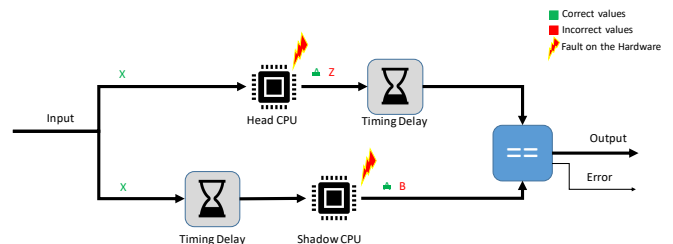

Fig. 1. Example of a Dual Core LockStep with Staggering Execution

This study tackles this challenge by identifying the main requirements to enable ASIL-D compliance for Commercial Off-The-Shelf (COTS) GPUs, assessing to what extent they have the potential to meet ASIL-D requirements, and providing a set of lowly-intrusive modifications that allow adhering to ASIL-D requirements without diminishing their performance for non-safety-related functionalities. Those modifications allow GPU vendors to reuse their designs avoiding a significant increase of their Non-Recurring Expenses (NRE). In particular, we perform our analysis on an NVIDIA COTS GPU, implement modifications on a GPU simulator where we can assess both performance and ASIL-D compliance, and evaluate what the performance impact would be on the COTS GPU.

REFERENCES

[1] International Standards Organization, *ISO/DIS 26262. Road Vehicles – Functional Safety*, 2009.

[2] "RENESAS R-Car H3," https://www.renesas.com/en-us/solutions/automotive/products/rcar-h3.html.

[3] D. Shapiro, "Introducing Xavier, the NVIDIA AI Supercomputer for the Future of Autonomous Transportation," *NVIDIA blog*, 2016. [Online]. Available: https://blogs.nvidia.com/blog/2016/09/28/xavier/

[4] NVIDIA, "NVIDIA Announces World's First Functionally Safe AI Self-Driving Platform," https://nvidianews.nvidia.com/news/nvidia-announces-worlds-first-functionally-safe-ai-self-driving-platform.

Sergi Alcaide received his BSc degree in Computer Engineering from Universitat Politècnica de Catalunya (UPC), Barcelona in 2016. Then he joined the Computer Architecture - Operative Systems (CAOS) group in Barcelona Supercomputing Center (BSC) while studying his master studies. In 2018, he completed his MSc degree, Master in Innovation and Research in Informatics (MIRI) from Universitat Politècnica de Catalunya (UPC). Now he is a PhD student from UPC in the Computer Architecture - Operative Systems (CAOS) group in BSC.

De novo binding prediction of peptides to MHC class I

Oriol Gracia[#1*], Pep Amengual-Rigo[#2*], Victor Guallar [#&3]

#*Barcelona Supercomputing Center (BSC)*

&*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

{[1]oriol.gracia, [2]jose.amengual, [3]victor.guallar}@bsc.es

Both authors contributed equally to this work

Keywords⸺ **Cancer, personalized medicine, immunotherapy, neoantigen, MHC class I**

Introduction

An emerging therapeutic application in immunology is the development of vaccines against cancer by using neoepitopes. A neoepitope is a peptide found only in the cancer state (coming from different gene expression or containing a mutation), that is able to generate an immune response. However, neoantigen discovery is hampered by the huge complexity of the immune system and the lack of experimental determinations.

The immune system is able to elicit an immune response against foreign antigens. Generation of the immune response depends on i) the degradation of proteins by the proteasome, and ii) the presentation of those peptides to the T-cells by the major histocompatibility center (MHC). Particularly in cancer, MHC class I role is to present self peptides to the CD8+ T-cells. There are described more than 10,000 MHC class I alleles[1], and it is expected that each of them is able to present around 5,000 different peptides. In contrast, a small fraction of peptides for less than 150 MHC class I alleles have been experimentally addressed. Current machine learning state-of-the art MHC binding predictors[2,3] are limited to deal only with the well characterized MHC class I alleles, hampering their application in personalized medicine. In this context we present uMHCpred, a de novo binding predictor that could give insights in the building of an universal model able to deal with the underrepresented MHC class I alleles.

Methodology

uMHCpred follows a frequency based approach that integrates the tridimensional structure of the MHC class I. More in detail, uMHCpred factorizes each residue of the peptide, building a position independent contact map. Then, the obtained contact map is used to classify each MHC by its structural similarity for each of the actorized peptide positions, archiving in this way a model that is no longer allele dependent.

In this way, uMHCpred is able to predict de novo MHC class I alleles using training data coming from those alleles that have a similar binding environment. In order to assess the performance of the de novo predictions, we compared them with a model containing all available data by computing the Receiver Operator Characteristic (ROC) and its Area Under the Curve (AUC).

Results

Our results showed that uMHCpred can predict de novo binding of MHC alleles, such as HLA-A*02:01 (**Figure 1**), reaching an accurate prediction (AUC of 0.88), with a minimal loss in comparison with the model that also uses data from the predicted allele (AUC of 0.91).
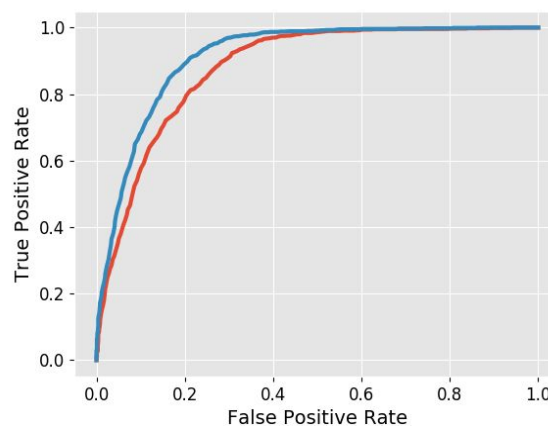

Figure 1*. ROC curves of the de novo prediction of HLA-A*02:01 (red) and the prediction that also uses data from the same allele (blue).*

Conclusions

uMHCpred can make accurate predictions for all the MHC variants including the ones that are underrepresented in the available data. Additionally, it is able to predict which peptides could be the best theoretical binders of each allele. Furthermore, this results could open a world of possibilities towards the neoepitope finding for personalized medicine.

References

1.   HLA Nomenclature @ hla.alleles.org.
     Available at:
     http://hla.alleles.org/nomenclature/stats.html.
     (Accessed: 2nd April 2019)
2.   O'Donnell, T. J. *et al.* MHCflurry:
     Open-Source Class I MHC Binding Affinity
     Prediction. *Cell Syst* **7**, 129–132.e4 (2018).
3.   Jurtz, V. *et al.* NetMHCpan-4.0: Improved
     Peptide–MHC Class I Interaction Predictions
     Integrating Eluted Ligand and Peptide
     Binding Affinity Data. *The Journal of
     Immunology* 199, 3360–3368 (2017).

Biography
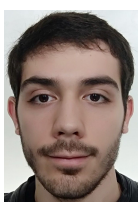
Oriol Gracia: BSc in Biochemistry (UB), MSc student in Bioinformatics for the Health Sciences (UPF), currently doing his Master's Thesis in the BSC supervised by Prof. Victor Guallar. His main interests are structural bioinformatics and binding energy prediction.
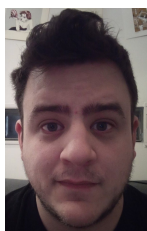
Pep Amengual-Rigo: BSc in Biochemistry (UIB), MSc in Bioinformatics (UAB), currently PhD student in UB under supervision of Prof. Victor Guallar at the BSC. His research lines are focusing antibody design and neoepitope prediction.

Local traffic contribution to black carbon horizontal and vertical profiles in compact urban areas

Jaime Benavides[1#], Marc Guevara[#] and Oriol Jorba[#]

[#]*Earth Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain*

[1]`jaime.benavides@bsc.es`

Keywords— **Vehicular traffic, source apportionment, urban areas**

EXTENDED ABSTRACT

Urban air pollution is characterized by strong spatial gradients produced by the presence of heavily trafficked streets. Given the negative health effects of air pollution, decision makers are implementing policies to reduce air pollution by modifying traffic flows near activity patterns of vulnerable populations. However, to our knowledge, there is quite a lack of appropriate decision tools to support such modifications at neighborhood to street levels. Measured horizontal and vertical distributions of traffic air pollutants can help understanding the variation of concentrations at increasing distances from emitting roads. Yet, models are still necessary to estimate the contribution of local traffic to measured concentrations.

The main objective of this work is to investigate the contribution of local traffic to black carbon horizontal and vertical profiles based on measurements and model simulations. We will discuss the degree of influence of the different streets on black carbon measured profiles in Barcelona.

A. CALIOPE-Urban

We apply the recently developed CALIOPE-Urban system for apportioning the black carbon traffic local sources. CALIOPE-URBAN [1] is the result of coupling the mesoscale modelling system CALIOPE [2] - based on WRF (meteorology), CMAQ (chemistry) and HERMES (emissions) - with R-LINE [3], a roadway dispersion model adapted to street canyons. The coupling considers atmospheric stability and street morphology to estimate local traffic dispersion and vertical mixing with background air (Fig. 1).
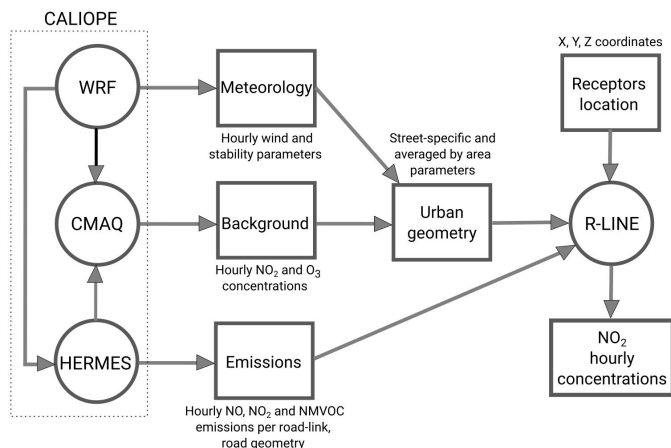

Fig. 1 CALIOPE-Urban workflow. Models are represented by circles and data by rectangular shapes. Meteorology and background from WRF and CMAQ are combined with urban geometry to create inputs for R-LINE.

B. Observational datasets

We use two datasets of observations to investigate the contribution of local traffic to Black Carbon (BC) profiles in Barcelona [4]. First, 29 horizontal profiles distributed along pedestrian streets perpendicular to trafficked streets are used to understand the dispersion at surface level. Measurements were collected starting from the trafficked street road edge until a maximum distance of 250 m. Monitors were located at different distances depending on local conditions to measure BC concentrations during approx. 30 minutes in four different days from 9 am to 6 pm. Second, 4 vertical profiles of BC starting from 2.5 to a maximum of 26 m height allow the analysis of the vertical dispersion within a street. Monitors were located depending on the number of floors of the buildings where measurements were undertaken. The period of measurements varied from 4 hours to 3 days.
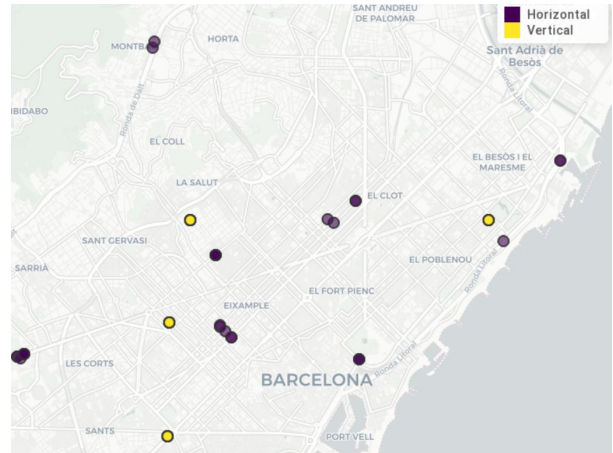

Fig. 2 Profile measurements sites represented by the measurement location closer to the road edge. Vertical measurements are depicted by yellow dots and horizontal are purple dots.

C. Experimental setup

Given that BC is a primary pollutant that is mostly emitted by vehicular traffic in Barcelona [5], there is a linear relationship between emissions sources strength and concentrations estimated at receptors. This fact facilitates the use of R-LINE model for source apportionment because it simulates the direct contribution of emission sources to receptors, labelling sources to study their specific influence at receptors locations. In order to investigate the streets contribution at each receptor, we establish two main contributors to BC: local vehicular traffic (i.e. emissions from streets within 1 km^2) and background concentrations. Local vehicular traffic streets are grouped depending on distance to the profile measurements (Fig. 3). We label the street segments forming four groups: nearest, which represents the closest street segment to the profile measurements; D_191, for street segments within a radial distance of 191 m consisting of the traffic flows that may surround four building blocks of the most spread geometrical pattern in the city (Eixample district); D_382, within a distance of 382 m serving as intermediate category between 382 m and 564 m; and D_564, within a

distance of 564 m accounting for street segments within 1 km^2.

Fig. 3 Street segment grouping for the source apportionment study. Green circles show the area considered in each group and yellow lines represent the radial distance of each circle.

D. Preliminary results

Fig. 4 shows the resulting modelled and observed profile in Diagonal Street close to Palau Reial Parc for the 14th of January 2015 between 12:06 and 12:22 local time. The measurement data (black dots) depict the concentration average during the 16 minutes period. We see a steep gradient within the first 20 meters, where concentrations decrease to less than 50% of the concentrations compared to the site close to the road edge. The model does not reproduce this strong initial decay, which may be due to the neglection of specific process occurring in the proximity of the road edge (e.g. traffic-induced turbulence). However, the model reproduces well the decay of concentrations from 20 meters to 200 meters in this location. With respect to local traffic contributions, the nearest street segment (from Diagonal Street) and the streets within 191 m are found to markedly influence concentration levels within the initial 50 meters. This result may indicate that if decision makers were aiming to reduce citizen exposure to BC in the initial 50 meters (e.g. if there were a school in the area), a traffic reduction from both nearest and streets within 191 meters should be considered.
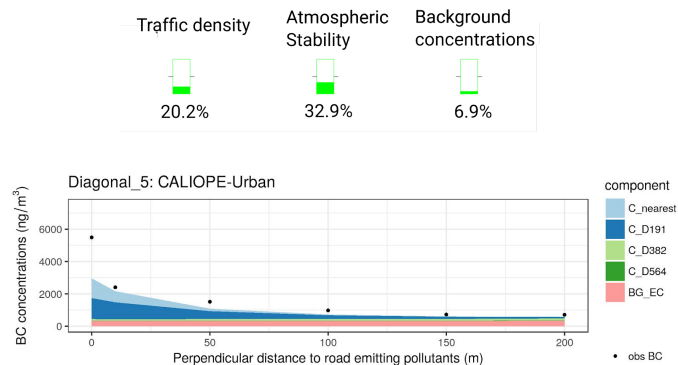

Fig. 4 Horizontal profile modelled and observed in Diagonal Street showing BC concentrations in ng/m^3 over distance to road edge. The white circles represent 2 m/s. On the top image, we see a description of the traffic density, atmospheric stability and background concentrations levels of the day and time when measurements were undertaken. In the graph below, we show the observed BC concentration levels (black dots) and the modelled contributions from local traffic (light blue for nearest, dark blue for streets within 191 m, light green for streets within 382 m, dark green for streets within 1 km ^{2} and pink for background concentrations. The wind speed and direction were: modelled (3.5 m/s; 315 degrees) and observed at Facultad de Fisica meteorological station (2 m/s; 320 degrees).

E. Conclusions and future enhancement

CALIOPE-Urban reproduces well horizontal profiles overall but it can't adequately reproduce the steep decay within the initial 20 meters from the road edge. We are working on extending the local traffic contribution analysis to all the horizontal and vertical profiles in the observational datasets.

F. ACKNOWLEDGEMENTS

References

[1] J. Benavides, M. Snyder, M. Guevara, A. Soret, C. Pérez García-Pando, F. Amato, X. Querol, and O. Jorba. "CALIOPE-Urban v1.0: Coupling R-LINE with a mesoscale air quality modelling system for urban air quality forecasts over Barcelona city (Spain)". Geosci. Model Dev. Discuss. In review, 2019.

[2] J.M. Baldasano, M.T. Pay, O. Jorba, S. Gassó and P. Jiménez-Guerrero. "An annual assessment of air quality with the CALIOPE modeling system over Spain". Science of the Total Environment 409.11, pp. 2163–2178. 2011.

[3] M.G. Snyder, A. Venkatram, D.K. Heist, S.G. Perry, W.B. Petersen, and V. Isakov. "RLINE: a line source dispersion model for near-surface releases". Atmospheric Environment 77, pp. 748–756. 2013.

[4] F. Amato, N. Perez, M. Lopez, A. Ripoll, A. Alastuey, M. Pandolfi, A. Karanasiou, A. Salmatonidis, E. Padoan, D. Frasca, M. Marcoccia, M. Viana, T. Moreno, C. Reche, V. Martins, M. Brines, M.C. Minguillón, M. Ealo, I. Rivas, J. Benavides, B. van Drooge, J.M. Craviotto, and X. Querol. "Vertical and horizontal decay of diesel pollutants within urban blocks". Submitted, 2019.

[5] X. Querol, A. Alastuey, M. Viana, T. Moreno, C. Reche, M.C. Minguillón, A. Ripoll, M. Pandolfi, F. Amato, A. Karanasiou, N. Pérez, J. Pey, M. Cusack, R. Vázquez, F. Plana, M. Dall'Osto, J. de la Rosa, A. Sánchez de la Campa, R. Fernández-Camacho, S. Rodríguez, C. Pio, L. Alados-Arboledas, G. Titos, B. Artíñano, P. Salvador, S. García Dos Santos, and R. Fernández Patier. "Variability of carbonaceous aerosols in remote, rural, urban and industrial environments in Spain: implications for air quality policy". Atmos. Chem. Phys., 13, 6185-6206. 2013.

Author biography

Jaime Benavides has a Msc. degree in Soft Computing and Intelligent Systems with emphasis on Data Mining and a BSc. degree in Civil Engineering both from University of Granada. Currently, he is doing a PhD under the supervision of Dr. Oriol Jorba, Dr. Albert Soret and Dr. Marc Guevara dealing with the development and evaluation of a street scale air quality modelling system over Barcelona within the Earth Sciences Department at BSC.

PluriZymes: new enzymes for new times

Marc Domingo-Cabasés[#][*], Sergi Rodà[#][*], Gerard Santiago[#], Victor Guallar [#][&]

#Barcelona Supercomputing Center (BSC)

& ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.

*Both authors contributed equally to this work

{marc.domingo, sergi.rodallordes, gerard.santiago, victor.guallar}@bsc.es

Keywords⸺ **PluriZyme, catalysis, active site, industry**

EXTENDED ABSTRACT

For the last decades, enzymes have been understood as globular proteins hosting one reactive region, called the active site. This state of fact has limited the enzymatic applications compared to heterogeneous catalysis, where a functionalized surface can host thousands of active sites. But, recent research has proved otherwise, enzymes can hold multiple active sites [1].

Applying a state-of-the-art modelling software, Protein Energy Landscape Exploration (PELE), detection of non-catalytic active sites is possible and, instead of blocking them (a classical approach), functionalize the region introducing the required mutations. The first attempt was done with esterases, because the catalysis requires only from three amino acids, the catalytic triad (serine, histidine and aspartic acid). Using a well-known esterase (PDB code: 5JD4) as initial model, a PluriZyme was *in silico* designed and *in vitro* tested (Fig. 1).
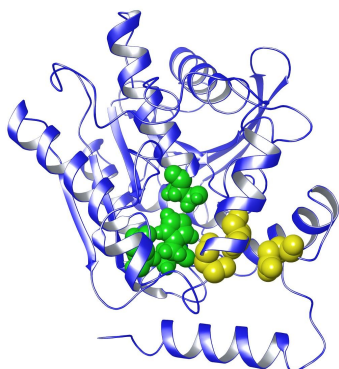

Figure 1: PluriZyme structure with an original active site (green) and an engineered one (yellow).

This first design succeed, both active sites (the natural and the engineered) were active and fully functional. Despite the positive results obtained, the final output was bittersweet, so the active site addition did not have the expected catalytic performance. The aftermatch was to be decided in further optimization rounds.

PluriZyme's engineered active site was deeply visual inspected aiming for potential mutation hotspots. Position 23 was found, an arginine that was close to the catalytic triad, reducing reaction space and potentially kidnapping the aspartate residue. An *in silico* saturated mutagenesis protocol was applied and for position 23 all short amino acids were tested, resulting glycine the most promising. Following *in vitro* validation proved a 7000 times activity fold in the artificial active sites and the obtention of an overall improved enzyme through additivity gain when both active sites are working at the same time. Furthermore, the analysis was extended to another esterases proving reproducibility.

Next natural step is the addition of different catalysis in the same enzyme scaffold and go through challenging problems. With this approach chain reaction can be placed together saving tremendous costs in industrial processes and reducing environmental harms. For example, enhanced PluriZymes for polyethylene terephthalate (PET) or auto-feeding unspecific peroxidases (UPOs), capable to produce their own required hydrogen peroxide.

References

1. Santiago, G. *et al.* Rational Engineering of Multiple Active Sites in an Ester Hydrolase. *Biochemistry* **57**, 2245–2255 (2018).

Marc Domingo is a native of Lleida and has lived there until 2013. Since then, he has lived in the surroundings of Barcelona where he studied a Bachelor's degree in Physics and Chemistry at Autonomous University of Barcelona (UAB) until 2018. Now, he is studying a Master in Bioinformatics (UAB) and he is doing the Professional Practice at EAPM group in BSC.
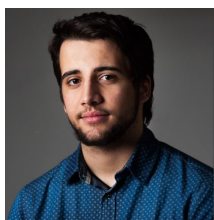
Sergi Rodà Llordés is a native of Terrassa and is currently a MSc research student at the BSC in the EAPM group led by Víctor Guallar. He got his BSc degree in Biochemistry at the "Universitat Autònoma de Barcelona". He has worked in the group of molecular bases of disease at the "Institut Investigació Biomèdica - Sant Pau"as a researcher. He started his MSc studies in Bioinformatics on October of 2018 with the "Becas Másteres de Excelencia - Fundació Catalunya-La Pedrera" grant.

Performance optimization of fully anisotropic elastic wave propagation on 2nd Generation Intel® Xeon Phi™ processors

Albert Farres*† Claudia Rosas* Mauricio Hanzich*
*Barcelona Supercomputing Center (BSC)
†Universitat Politècnica de Catalunya (UPC)
Spain
{albert.farres,claudia.rosas,mauricio.hanzich}@bsc.es

Alejandro Duran
Intel Corporation Iberia
Spain
alejandro.duran@intel.com

Charles Yount
Intel Corporation
USA
chuck.yount@intel.com

Keywords—Stencil-based wave propagation, Performance optimizations, Intel Xeon Phi, Fully Staggered Grid

I. Extended Abstract

In the last few years, acoustic isotropic wave propagation has been the preferred simulation engine for 3D full-wave field modelling-based applications. The reason for its success over better approximations is the lower computational cost it entails. However, recent trends in seismic imaging rely on an improved physical model that represents the Earth no more as a rigid body but as an elastic and anisotropic one. Also, moving the wave propagation simulation closer to the real physics of the problem results in a significant increment of the needed computational resources. New hardware alternatives appear as a potential solution to satisfy the high demands of the computing power of the elastic, anisotropic, wave propagation engine. Also, the last decade has seen a trend on building systems with dedicated devices and accelerators, which produce a good FLOPs/Watt ratio. One of the most promising HPC alternatives comes from Intel® Phi™ product family.

This work shows several optimization strategies evaluated and applied to an elastic wave propagation engine, based on a Fully Staggered Grid, running on the latest Intel Xeon Phi processors, the second generation of the product (code-named Knights Landing). The developed propagator is able to reproduce elastic wave propagation, even for an arbitrary anisotropy.

A. Elastic FSG wave propagator algorithm

The Algorithm 1 describes our implementation of the FSG wave propagator using finite differences. It has been implemented from scratch inside YASK [1] framework using its domain-specific language (DSL), and it is now available as one of the example solutions included in the open-source package.[1]

B. The Intel Xeon Phi processor

For this study we are using a system with a CPU from the Intel Xeon Phi x200 processor family (code-named Knights

---

Algorithm 1 FSG wave propagator using finite differences

```
 1: function UPDATE_VELOCITIES
 2:     for all cells do
 3:         12 × density interpolations
 4:         12 × derivative calcs. based on stresses
 5:         12 × velocity updates
 6:     end for
 7: end function
 8: function UPDATE_STRESSES
 9:     for all cells do
10:         84 × coefficient interpolation
11:         28 × derivative calcs. based on velocities
12:         24 × stress updates
13:     end for
14: end function
15: for all timesteps do
16:     update_velocities ()
17:     update_stresses ()
18: end for
```

Landing). These processors have a multi-core architecture with up to 36 tiles per package connected through a 2D mesh between them and the memory controllers. Each tile is composed of two cores that share a 512 MB L2 cache and an agent that connects the tile to the mesh. Each core appears as four logical CPUs through hyper-threading. All hyper-threads on a core share a first-level (L1) cache. The cores are capable of issuing two instructions per cycle out-of-order, including vector and memory instructions. The Intel Xeon Phi architecture implements the Intel® Advanced Vector Extensions 512 (Intel® AVX-512) instruction set. To be able to provide the cores with enough data to feed the computing capabilities, the Intel Xeon Phi processors may be configured with an integrated on-package Multi-Channel DRAM (MCDRAM) memory of up to 16 GiB, which can deliver up to 490 GB/s of bandwidth. The main DDR4 memory on the same platform can deliver about 90 GB/s.

C. Evaluation Methodology

The set of optimizations applied to the code combine classic and widely used performance improvement approaches, such as: blocking, prefetching and scheduling, with

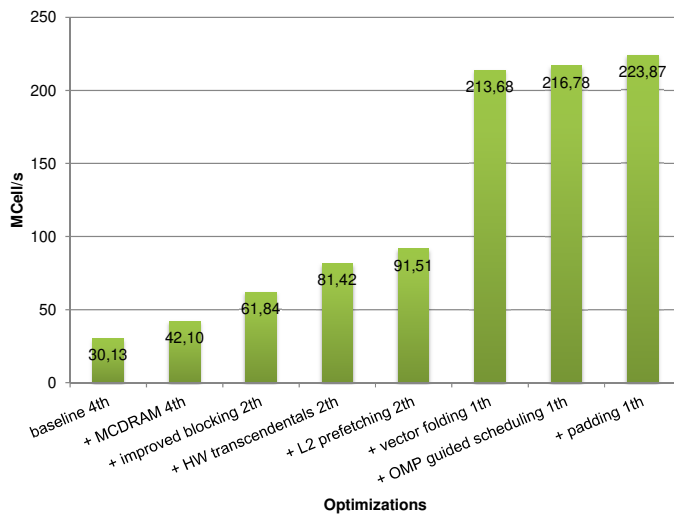---

[1] https://github.com/intel/yask

Fig. 1. FSG wave propagator throughput on Intel Xeon Phi

architecture-dependent optimizations like approximate reciprocal, and mechanisms from the state-of-the-art like vector folding. Additional details of the optimization methods are described below.

Baseline. The reference performance baseline was obtained by executing a fairly optimized version of FSG with parallelization using hyper-threading and blocking. This version uses the default compilation and run-time configuration provided by YASK. **MCDRAM memory.** Bandwidth-bound applications can easily benefit from the MCDRAM of the Intel Xeon Phi processors, exhibiting better execution times than with conventional DDR DRAM. **Improved blocking.** To improve temporal data reuse and reduce the memory bandwidth requirements per updated grid point, we implemented a loop blocking strategy transforming the memory domain of our problem into smaller chunks, rather than sequentially traversing through the entire memory domain. **Threads per core.** On Intel Xeon Phi processors is possible to place up to 4 threads per core. Reducing the numbers of threads per core to just 1 gave us the best performance. **Approximate reciprocal.** We approximated divisions using the reciprocal intrinsic instructions available in Intel AVX-512-ER which provide a faster implementation restraining the error. **Prefetching.** While the Intel Xeon Phi processor automatically prefetches data into the L1 and L2 caches it is possible to improve the memory behavior by using software prefetching. We have enabled aggressive software prefetching for L1 data cache. **Vector folding**. This method stores a small multi-dimensional block of data in each vector size memory region compared to the single dimension in the traditional approach. Therefore memory accesses required are reduced by increasing overlap between points for the stencil computation [1]. **OpenMP scheduling.** In most implementations the default loop scheduling algorithm in OpenMP is static, but it is not necessarily the best one. We evaluated the three most relevant loop scheduling strategies: static, dynamic and guided. The best performance in our case resulted from using dynamic scheduling. **Padding.** Padding improves performance by carefully aligning data in memory at the expense of using extra memory.

D. Results

Optimizations were enabled incrementally in the given order. Major changes, such as total number of threads used per core, or a different, sometimes additional optimization are clearly stated. To avoid potential outliers, all the experiments were executed in an exclusive node and repeated several times. The algorithm and the machine were found to be stable, showing only negligible variations between experiments. The throughput results of FSG are summarized in Fig. 1. It represents the maximum throughput obtained for each optimization and for the baseline implementation using DDR memory. The number of threads used in each case is shown after each configuration (e.g., "4th" indicates four threads per core). Baseline and MCDRAM ran with 4 threads per core; improved blocking, hardware trascendentals and L2 prefetching used 2 threads per core; and the remaining experiments ran with 1 thread per core.

E. Conclusion

We have shown a set of optimizations, applied to a Finite Difference Numerical method solving elastic wave propagation equations on the second generation Intel Xeon Phi processor. Moreover, the proposed scheme for solving the elastic equation supports arbitrary anisotropy at a higher computational cost when compared to more traditional approaches to address this problem. The evaluated set of optimizations ranges from memory to compute optimizations. Our findings indicated that, compared to a conventionally-optimized version, we were able to achieve $7\times$ more FLOPS and $8\times$ more bandwidth with all implemented optimizations. In consequence, we reached 75% of the maximum attainable FP performance at our current operational intensity.

II. ACKNOWLEDGMENT

REFERENCES

[1] C. Yount, "Vector folding: Improving stencil performance via multi-dimensional simd-vector representation," in *IEEE 17th International Conference on High Performance Computing and Communications (HPCC)*, Aug 2015, pp. 865–870.

[2] A. Farres, C. Rosas, M. Hanzich, A. Duran, and C. Yount, "Performance optimization of fully anisotropic elastic wave propagation on 2nd Generation Intel Xeon Phi processors," in *Proceedings of the 32nd IEEE International Parallel and Distributed Processing Symposium Workshops*, 2018, pp. 1033–1042.

Albert Farres is an engineer at Barcelona Supercomputing Center, the Spanish National Supercomputing Institute. He is currently researching and developing seismic imaging tools for the oil & gas industry. He has a MSc degree and a Bachelor in Computer Science from the Universitat Politècnica de Catalunya.

Deep Learning Phase Picking of Large-N Experiments

Luis Fernandez-Prieto[#1], Antonio Villaseñor[#2]

#*Earth's Structure and Dynamics and Crystallography, Institute of Earth Sciences Jaume Almera, C/Lluís Solé i Sabarís s/n 08028, Barcelona, Spain*

[1]`lmfernandez@ictja.csic.es`, [2]`antonio.villasenor@csic.es`

Keywords⸺ **Deep Learning, Phase Picking, Seismic Networks**

EXTENDED ABSTRACT

A. Introduction

The popularization of the use large-N arrays of seismometers has resulted in a significant increase of the size of the datasets recorded during these experiments. Therefore, new challenges have arisen on how to process all these data efficiently, and in an automated fashion. This is particularly true in the case of induced seismicity monitoring, where often a large number of number of events are recorded at high frequency sampling rates.

Several methods of automatic picking have been developed during recent years, from triggering algorithms (e.g. STA/LTA) to higher order statistics or waveform similarity. Latest development in computational power and the popularization of GPUs have made possible to apply machine learning methods to several problems, from arrival picking and phase detection to earthquake location.

B. Materials and Methods

We have developed a deep neural network to detect the arrivals of seismic body waves, using an architecture based on convolutional layers. This type of models are widely used in computer vision applications, which is the most similar case to the phase picking by an operator. The network was trained using the data of the Southern California Seismic Network, consisting in 4.8 Million seismograms labelled as P-waves, S-waves and pre-event noise. The data was divided into Training data (70%), Validation data (10%) and Test data (20%).

C. Results

This network is able to differentiate P and S waves from background noise with a precision of 98.3 %. The results obtained over the Test Data are shown in Table 1.

TABLE I
RESULTS OVER TEST DATA

	P-wave	S-wave	Noise
True positives	32.72 %	32.79 %	32.77 %
True negatives	66.13 %	66.13 %	66.01 %
False positives	0.56 %	0.51 %	0.66 %
False negatives	0.59 %	0.57 %	0.56 %

D. Conclusions

We have applied this neural network to other large-N experiments in other regions (Europe and Asia) and found that the network localizes the events with a precision comparable or superior to an human operator, even in the case of low signal-noise ratio and superposition of earthquakes. Also, the network is able to correctly identify non-seismic signals as noise, even if their amplitude is greater than the seismic events present.

E. ACKNOWLEDGEMENTS

References

[1] S. Chevrot and M. Sylvander, "Maupasacq" International Federation of Digital Seismograph Networks. Dataset/Seismic Network, 2017.

[2] Z. E. Ross, M. Meier, E. Hauksson, and T. H. Heaton, "Generalized Seismic Phase Detection with Deep Learning", Bulletin of the Seismological Society of America, Vol 108, No. 5A, Pp. 2894–2901, Sept. 2018.

[3] V. Salinas, A. Ugalde, A. Kamayestani, M. Jokar, M. M. Gharibvand, A. Villaseñor, and G. Heidari, "Designing and testing a network of passive seismic surveying and monitoring in Dehdasht (South Western Iran)", Geophysical Prospecting, Feb. 2019

[4] SCEDC, Southern California Earthquake Data Center, California Institute of Technology. Dataset, 2013.

Author biography

Luis Fernandez-Prieto was born in Madrid, Spain, in 1985. He received the degree in physics from the Universidad Complutense de Madrid, Spain, in 2015, and the master's degree in geophysics from the Universidad Complutense de Madrid, Spain, in 2018.
Since September 2018, he has been with the Department of Earth's Structure and Dynamics and Crystallography, in the Institute of Earth Sciences Jaume Almera, Barcelona, Spain. His current research interests include the study of machine learning algorithms, with applications in induced and natural seismicity.

A FM-index transformation to enable large k-steps

Rubén Langarita[*], Adrià Armejach[*†], Miquel Moretó[*†]

[*]Barcelona Supercomputing Center (BSC), Barcelona, Spain

[†]Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

E-mail: {ruben.langarita, adria.armejach, miquel.moreto}@bsc.es

Keywords—FM-index, sequence alignment, genomics.

I. EXTENDED ABSTRACT

The high demand for fast and low-cost genomic sequencing has pushed onward the rapid development of next-generation sequencing (NGS) technologies. These systems are able to produce huge amounts of short reads (in the order of giga base-pairs) per day of operation. Usually the first step in NGS corresponds to sequence alignment, where sequence reads must be aligned or compared to a genomic reference to identify regions of similarity. Most popular alignment methods are based on two types of index structures: suffix trees and hash tables [1]. When dealing with large reference genomes, great efforts were devoted to reduce memory requirements for sequence alignment. As a result, a set of alignment algorithms based on the FM-index (Full-text index in Minute space) structure have been developed.

The following section introduces the widely used FM-index technique. Then, we introduce our proposal, that uses a different data structure organization to enable search steps of longer symbols. Finally, we present our findings when evaluating our proposal on an Intel Knights Landing (KNL) machine.

A. Background - FM-index

FM-index is well suited for fast exact matches of short reads to large reference genomes while keeping a small memory footprint. The objective is to find short sequences of around 200 bases in a large reference genome. Normal problem sizes involve the human genome, which is around 3 gigabases. Therefore, to keep the memory footprint of these search algorithms manageable, FM-index uses the Burrows-Wheeler transform (BWT); a method for rearranging a character string that is useful for data compression.

The first step is to make all possible permutations of the reference genome, moving one letter from the start to the end each time (Figure 1). Then, we have to sort all the permutations alphabetically and extract the last column (Figure 2), usually also called BWT, which will be used to build the index.
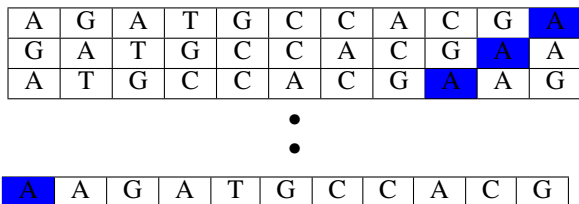

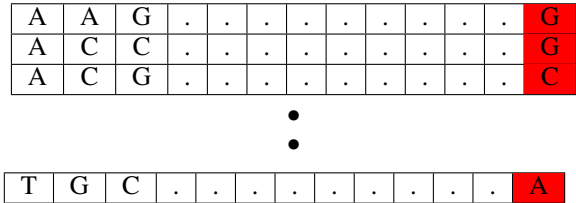Fig. 1. BWT step 1: Permutations.

Fig. 2. BWT step 2: Sort alphabetically and extract last column.

FM-index consists of two structures, called *C* and *Occ*. The structure *C* indicates where each symbol (letter of the alphabet) starts in the sorted permutation matrix. The Occ structure has a row for each entry of the BWT column, and one column for each symbol. Each entry of Occ contains the number of occurrences of each symbol until that point in the BWT column (Figure 3).
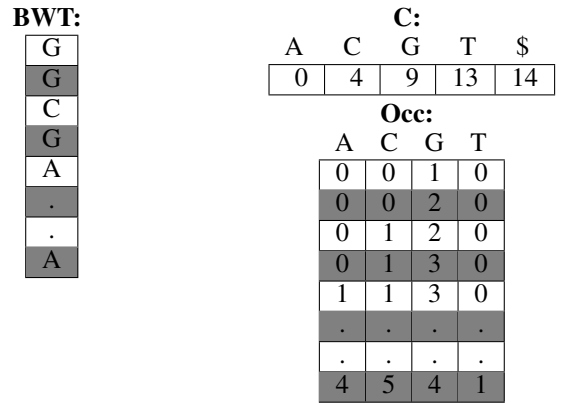

Fig. 3. *C* and *Occ* structures built from BWT.

FM-index uses two pointers (start pointer and end pointer) that indicate in each moment where are all the possible candidates in the sorted matrix. FM-index starts searching from the end of the search string. For the initialization, we have to read C and setup the pointers according to the first symbol read. Then, each iteration has to perform the following operations for each pointer: i) read the next symbol, ii) index C with that symbol, iii) index Occ with the symbol and the previous pointer and iv) add the entries of C and Occ and store the result in the pointer. This process repeats until all the sequence we are searching is read.

B. Changes - An Occ representation with constant size

A way to improve this algorithm is to increase the number of bases per symbol (K). For example, with K equals 2, the

alphabet would be: {AA, AC, AG, AT, CA, CC ... TT}; instead of {A, C, G, T}. Increasing K also means the size of the structures grows exponentially, since these are dependent on the number of symbols in the alphabet, which is 4^K. In order to maintain the size more or less constant while increasing K, we propose a novel method to store the entries of Occ.

Each entry of the BWT increases by one a single counter of the Occ table per row (see Figure 3). For example, with a K of 15 bases per symbol, only one column out of 4^{15} will change for each entry of BWT. Because of that, we propose to store the indexes when the counters change. We will call this new structure Changes. If the reference genome is of 3 gigabases, Changes would occupy 12 GB of memory no matter the K. The size of C does increase with K, as in the original FM-index proposal. For K equal 15, C would occupy 4 GB of memory, i.e., 4^{15} entries x 4 bytes/entry. C also indicates the offsets where each column in the Changes array starts (Figure 4).
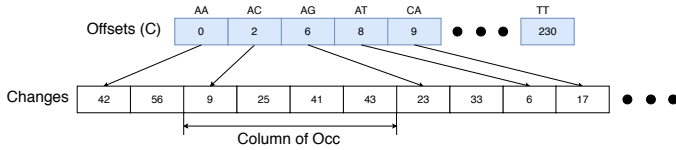

Fig. 4.   Proposal structures.

In our proposal, instead of indexing Occ, we would have to perform a search over the column of the symbol we are searching. Using Figure 4 as an example, suppose that we want to find the symbol "AC", and the previous pointers were 31 for start and 42 for end. First, we obtain the boundaries of the column by indexing C. That is 2 and 6. Then, we have to count the number of elements lesser than the pointers of the previous iteration in the column, that is 2 for start (9 and 25) and 3 for end (9, 25 and 41). Finally, we have to add the value obtained indexing C and the counted elements. That would be $2 + 2 = 4$ for start and $2 + 3 = 5$ for end. This process is repeated for the next symbol in the search string with start and end pointers 4 and 5, respectively. Once the entire search string is processed the start and end pointers identify the matches.

With K equals 15, we find that some symbols are repeated significantly, while others do not appear in the whole reference. This leads to columns that are large, while others are empty. The following histogram (Figure 5 left) shows the distribution of column sizes, i.e., repetition of patterns with K equals 15, for a reference human genome. The x axis is the column size and the y axis the number of occurrences of columns with that size. Figure 5 right shows the number of columns accessed during time execution multiplied by the column size, which to illustrates the time spent processing the different column types. This shows the impact of the large columns in terms of execution time if we perform sequential search. As can be seen the distribution is clearly non normal, as certain patterns are much more recurrent in nature than others. In order to improve search speed over the larger columns we perform binary search, as by definition our structure has sorted columns.

C. Evaluation

Recent work by J.M. Herruzo et. al [2] proposed an optimization of the FM-index algorithm based on sampling and
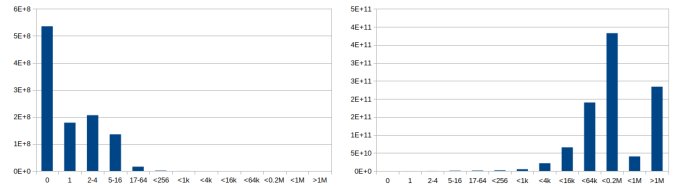

Fig. 5.   Number of columns for each size with K = 15. Left figure shows the number of columns of the original structure, while right shows the columns sizes accessed during execution weighted (multiply by the column size).

the use of bit vectors. They employ K equals 2 and sample 64 entries of the Occ by employing a 64 bit vector. We name this proposal from now on k2bv64. To evaluate this proposal they employ an Intel Knights Landing (KNL) system.

We use this as the baseline for comparison by evaluating our proposal using the same system setup also on a KNL machine. We employ the stacked MCDRAM in flat mode to map the data structures into this memory. In addition, the machine is configured to use 1GB huge pages and Hyper Threading, which means a total of 256 hardware threads are available.

For our proposal we use a k of 15, we will call this version k15inc. In order to measure the performance of the different versions we use a metric called "LF operations per second" (LFOPS). This measures the number of bases found per pointer. With a k of 2, each iteration performes 4 LFs, 2 for start and 2 for end, with a k of 15 would be 30 LFs... The best performe we obtained was 12.643 GLFOPS compare with the best version of the baseline 10.716 GLFOPS. We found that the algorithm performes better with sequences multiple of 15. In Figure 6 we can see the performance for the k2d64bv version and for our proposal with and without sequences multiple of 15.
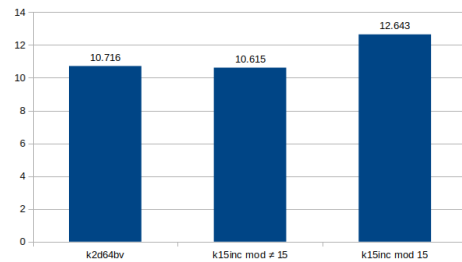

Fig. 6.   Performance for different versions.

D. Future work

We are planning to perform further space reduction optimizations over the C structure, as it is currently limiting the use of larger K values. After evaluating the pending optimizations on the KNL machine, we plan to evaluate the proposal in other systems that do not feature MCDRAM. In addition, we are interested to evaluate our proposal using vectorization, i.e., using the recently proposed vector extension (SVE) by Arm.

REFERENCES

[1]   H. Li and N. Homer, "A survey of sequence alignment algorithms for next-generation sequencing," *Briefings in bioinformatics*, vol. 11, no. 5, pp. 473–483, 2010.

[2] J. M. Herruzo, S. G. Navarro, P. Ibánez, V. V. Yufera, J. Alastruey, and O. Plata, "Accelerating sequence alignments based on fm-index using the intel knl processor," *IEEE/ACM transactions on computational biology and bioinformatics*, 2018.

Rubén Langarita received his BSc degree in Computer Science from University of Zaragoza in 2018. Now, he is doing a Master in Innovation and Research in Informatics at Universitat Politècnica de Catalunya (UPC), while working at Barcelona Supercomputing Center (BSC).

Development of HPC Multiphysics Framework for HTS Magnets in Fusion

José Lorenzo*, Octavio Castillo-Reyes*, Xavier Granados‡, Mervi Mantsinen*†, Xavier Sáez*, Alejandro Soba*

*Barcelona Supercomputing Center (BSC), Barcelona, Spain
‡Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Barcelona, Spain
†ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain
E-mail: jose.lorenzo@bsc.es

Keywords—High Temperature Superconductors, Fusion, Multiphysics, High Performance Computing, Edge Finite Element Method.

Extended Abstract

High-Temperature Superconductors (HTS) technology show promise for the next generation of fusion devices based on magnetic confinement, enabling higher magnetic fields and current densities, and thus a reduction of overall sizes and running costs due to lower heat load in the cryogenic system. Nevertheless, the use of HTS materials presents some drawbacks concerning quench protection and mechanical stability so that the application in large scale fusion magnets requires the detailed assessment of their thermal, electromagnetic (EM) and mechanical behavior. Computer modeling has played an important role to date in the design of superconducting magnets, but due to the lack of specific tools and to the inherent complexity of such structures, it is often needed either to study the involved phenomena separately or to simulate coupled multiphysics problems on rather simple geometries and models. Our goal is therefore to provide a new High Performance Computing (HPC) unified framework for the multiphysics simulation of HTS devices. The development is intended to build on an existing HPC multiphysics tool and will enable in the near future the design of the most suitable architectures of HTS cables for fusion applications considering mechanical robustness, losses, and quench propagation, detection and protection.

The present abstract is organized as follows: Section I briefly introduces the potential role of HTS technology in Fusion, Section II discusses the current state of simulation tools for HTS and our research objectives, and Section III describes the EM model used to develop a first prototype giving rise to first preliminary results.

I. HTS Cable Assemblies for Fusion Devices

HTS can withstand high current densities and magnetic fields with wide temperature margins compared to conventional Low-Temperature Superconductors (LTS) at the same operating temperature; therefore they emerge as a reasonable option to be considered in the design of large fusion magnets.

The usual HTS in the form of wires (Bi2212) and tapes (Bi2223, REBCO) can only carry a few hundreds of amperes in high field at low temperature. Since the current of ITER cables and several DEMO designs ranges from 30 kA to 100 kA [1], high current cables with several tens of superconductor wires or tapes need to be assembled in cables for its use in fusion magnets.

Several cable assemblies accounting for the flat geometry of thin REBCO tapes have been proposed for a wide range of applications: the stacked-tape conductor, the Conductor On Round Core (CORC) design [2], and the Roebel cable [3]. The latter two layouts present some features that are not suitable for high field applications [1], but cables based on stacks of REBCO tapes are being considered in some designs of the DEMO (DEMOnstration Fusion Power Station) [4, 5, 6, 7], a nuclear fusion power plant intended to follow the ITER experimental tokamak nuclear fusion reactor by 2050 [8].

II. State of the Art and Objectives

Since cabling layouts are currently under research [9] in order to optimize the mechanical, thermal and electromagnetic behavior, it is necessary to develop suitable support for magnets and cables design based on multiphysics computer-aided design.

At present, most of the simulation work for superconducting devices is carried out with a large number of in-house codes and commercial softwares due to the existence of various phenomena at different scales that make some tools more appropriate than others depending on the specific physics to be solved. This is further encouraged by the lack of tools able to address the whole picture at a competitive cost both in terms of time and budget.

The project objective is to deliver a HPC unified framework for the multiphysics simulation of large scale HTS devices that, in particular, can be applied to perform a comprehensive analysis of all the physical processes that occur during the operation of superconducting magnets for fusion. This will be achieved by enhancing the current capabilities of Alya multiphysics HPC platform [10] developed at Barcelona Supercomputing Center. More specifically, the project aims at adding to Alya a new electromagnetic module with HTS features that makes use of advanced HPC resources efficiently and that can be coupled to the already existing thermal and mechanical modules.

III. FEM Prototype and Preliminary Results

A first standalone prototype has been successfully tested against approximated analytical models, see Fig. 1. We have

implemented the 2-D H-formulation of Maxwell's equations to model the highly nonlinear electromagnetic behavior of superconductors,

$$\mu\,\partial_t \boldsymbol{H} + \nabla \times (\rho\,\nabla \times \boldsymbol{H}) = 0 \qquad \text{in } \Omega \times (0, T]\,,$$
$$\boldsymbol{n} \times \boldsymbol{H} = \boldsymbol{g}_D \qquad \text{on } \Gamma \times (0, T]\,, \quad (1)$$
$$\boldsymbol{H}\,(t = 0,\, \boldsymbol{x}) = \boldsymbol{H}_0\,(\boldsymbol{x}) \qquad \text{in } \Omega\,.$$

$\boldsymbol{H}\,(t, \boldsymbol{x})$ is the magnetic field intensity, μ the magnetic permeability, and ρ the electric resistivity given by a power law in superconducting materials,

$$\rho = \frac{E_c}{J_c}\left(\frac{\|\nabla \times \boldsymbol{H}\|}{J_c}\right)^{n-1}, \qquad (2)$$

where E_c is the critical electric field, J_c is the critical current density, and n is the exponent of the power law.
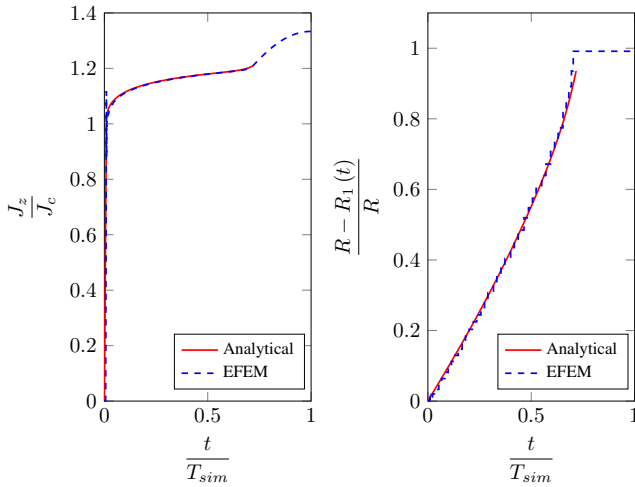

Fig. 1: Magnetization of a superconducting wire, $R = 1\,mm$, $I_c = 300\,A$, $E_c = 1\,\mu V/cm$, $n = 30$, transporting a time-varying current, $I\,(t) = I_0 \sin\,(2\pi f t)$.

Due to the mathematical properties of the problem, the Edge Finite Element Method (EFEM) [11, 12] has been preferred over the classical nodal FE approaches. EFEM is the common approach in the field of numerical modeling for applied superconductivity [13, 14], and will be adopted in the near future to study 3-D geometries and the foreseen coupling with the thermal model.

REFERENCES

[1] P. Bruzzone, W. H. Fietz, J. V. Minervini, M. Novikov, N. Yanagi, Y. Zhai, and J. Zheng, *Nucl. Fusion*, no. 10, p. 103001.
[2] J. D. Weiss, T. Mulder, H. J. ten Kate, and D. C. van der Laan, *Superconductor Science and Technology*, vol. 30, no. 1, p. 014002, 2016.
[3] W. Goldacker, F. Grilli, E. Pardo, A. Kario, S. I. Schlachter, and M. Vojenčiak, *Superconductor Science and Technology*, vol. 27, no. 9, p. 093001, 2014.
[4] R. Heller, P. V. Gade, W. H. Fietz, T. Vogel, and K. P. Weiss, *IEEE Trans. Appl. Supercond.*, vol. 26, no. 4, pp. 1–5, 2016.
[5] W. H. Fietz, R. Heller, M. J. Wolf, and P. V. Gade, *Fusion Eng. Des.*, no. February 2016, pp. 290–293.
[6] R. Wesche, N. Bykovsky, X. Sarasola, K. Sedlak, B. Stepanov, D. Uglietti, and P. Bruzzone, *Fusion Eng. Des.*, vol. 124, pp. 82–85, 2017.
[7] R. Wesche, X. Sarasola, K. Sedlak, N. Bykovsky, B. Stepanov, D. Uglietti, and P. Bruzzone, *IEEE Trans. Appl. Supercond.*, vol. 28, no. 3, pp. 1–5, 2018.
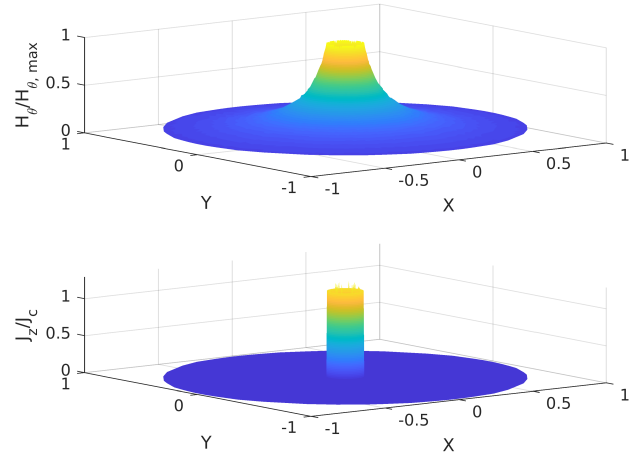[8] F. Romanelli, P. Barabaschi, D. Borba, G. Federici, L. Horton, R. Neu, D. Stork, and H. Zohm, 2012.

Fig. 2: Magnetic field and current distribution in a superconducting wire surrounded by air.

[9] N. Bykovskiy, Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2017.
[10] M. Vázquez, G. Houzeaux, S. Koric, A. Artigues, J. Aguado-Sierra, R. Arís, D. Mira, H. Calmet, F. Cucchietti, H. Owen, A. Taha, E. D. Burness, J. M. Cela, and M. Valero, *J. Comput. Sci.*, vol. 14, pp. 15–27, 2016.
[11] J.-M. Jin, *The finite element method in electromagnetics*. John Wiley & Sons, 2015.
[12] P. Monk *et al.*, *Finite element methods for Maxwell's equations*. Oxford University Press, 2003.
[13] R. Brambilla, F. Grilli, and L. Martini, *Supercond. Sci. Technol.*, vol. 20, no. 1, pp. 16–24, 2007.
[14] G. G. Sotelo, M. Carrera, J. Lopez-Lopez, and X. Granados, *IEEE Trans. Appl. Supercond.*, vol. 26, no. 8, 2016.

José Lorenzo received his MSc degree in Industrial Engineering from Universidad de Sevilla, Spain, in 2014. In collaboration with CIEMAT (Madrid, Spain), he joined the Magnets, Superconductors and Cryostats Group at CERN (Geneva, Switzerland), where he worked in the analysis and numerical modeling of quench in Nb$_3$Sn cables from 2015 to 2017. Later, he has worked at Universidad Autónoma de Madrid, Spain, in the implementation of control for Partial Differential Equations, shape optimization for Fluid Dynamics and mesh morphing techniques.

Techniques for reducing and bounding OpenMP dynamic memory

Adrian Munera*†, Sara Royuela*†, Eduardo Quiñones*

*Barcelona Supercomputing Center, Barcelona, Spain
†Universitat Politècnica de Catalunya, Barcelona, Spain
E-mail: {adrian.munera, sara.royuela, eduardo.quinones}@bsc.es

Abstract—**OpenMP offers a tasking model very convenient to develop critical real-time parallel applications by virtue of its time predictability. However, current implementations make an intensive use of dynamic memory to efficiently manage the parallel execution. This jeopardizes the qualification process and limits the use of OpenMP in architectures with limited amount of memory. This work introduces an OpenMP framework that statically allocates the data structures needed to efficiently manage parallel execution in OpenMP programs. We achieve the same performance than current implementations, while bounding and reducing the dynamic memory requirements at runtime.**

Keywords—*OpenMP, dynamic memory, critical-systems.*

I. INTRODUCTION

Current embedded architectures for critical real-time systems such as the Kalray MPPA [1], and the TI Keystone II [2] deliver levels of parallelism that can efficiently by exploited by means of parallel programming models. In fact, these architectures include support for OpenMP, a well-known parallel programming model from the high-performance domain, in their software development kits.

There has been a significant effort to evaluate the time predictability of the OpenMP tasking model, which enables exploiting fine-grain and highly dynamic parallelism. These works are based on the extraction of a Direct Acyclic Graph (DAG) representing the parallel execution of an OpenMP program [3], and upon which timing and schedulability analyses can be applied for both dynamic [4] and static [5] scheduling approaches. Furthermore, the functional safety of the OpenMP specification has also been analyzed [6].

However, current OpenMP implementations require complex data structures to efficiently orchestrate the parallel execution. For example, in *libgomp* [7], the OpenMP runtime of GNU, these include: (1) a structure per task instance with the code to execute and the data environment; (2) a *hash table* to store the elements in the dependence clauses (i.e., `out` and `inout`), and the list of tasks depending on them, so when a task is created its dependences (i.e., `in` and `out`) are matched against those of the existing tasks; and (3) several linked lists to quickly identify the tasks whose dependences are honored when a task finishes.

Previous works reduce the complexity of the structures in the OpenMP runtime by statically storing the complete *task dependency graph* (TDG) [3]. Nonetheless, the runtime still makes an intensive use of dynamic memory due to: (a) dynamically capturing the data environment, and (b) the task creation policy (which can produce a number of task structures proportional to the total number of tasks of the application).

We have developed an OpenMP framework that pushes the creation of dynamic task data structures to the initialization phase, efficiently executes an OpenMP program, and reduces the amount of memory necessary to exploit parallelism. The framework is composed of: (a) a compiler that expands the complete TDG of an OpenMP application and computes the amount of parallelism exposed in the application to statically allocate the data structures needed be the runtime depending on the task creation policy; and (b) a runtime that implements different policies regarding the moment at which tasks are created and initialized.

Our evaluation shows that we provide the same performance speedup than the one provided by the OpenMP framework included in the SDK of the targeted processor architecture, while significantly reducing the memory requirements.

II. PREALLOCATION AND TASK CREATION POLICIES

With the objective of reducing and bounding the amount of memory needed at runtime, our framework allows different policies with regard to the static allocation of task data structures and the moment at which tasks are created.

Task creation policies. Current OpenMP runtimes usually create tasks as they are encountered. However, these tasks may not be ready for execution due to unmet dependences, in which case they will be using dynamic memory even though they cannot yet be executed. We have implemented a *lazy* policy in libgomp based on a proposal to "un-inline" tasks in a parallel Lisp system [8]. In our case, tasks are not un-inlined, but their creation is pushed until the task is ready to be executed. This means that the amount of on-the-fly memory used to manage task structures is bound by the maximum amount of parallelism exposed in the TDG.

Preallocation policies. We have implemented a series of analysis and optimizations in the Mercurium compiler to expand the complete TDG of an OpenMP application. This enables the static allocation of the structures needed to manage tasks at runtime following two different approaches:

– Allocate as many structures as tasks in the TDG. This reduces the use of dynamic memory, although the total amount of memory needed remains the same.

– Allocate as many structures as parallelism is exposed in the TDG by using a Matlab based algorithm that computes the maximum degree of parallelism of the TDG. This drastically reduces the total amount of memory required by the application.
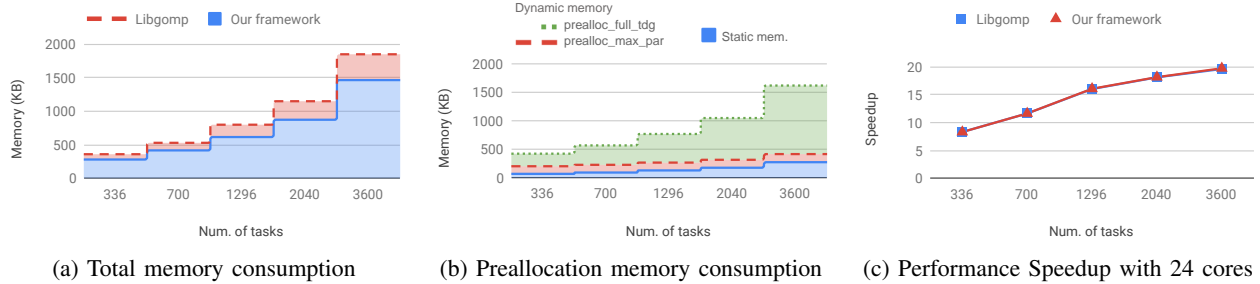
(a) Total memory consumption   (b) Preallocation memory consumption   (c) Performance Speedup with 24 cores

Fig. 1: Runtime evaluation for the HoG application depending on the number of tasks.

III. EVALUATION

A. Experimental Setup

Processor Architecture. Intel Xeon Platinum 8160 [9] featuring 24 cores.

Framework. GCC 7.3, implementing OpenMP4.5 [10].

Application. A pedestrian detector from the real-time embedded domain based on the computation of histogram of oriented gradients (HoG), available in the VLFeat library.

B. Performance speedup

Figure 1c shows the performance of the native OpenMP framework (labeled as libgomp) and our framework preallocating only the required amount of parallelism and using lazy tasks creation, for the HoG application. The speedup of the two versions is almost the same when using 24 cores (considering the average time of 50 executions with the same data input).

C. Memory usage

Figure 1a shows the maximum memory space (in KBs) required to execute the HoG application with libgomp, compared with our framework without preallocation or lazy tasks. Here, our framekork reduces the total amount of memory used due to the static management of task dependences, which is more memory efficient than the libgomp hash table implementation.

The lazy tasks policy allows preallocating only as many task structures as parallelism is exposed in the TDG. Figure 1b shows the memory consumption of the HoG application with our framework, for several numbers of tasks. The lazy task creation policy reduces drastically the use of dynamic memory, compared to the full allocation of the TDG. Also, the static memory usage increments with the number of tasks, as task dependences are stored statically by the compiler.

Our preallocation mechanism calls a runtime routine to do the allocation at initialization phase. In the future, we will use static structures, eliminating the use of dynamic memory.

IV. CONCLUSION

Despite the proven benefits of OpenMP to develop critical real-time applications, current implementations are not suitable due to (a) the potentially high amount of memory used to hold tasks, and (b) the intensive use of dynamic memory to allocate task data structures. This limits the use of OpenMP in systems where memory is restricted, and also jeopardizes the qualification for critical real-time systems.

Our work allows limiting the amount of memory required at runtime to handle on-the-fly tasks while showing the same

performance as the original runtime. Furthermore, our framework can statically allocate the data structures, so the amount of dynamic memory used by the runtime can be drastically reduced, hence shortening the gap towards the qualification of OpenMP frameworks for critical real-time systems.

V. ACKNOWLEDGMENT

REFERENCES

[1] B. D. De Dinechin, D. Van Amstel, M. Poulhiès, and G. Lager, "Time-critical computing on a single-chip massively parallel processor," in *DATE*, 2014.

[2] Texas Instruments, *66AK2Hxx Multicore Keystone II System-on-Chip (SoC)*, 2012. [Online]. Available: www.ti.com/product/66AK2H12

[3] R. E. Vargas, S. Royuela, M. A. Serrano, X. Martorell, and E. Quiñones, "A lightweight openmp4 run-time for embedded systems," in *ASP-DAC*, 2016, pp. 43–49.

[4] M. A. Serrano, A. Melani, R. Vargas, A. Marongiu, M. Bertogna, and E. Quiñones, "Timing characterization of openmp4 tasking model," in *CASES*, 2015, pp. 157–166.

[5] A. Melani, M. A. Serrano, M. Bertogna, I. Cerutti, E. Quiñones, and G. Buttazzo, "A static scheduling approach to enable safety-critical OpenMP applications," in *ASP-DAC*, 2017.

[6] S. Royuela, A. Duran, M. A. Serrano, E. Quiñones, and X. Martorell, "A functional safety openmp* for critical real-time embedded systems," in *IWOMP*, 2017, pp. 231–245.

[7] GNU, "libgomp," 2018. [Online]. Available: gcc.gnu.org/onlinedocs/libgomp/

[8] E. Mohr, D. A. Kranz, and R. H. Halstead, "Lazy task creation: A technique for increasing the granularity of parallel programs," *TPDS*, vol. 2, no. 3, pp. 264–280, 1991.

[9] Intel, "Intel Xeon Platinum 8160," 2018. [Online]. Available: ark.intel.com/products/120501/Intel-Xeon-Platinum-8160-Processor-33M-Cache-2-10- GHz-

[10] OpenMP ARB, "OpenMP 4.5 Specification," 2015. [Online]. Available: www.openmp.org/wp-content/uploads/openmp-4.5.pdf

Adrian Munera received his BSc in Computer Engineering (Computer Architecture specialization) from Universitat Politècnica of Valencia (UPV) in 2018. He is finishing his MSc in Innovation and Research in Informatics at Universitat Politècnica de Catalunya (UPC), developing his thesis in the field of OpenMP and real-time systems. He is also working as a research student at BSC in the CAOS (Computer Architecture - Operating Systems) group.

A multilayer network approach to elucidate severity in Congenital Myasthenic Syndromes

I. Núñez-Carpintero[*], D. Cirillo[*], A. Valencia[*]

[*]*Barcelona Supercomputing Center (BSC)*

E-mail {iker.nunez, davide.cirillo, alfonso.valencia}@bsc.es

EXTENDED ABSTRACT

A. Introduction

Congenital Myasthenic Syndromes (CMS) conform a group of diverse rare diseases characterized by neuromuscular junction (NMJ) dysfunctions resulting in muscle weakness, which is a common hallmark in these conditions. While causative genes have been previously described , a molecular explanation of the observed large range of phenotypic severity remains unclear.

In this work we undertook a Personalized Medicine approach with the goal of elucidating the molecular processes causing severity in CMS. Personalized Medicine is defined as the integration of molecular research with clinical data in order to deliver better diagnoses and treatments tailored to the individual characteristics of each patient.

We jointly analysed DNA and RNA from 20 samples from an isolated region in Eastern Europe (Bulgaria), integrating RNA sequencing (RNAseq) from fibroblast culture and Whole Genome Sequencing (WGS) from blood, identifying a set of candidate genes. Since none of these genes' features (annotations, mutations, etc.) were able to explain severity , we performed a multilayer network analysis of several aspects (interactome, reactome and metabolome) involving the damaged genes of each CMS patient in our cohort.

B. Methods

1 – RNAseq & WGS

Whole genome sequencing (WGS) data have been obtained from blood, while RNA sequencing (RNA-seq) data have been obtained from fibroblasts. All analyses have been performed using RD-Connect project (https://rd-connect.eu/) specific pipeline and the human genome GRCh37d5 as reference.

Copy Number Variants (CNVs) have been extracted using ClinCNV (https://github.com/imgag/ClinCNV) by employing a set of Eastern European samples as background control group. Heterozygous compound mutations have been obtained by phasing the WGS and RNA-seq data, removing variants with allele frequency > 3%, outsde exonic and splicing regions (Ensembl annotation), synonymous, and with read depth (coverage) smaller than 8.

2 – Multilayer Community Detection

We generated a multilayer network using reactome [2], the metabolome [3], and the interactome [4] layers. The multilayer community detection analysis was performed by using MolTi software [5]. To study the relationship between disease-associated genes multilayer network communities' membership, we analyzed the curated gene-disease associations of the DisGeNET database (http://www.disgenet.org/).

C. Results

Albeit all affected individuals shared the very same causal mutation (a deletion within the acetylcholine receptor -AChR- ε subunit, CHRNE c.1327delG), the severity of symptoms across this cohort varies considerably regardless of age and sex, pointing towards causes to the disease phenotype other than the causal variant. In this work, we performed an in-depth characterization of 20 CMS samples from patients from this cohort by analyzing multi-omics data (see Methods: *RNAseq & WGS*). Distinct CMS severity levels have been classified by specialized physicians , namely severe (8 patients) and not-severe (2 intermediate and 10 mild patients) disease phenotypes. In particular, the clinical phenotyping reports the outcomes of medical tests such as swallowing, speech, respiratory dysfunction, as well as the prescription of pharmacological treatments.

We jointly analyzed DNA and RNA from 20 patients, with the objective of validating whether the severity was determined by the accumulation of damaging mutations hampering the neuromuscular activity, on top of the CHRNE damaging mutation. By analysing segregating single nucleotide polymorphisms (SNPs) and copy number variations (CNVs), we could not find any unique cause of severity. Nevertheless, we observed that heterozygous compound mutations are enriched in pathways related to the extracellular matrix (ECM) receptor, which has been recently proposed as a target for CMS therapy [1]. As a result, we used CNVs and heterozygous compound mutation gene sets in the multilayer network analysis (**Figure 1**).
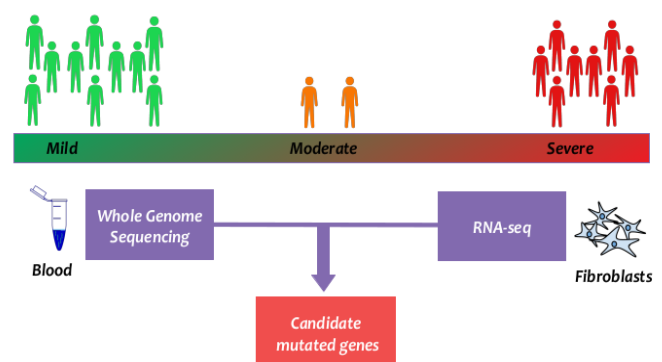

Fig. 1: Phenotypic severity spectrum and genomic work performed on the CMS cohort.

Once assessed that compound heterozygous variants are associated with functionally related pathways, we sought to analyze the relationships among the affected genes at different scales . We generated a multilayer network, a complex system representation in which distinct features of nodes and edges are described by several layers, containing different facets of molecular biology (**Figure 2**), namely the reactome [2], the metabolome [3], and the interactome [4].

We performed a multilayer community analysis using MolTi, a software that adapts classical Louvain clustering analysis into multilayer networks [5] and looked for topological relationships between the affected genes and previously known CMS causal genes [6].

We identified a largest component module of damaged genes that is specific of the severe group of patients, composed of 15 genes **(Figure 2)**. 6 out of these 15 are previously described CMS causal genes, while the other genes are mutated with damaging compound heterozygous mutations (9 out of 15, as agrin, a well-known CMS causal gene present this type of mutations in one of the patients) or CNVs (1 out of 15). The mutated genes are involved in a varied spectrum of functions at the neuromuscular junction, such as neuromuscular synapsis development and AChR clustering at the skeletal muscle fiber. Our results show that alterations in proteoglycans (AGRN, HSPG2, VCAN, COL15A1), tenascins (TNC, TNXB), and chromogranins (CHGB) are specific of the severe group. Strikingly, no proteoglycans are damaged in the not-severe group, suggesting a direct involvement of ECM in CMS severity.
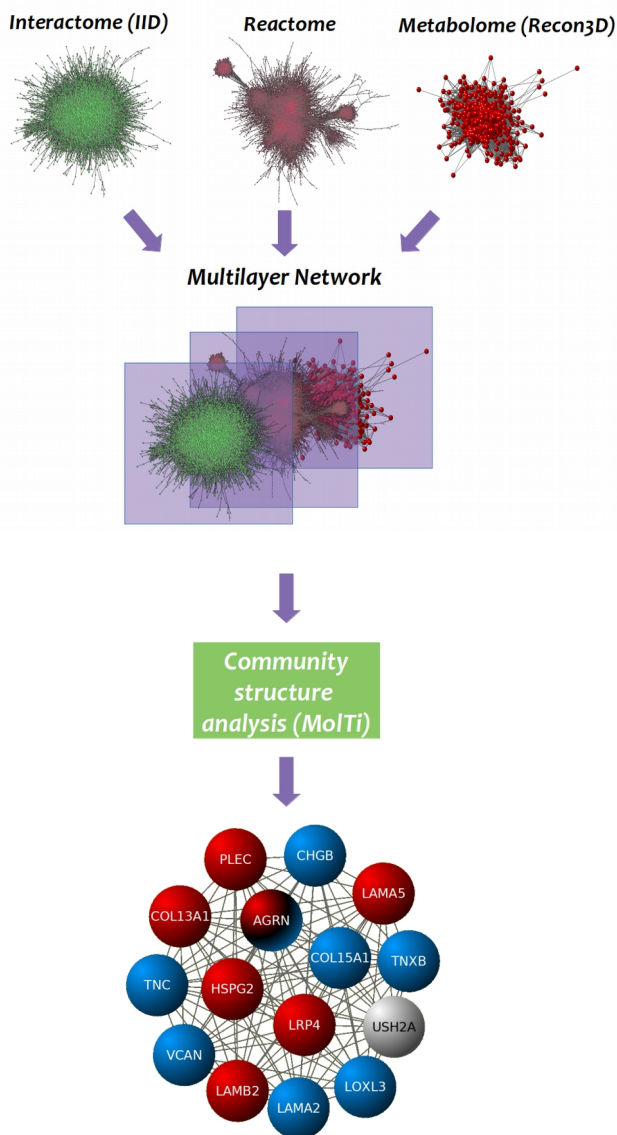

Fig. 2: Multilayer network components, community analysis workflow and the identified largest component module of genes functionally related to neuromuscular junction function.

In red, previously known CMS genes. In blue, genes mutated with compound heterozygous mutations. In white, genes mutated with copy number variants (CNVs).

D. Discussion

In this work, we tested the hypothesis that the severity of a rare disease with a known causative mutation was determined by the accumulation of additional damaging alterations hampering a specific impaired functional process. In particular, we performed an in-depth analysis of 20 CMS patients, from a narrow and geographically isolated population, who share the same causative mutation in the acetylcholine receptor ε subunit (CHRNE), presenting different levels of disease severity, as assessed by expert physicians.

Our results show that CMS severity can be ascribed to the personalized impairment of specific classes and localizations of NMJ activities, namely extracellular matrix components (proteoglycans, tenascins, chromogranins) and postsynaptic modulators of AChR clustering. Moreover, this work shows that coupling multilayer network analysis with personalized omics information helps give a molecular explanation of the phenotypic severity of rare diseases.

Finally, as our results suggest that severity is related to AChR clustering at the AGRN-PLEC-LRP4-Laminins axis level, severe affected patients may potentially benefit from pharmaceutical interventions enhacing the AChR clustering process. For example, beta-2 adrenergic receptor agonists like ephedrine and salbutamol have been documented as capable of enhancing AChR clustering and proved to be successful in the treatment for severe AChR deficiency syndromes [7].

References

[1]   Ito M & Ohno, K. *Protein-Anchoring Therapy to Target Extracellular Matrix Proteins to Their Physiological Destinations.* Matrix Biol. 2018, 68-69, 628-636.

[2]   Fabregat et al. *The Reactome Pathway Knowledgebase.* Nucleic Acids Res. 2018 46 (D1): D469-55

[3]   Brunk et al. *Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism.* Nat. Biotechnol 36 (3): 272-81

[4]   Kotlyar et al. *Integrated Interactions Database: Tissue-Specific View of the Human and Model Organism Interactomes".* Nucleic Acids Res 2016. 44 (D1): D536-541

[5]   Didier et al. . *Identifying communities from multiplex biological networks.* PeerJ 2015. 3, e1525

[6]   Bonne et al. *Diagnostic Approach to the Congenital Muscular Dystrophies.* Neuromuscul. Disord. 2017. 27(12): 1152-83

[7]   Rodríguez Cruz et al. *Salbutamol and Ephedrine in the Treatment of Severe AChR Deficiency Syndromes.* Neurology 85(12): 1043-47

Author biography

Iker Núñez was born in Valladolid, Spain, in 1995. He received the BSc degree in Biology from Universidad de Alcalá (UAH), Spain, in 2017, and the MSc degree in Biomedical Research from Universidad de Valladolid (UVA) and Molecular Biology and Genetics Institute (IBGM) of Valladolid, Spain, in 2018. Since October 2018, he is a Ph.D. student at Computational Biology Group within the Life Sciences Department of Barcelona Supercomputing Center (BSC), Spain. His current main research interests includes Network Analysis, Deep Learning and Personalized Medicine.

C/R Support for Heterogeneous HPC Applications

Konstantinos Parasyris, Leonardo Bautista Gomez

Barcelona Supercomputing Center (BSC)

E-mail: {konstantinos.parasyris,leonardo.bautista }@bsc.es

Keywords—Fault Tolerance, High-performance computing, Reliability

I. Extended Abstract

As we approach the era of exa-scale computing, fault tolerance is of growing importance. The increasing number of cores as well as the increased complexity of modern heterogenous systems result to substantial decrease of the expected mean time between failures. Among the different fault tolerance techniques, checkpoint/restart it is vastly adopted in supercomputing systems. Although, many supercomputers in the TOP 500 list use GPUs, only a few checkpoint restart mechanism support GPUs.

In this paper, we extend an application level checkpoint library, called fault tolerance interface (FTI), to transparently support multi-node/multi-GPU checkpoints. Upon a checkpoint invocation the extended library tracks the actual location of the data to be stored and handles the data accordingly. When data are stored in the GPU side, to hide the extra latencies, we overlap the copying of device memory to host memory with the writing of the data to the checkpoint file.

II. Importance of Fault Tolerance

The last decades the supercomputers have increased in size and computing cababilities. Exascale computing is the next objective, which will bring even more computing power to scientific applications and to industries. However several challenges are raised with exascale computing, the most important ones are the power consumption and the error resiliency. As the number of components increase in large scale systems, the systems become more error prone, and thus more prone to failures. It is expected that the next generation of high performance computing machines will experience failures up to several times an hour, making the need for effective fault resilience effective for building tomorrows HPC systems [1].

Another important consideration for the fault resiliency of extreme-scale HPC systems is the increasingly heterogenity of the components within the system. The future exascale will consit of multiple nodes with each node consisting of a high-performance system that combines a balance of high throughput general-purpose (GPGPU) pipelines for extreme high performance, coupled with high performance multicore CPUs targeting single-thread performance. The GPUs provide the high throughput required for exascale levels of computation, whereas the CPU cores handle hard-to-parallelize code sections and provide support for legacy applications. However, GPUs are more error prone than CPUs. In *TSUBAME* 40% of the total number of failures are caused by GPU errors, however the number of CPU related failures is below 5% [2]. When injecting faults [3] to a applications executing on CPUs, only 2.3% of the injected faults manifest as silent data corruptions, in GPUs this percentage rises to 16-33%. For all these reason the mean time between failures (MTBF) is expected to decrease even more in future systems.

To overcome failures, supercomputers typically use checkpoint restart techniques, by storing the state of the computation in reliable storage. Upon a failure, the most recent state is used to restart the computation. Unfortunately, the amount of data to be checkpoint increase, since HPC applications nowadays are able to process more information. On the one hand, the decrease of the MTBF results to higher checkpoint frequency to reduce the amount of re-computation. On the other hand, the increase of the data to store, increase the overhead of the checkpoint procedure. To make things even worse, typically, in GP-GPU HPC applications portions of the application data are stored in the CPU-memory, whereas other portions are stored in the GPU main memory. This distribution of data increases the overhead of the checkpoint procedure. For all these reasons, checkpoints reduce heavily the system's efficiency. To maintain high productivity in supercomputers and large data centers, it is important to: i) reduce the programmers effort to implement checkpoints ii) reduce as much as possible the data to be stored iii) reduce the total overhead of the checkpointing procedure.

We have extended a checkpoint library called Fault Tolerance Interface (FTI) [4] to support transparent checkpoint of data stored in differen CUDA-enabled GPU devices. Our method does not extend the library's API, but automatically tracks the physical memory location of user defined virtual addresses. The functionality transparently handles CPU, GPU as well as unified memory addresses. The methodology supports all the checkpointing techniques of the library, namely incremental checkpointing, hash based differential checkpointing and normal checkpointing.

III. FTI GPU/CPU implementation

FTI handles checkpoints in three different phases. The first, called initialization-phase, corresponds to the initialization of the library and the definition of the protected memory regions. The second layer, (Checkpoint/Restore(C/R))-phase, corresponds to the actual C/R procedure, hence it moves the data from the device and host memory to the stable local storage device (SSD, NVMe) and vice versa in the case of the recovering. When the checkpoint-phase is terminated, the application resumes the normal execution, and the async-phase starts. In the async-phase the FTI managers start in the background encoding and transferring the checkpoint files to the respective checkpoint level.

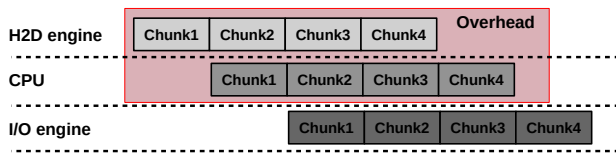To support Hybrid GPU/CPU support to FTI we extend

Fig. 1. Checkpoint communication scheme to overlap data transfers.

Fig. 2. Execution time of the checkpoint procedure for the configurations

the initialization-phase and the C/R-phase. In the initialization-phase we identify the physical location of the address. The library through an FTI-call (*FTI_Protect*) identifies the physical location of the data. This is done through the CUDA driver API support, namely the function *cudaPointerGetAttributes(&attributes, address)*. The function raises an error when called with a host address, whereas it returns normally with a device or an UVA address. In the second case we further check the values of *attributes* field which provides information whether the address is UVA or not. In the end we tag each address as *CPU,GPU,MANAGED*.

During C/R phase, depending on the tag of each address we perform a different action. In the case of CPU or UVA addresses, we invoke the normal FTI C/R procedure. In the case of UVA addresses we use the CUDA driver to fetch the data from the GPU and move them to the stable local storage. Finally, in the case of *GPU* addresses, we overlap the writing of the file with the data movement from the GPU side to the CPU side. This is done through streams and asynchronous memory copies of chunks from GPU memory to host pinned memory. The procedure of transferring data from the GPU memory to the CPU is depicted in detail in Figure 1. Each protected memory region is divided into smaller blocks. The size of the block, from now on called communication block (*cBlock*), is controlled through a configuration option. The CPU requests from the Host to Device (H2D) engine an asynchronous transfer of the first *cBlock*, when the *cBlock* is copied to the host memory, the CPU requests the next *cBlock* and starts performing the necessary actions with the current cBlock. The main actions are the following: i) Update the checkpoint integrity checksum ii) Copy the *cBlock* to the I/O layer through the respective I/O library call.

When the data are copied to the I/O layer the CPU starts processing the next chunk, which ideally should already be copied in the host memory. The application process does not wait for the I/O operations to finalize, when all data are moved to the I/O layer it informs the FTI-managers to start the background actions and resumes the user code execution. The described scheme is optimal only if the execution time to compute the integrity checksum and copy the *cBlock* to the I/O layer is equal to the execution time needed to copy the data from the device to the host. In any other case, either the H2D engine or the CPU is idle.

A. Evaluation

In this section we analyze the FTI GPU checkpoint scheme. We use a micro-benchmark for profiling and analysis purposes. The micro-benchmark checks the strong/weak scaling of our approach using different mixtures of device/host memory allocations. The micro-benchmark allocates two memory buffers, the first buffer, called *hBuff*, is allocated on the host memory, whereas the second one, called *dBuff*, is allocated on the device
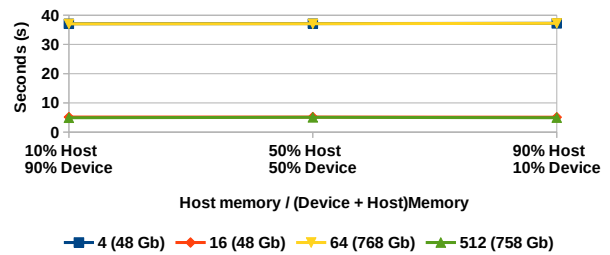
memory. The size of each memory buffer is user defined. The application protects these two buffers and performs a checkpoint every 5 minutes. To test both the weak and the strong scaling of our approach we execute each benchmark on 4 different node/process mappings. Specifically, we executed experiments using 1 and 16 nodes, with 4 and 32 processes in each node. The checkpoint size of each node is 48 Gb regardless the number of processes. On each of these mappings we execute 3 different memory allocation schemes. In the first scheme, for each process we allocate 10% memory on the host and the remaining 90% to the device memory. The second scheme allocates 50-50% on the respective memories and the final scheme allocates 90-10%. In Figure 2 we depict the results of the executed experiments the X-axis represents the memory schemes, the Y-axis the amount of time spend by the user process to perform a single checkpoint and the different lines correspond to the different node/process mappings.

IV. CONCLUSIONS

Interestingly, regardless the actual memory location of the checkpoint data the checkpoint overhead remains the same. The implementation demonstrates nice weak and strong scaling. We profiled the execution time of the checkpoint procedure. The communication between the GPU and the CPU is completely overlapped, therefore there is almost no-overhead to move the data from the GPU to the CPU. Interestingly, the execution time is mainly spend in computing the integrity checksum of the checkpoint file. We plan to move the computation of the checksum on the GPU and overlap it with the actual writing of the C/R file. This will dramatically decrease the execution time of the checkpoint procedure.

REFERENCES

[1] F. Cappello, "Fault tolerance in petascale/ exascale systems: Current knowledge, challenges and research opportunities," *The International Journal of High Performance Computing Applications*, vol. 23, no. 3, pp. 212–226, 2009.

[2] B. Pourghassemi and A. Chandramowlishwaran, "cudacr: An in-kernel application-level checkpoint/restart scheme for cuda-enabled gpus," in *2017 IEEE International Conference on Cluster Computing (CLUSTER)*, Sep. 2017, pp. 725–732.

[3] K. S. Yim, C. Pham, M. Saleheen, Z. Kalbarczyk, and R. Iyer, "Hauberk: Lightweight silent data corruption error detector for gpgpu," in *2011 IEEE International Parallel Distributed Processing Symposium*, May 2011, pp. 287–300.

[4] L. Bautista-Gomez, S. Tsuboi, D. Komatitsch, F. Cappello, N. Maruyama, and S. Matsuoka, "Fti: High performance fault tolerance interface for hybrid systems," in *SC '11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov 2011, pp. 1–12.

Konstantinos Parasyris is a Postdoctoral researcher at the Barcelona Supercomputing Center where he works on Resilience and Optimization. He received the B.Sc., M.Sc., and Ph.D. at the electrical and computer engineering at the university of Thessaly (Greece). His research is focused on fault tolerance and application level error resiliency. During his PhD he developed GemFI a simulation based fault injection tool as well as XM2 a hardware level fault injection. He has been involved in FP7 ScoRPiO and H2020 Uniserver European projects. In ScoRPiO he actively participated in the development of a task-based approximate programming model, whereas in Uniserver he participated on dynamically identifying the voltage margins of modern processors for reliable operation.

Deciphering the interactions between the immune system and cancer cells to enable precision medicine

Victoria Ruiz-Serra*, Eduard Porta-Pardo*,
*Barcelona Supercomputing Center (BSC)
E-mail: {victoria.ruizserra, eduard.porta}@bsc.es

Keywords—Personalized Medicine, Cancer Disease, Immunogenomics.

I. EXTENDED ABSTRACT

One of the missions of Precision Medicine (PM) is to tackle the issue of interpreting the mass of malignant cells or tumor nature to design the most appropriate strategy to treat patients [1], [2]. To this extent, recent therapies focus on the tumor microenvironment (TME) and, more specifically, on the role of the immune cells in it [3]. Such treatments take advantage of the intrinsic properties of the immune system: malignant cells are recognized as a pathogen and in turn will be utterly destroyed [4]. Therefore, successful-to-thrive tumor means that cancer cells are evading immunological attacks among other anti-tumoral mechanisms.

The field of PM has made progress thanks to the evolution of the concept of disease. By leaving behind the application of one-layered approach studies, diseases are now understood as an interconnected and multidimensional multilayerered network of molecular mechanisms of biological processes [1], [5]. Particularly, Cancer disease was traditionally explained as the aftereffect of somatic mutations of cancer related genes [6], [7], [8]. Actually, most typical cancer-associated driver genes such TP53 or BRCA2 are usually found mutated in tumors [6], [7], [8] as they happen to be key regulators of cell death processes [9], [10]. However scientific studies solely based on isolated mutations produced on these type of genes are not able to explain nor predict the whole process and development of the disease [11].

In a further complication, every tumor is unique [12]. This is due to the convoluted relationship existing between cellular processes such as genetic and epigenetic modifications that may alter the intracellular and extracellular composition and, by extension, the TME [13]. For instance, results obtained by The Cancer Genome Atlas (TCGA) from the transcriptional analysis of 10,000 tumor samples from different subjects and 33 different types of cancer pointed out the existing variation of immune cells population across tumors (Figure 1) [14].

Against the general tendency to study the effect of cancer-associated somatic mutations on a genetic level, the present work pays attention to their impact from a structural point of view in the TME. Clearly, it is not the same to find a mutation affecting the catalytic site of a protein than in a non-functional area of the structure [15], [16]. The goal is to analyze the Biology underlying variation on the tumor immune infiltration and to what degree this can be explained by the location of missense mutations on the protein structure. More specifically,

major emphasis has been put on the mutation occurring on the interaction or interface area of these proteins.

A. Methods

We hypothesized that all protein residues involved in the same interface create a functional region within a protein. Mutations within the same functional region are more likely to have the same effect than mutations in different functional regions. To identify such interfaces, we analyzed 421582 protein coordinate files from the Protein Data Bank. We defined protein interfaces as all the residues in close proximity to either other proteins, nucleic acids or small ligands. We then generated used a linear model to analyze 1839298 missense somatic mutations from 10224 patients from The Cancer Genome Atlas:

$$
\begin{aligned}
\text{Leukocyte fraction}_i = \beta_0 \\
+ \beta_1 \text{Cancer Type}_i \\
+ \beta_2 \text{Tumor Mutation Burden}_i \\
+ \beta_3 \text{Mutation 3D Location}_i \\
+ \beta_4 \text{Biological Sex}_i \\
+ \epsilon_5
\end{aligned}
$$

with i = 1, ... , n; being n the total number of TCGA patients.

B. Results

We generated a catalogue of 145046 protein interfaces (52237 protein-protein, 5329 protein-nucleic acid and 87480 protein-ligand). Our preliminary results show that 13379 of these interfaces, when somatically mutated, correlate with changes in the quantity of the immune infiltrate in the tumor microenvironment.

For example, mutations on the interface between peptidylarginine deiminase 1 (PAD1) and Ca2+ atoms, which correspond to the catalytic site of the enzyme, were found in patients with higher levels of immune infiltrate than cancer patients with other mutations in the same gene or no mutations in the gene at all. Interestingly, PAD1 is in charge of protein citrullination since it converts arginine to citrulline, being these citrullinated proteins target of antibodies [ref]. Perhaps, in these patients, if the levels of citrulline is low due to the mutation on the catalytic site, less number of citrullinated
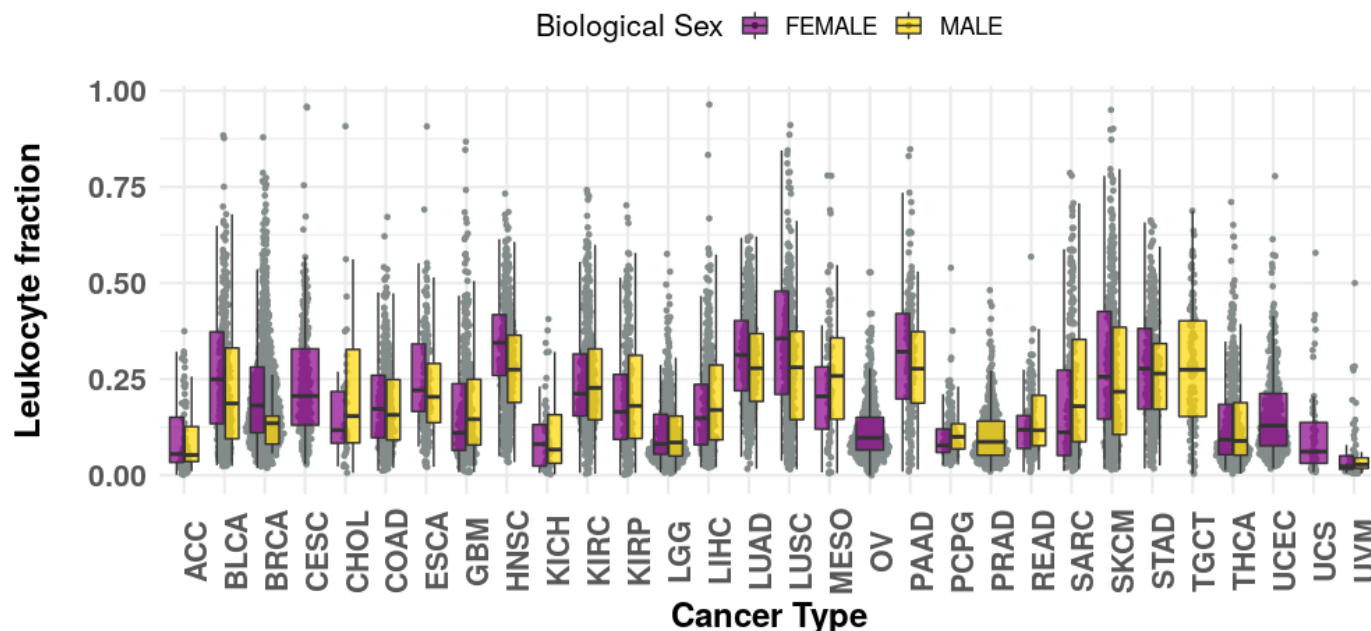
Fig. 1. Distribution of leukocyte fraction (synonym of immune infiltrate) across the different TCGA samples. Each dot corresponds to an individual tumor sample and the samples are grouped according to the Cancer type classification. Box-plots summarize each sex contributions to the leukocyte fraction distribution.

proteins will be found and lead to lower immune attack to the cancer cell.

C. Future work

Next steps will focus on the global study of all the mutations of the protein interfaces. It would be interesting to be able to discriminate which mutations have higher effects and why, in terms of tumor immune-infiltrate. Moreover, linear models will be cross-validated to rise the statistical power of the results.

REFERENCES

[1] K. Matchett, N. Lynam-Lennon, R. Watson, and J. Brown, "Advances in Precision Medicine: Tailoring Individualized Therapies," *Cancers*, vol. 9, no. 12, p. 146, oct 2017.

[2] G. Bindea, B. Mlecnik, H. K. Angell, and J. Galon, "The immune landscape of human tumors: Implications for cancer immunotherapy." *Oncoimmunology*, vol. 3, no. 1, p. e27456, jan 2014.

[3] M. Binnewies, E. W. Roberts, K. Kersten, V. Chan, D. F. Fearon, M. Merad, L. M. Coussens, D. I. Gabrilovich, S. Ostrand-Rosenberg, C. C. Hedrick, R. H. Vonderheide, M. J. Pittet, R. K. Jain, W. Zou, T. K. Howcroft, E. C. Woodhouse, R. A. Weinberg, and M. F. Krummel, "Understanding the tumor immune microenvironment (TIME) for effective therapy," *Nature Medicine*, vol. 24, no. 5, pp. 541–550, may 2018.

[4] C. L. Ventola, "Cancer Immunotherapy, Part 1: Current Strategies and Agents." *P & T : a peer-reviewed journal for formulary management*, vol. 42, no. 6, pp. 375–383, jun 2017.

[5] L. J. Frey, "Data integration strategies for predictive analytics in precision medicine," *Personalized Medicine*, vol. 15, no. 6, pp. 543–551, nov 2018.

[6] L. Luzzatto, "Somatic mutations in cancer development," *Environmental Health*, vol. 10, no. Suppl 1, p. S12, 2011.

[7] L. R. Yates and P. J. Campbell, "Evolution of the cancer genome," *Nature Reviews Genetics*, vol. 13, no. 11, pp. 795–806, nov 2012.

[8] I. R. Watson, K. Takahashi, P. A. Futreal, and L. Chin, "Emerging patterns of somatic mutations in cancer," *Nature Reviews Genetics*, vol. 14, no. 10, pp. 703–718, oct 2013.

[9] E. R. Kastenhuber and S. W. Lowe, "Putting p53 in Context." *Cell*, vol. 170, no. 6, pp. 1062–1078, sep 2017.

[10] R. Roy, J. Chun, and S. N. Powell, "BRCA1 and BRCA2: different roles in a common pathway of genome protection," *Nature Reviews Cancer*, vol. 12, no. 1, pp. 68–78, jan 2012.

[11] K. M. Coyle, J. E. Boudreau, and P. Marcato, "Genetic Mutations and Epigenetic Modifications: Driving Cancer and Informing Precision Medicine," *BioMed Research International*, vol. 2017, pp. 1–18, 2017.

[12] R. J. Hartmaier, J. Charo, D. Fabrizio, M. E. Goldberg, L. A. Albacker, W. Pao, and J. Chmielecki, "Genomic analysis of 63,220 tumors reveals insights into tumor uniqueness and targeted cancer immunotherapy strategies," *Genome Medicine*, vol. 9, no. 1, p. 16, dec 2017.

[13] F. Petitprez, C.-M. Sun, L. Lacroix, C. Sautès-Fridman, A. de Reyniès, and W. H. Fridman, "Quantitative Analyses of the Tumor Microenvironment Composition and Orientation in the Era of Precision Medicine," *Frontiers in Oncology*, vol. 8, sep 2018.

[14] V. Thorsson, D. L. Gibbs, and E. Al., "The Immune Landscape of Cancer," *Immunity*, vol. 48, no. 4, pp. 812–830.e14, apr 2018.

[15] J. Ashworth, B. Bernard, S. Reynolds, C. L. Plaisier, I. Shmulevich, and N. S. Baliga, "Structure-based predictions broadly link transcription factor mutations to gene expression changes in cancers," *Nucleic Acids Research*, vol. 42, no. 21, pp. 12 973–12 983, dec 2014.

[16] S. Khan and M. Vihinen, "Spectrum of disease-causing mutations in protein secondary structures," *BMC Structural Biology*, vol. 7, no. 1, p. 56, 2007.

Victoria Ruiz-Serra received his BSc degree in Biochemistry from University of Seville in 2014. The following 2 years, she worked at the Biochemistry and Molecular Biology department of the Faculty of Pharmacy at the University of Seville. She completed his MSc degree in Bioinformatics from Vrije University and Amsterdam University, The Netherlands in 2018. Since 2018, she is a PhD student in the Life Science group of Barcelona Supercomputing Center (BSC), Spain.

An ILP-based Real-Time Scheduler for Distributed and Heterogeneous Computing Environments

Eudald Sabaté[1], Maria A. Serrano[2], Eduardo Quiñones[3]

Barcelona Supercomputing Center (BSC), Spain

[1]eduald.sabate@bsc.es, [2]maria.serranogracia@bsc.es, [3]eduardo.quinones@bsc.es

Keywords—**real-time, schedulability analysis, distributed computing environment, ILP**

ABSTRACT

The digitalization process is making cities to rapidly increase the amount of data to be processed upon which data analytics can extract valuable knowledge. However, this phenomenon is facing many important challenges. On one side, the advent of connected and autonomous vehicles challenges data analytics methods due to the need of accomplishing real-time requirements. On the other side, the dispersion nature of data sources makes current big data analytics methods, commonly designed to execute in centralized and computationally intensive (cloud-based) environments, not suitable for smart cities. The use of distributed computing environments composed of advanced parallel embedded processor architectures at the edge, e.g., NVIDIA Jetson, Kalray MPPA, can help alleviating the pressure on centralized cloud-based solutions, while providing the real-time guarantees needed to implement advanced mobility functionalities on cars and cities.

To do so, this work presents a novel scheduler (based on ILP formulation) to optimally distribute the computation across the compute continuum composed of multiple edge devices, while providing real-time guarantees. Our scheduler, implemented in the COMPSs distributed programming model developed at BSC, statically assigns tasks to those edge devices so that the overall response time of the workflow is minimized. It takes into account an execution time upper bound of the computation and communication existing in the workflow.

A. Introduction

Critical real-time systems are commonly modelled as a set of concurrent and periodic real-time tasks that implement the system functionalities [1]. Each real-time task is characterized with a *period,* a *deadline* and *direct acyclic graph (DAG)* composed of a set of nodes, implementing sub-functionalities, and a set of edges connecting the nodes, representing the precedence constraints among them. In this context, real-time scheduling techniques are key to: (1) efficiently execute the real-time tasks into computing resources, and (2) guarantee that the deadlines of the real-time tasks are met by means of schedulability analysis [2]. Interestingly, workflows in COMPSs [3], a task-based programming framework for distributed computing environments, are represented as a DAG, in which nodes correspond to COMPSs tasks and edges to the data dependencies existing among them. Unfortunately, current schedulability analysis are only applicable to systems executed on the same computing node and therefore, cannot be applied to COMPSs.

This work tackles the problem of efficiently distributing a COMPSs workflow across the compute continuum, while guaranteeing its timing constraints. To do so, we develop a schedulability analysis based on *Integer Lineal Programming* (ILP) formulation with a twofold objective: (1) to minimize the execution time of the COMPSs workflow by statically assigning COMPSs tasks to the computing resources that form the compute continuum; and (2) to provide an upper bound response time of the overall workflow by considering the timing characterization of each COMPSs task and the data transfer among them. This static allocation is then implemented by a dedicated COMPSs scheduler.

B. COMPSs

COMPSs offers a task-based and portable programming framework that facilitates the distribution and parallelization of sequential source code (written in Java, C/C++ or Python) in a distributed and heterogeneous computing environment, such as those existing in smart cities. In COMPSs, the programmer is responsible of identifying *COMPSs tasks* and the data dependencies existing among them, by annotating the sequential source code. The run-time scheduler is then in charge of distributing the COMPSs tasks among the available computing resources across the compute continuum, from on edge to cloud.

The COMPSs runtime incorporates several task schedulers implementing different allocation policies: (1) FIFO selects the first ready task and the first available computing resource; (2) FIFO + Data locality selects the first ready task and the computing resource that better exploits the data locality; (3) LIFO selects the last ready task and the first available computing resource; and (4) FIFO + Load Balancing selects the first ready task and the computing resource that better balances the overall load of the system. These strategies are intended to exploit the performance in HPC environments, but do not guarantee the real-time constraints of the system.

We tackle the problem of efficiently distributing a COMPSs workflow and guaranteeing the timing constraints, by developing a dedicated COMPSs scheduling strategy based on a static allocation of COMPSs tasks to computing resources derived with an ILP formulation.

C. ILP Scheduling Strategy

With the objective of incorporating real-time requirements while efficiently distributing the workflow, we propose a new scheduling strategy that allocates COMPSs tasks to computing resources, based on a predefined static allocation given by an ILP formulation. ILP is a well-known technique already used in real-time systems to compute the minimum makespan of real-time tasks represented as a DAG [4][5].

Concretely, our ILP scheduling strategy relies upon a system model that considers the following information: (1) a representation of the COMPSs workflow by means of a DAG, that includes the execution time upper bound of each COMPSs task and the characterization of the size of the data dependencies between tasks; and (2) a description of the compute continuum, including the computing resources and the communication network. Next, we briefly describe the system model considered.

C.1. Compute continuum model

The compute continuum model represents the available (and heterogeneous) computing resources interconnected by different network links. It is represented as a directed graph (digraph) where each node represents a computing resource characterized

by its type and computation capabilities, (e.g., GPU, CPU, Instruction Set Architecture (ISA) supported), and each edge represents the communication link between nodes, characterized by the transport bandwidth. Figure 1(a) shows an example of a compute continuum model representing an edge-cloud system, consisting of four computing resources: a computing resource available on a car, two computing nodes located within the street and a data center representing the city cloud services
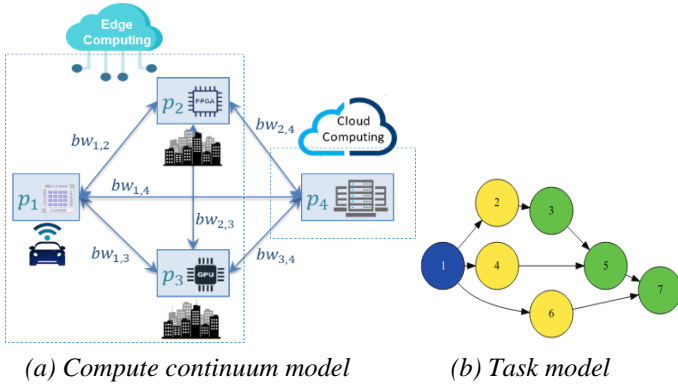

(a) Compute continuum model   *(b) Task model*

Figure 1. System model.

C.2. Task Model

A COMPSs workflow is represented by a DAG, where each node represents a COMPSs task, and each edge represents a data dependency existing between COMPSs tasks. Edges are characterized by the total size of the data involved in the dependency. If two dependent tasks execute in different computing resources, then this data must be transferred through the computer network. Edges also represent the execution order constraints, i.e., if there exists an edge between two nodes, then predecessor node must complete before successor node can begin its execution. Since each task may have different implementations, for each of the potential computing resources in which it can execute, each node is characterized by set of values corresponding to an upper bound of the execution time of the task in a given computing resource. Figure 1(b) shows an example of a task workflow, represented as a DAG, composed of seven nodes and eight edges.

D. Evaluation

This section provides a preliminary evaluation of the ILP scheduling strategy proposed.

D.1. Experimental Setup

We consider a Cholesky factorization [6] application (implemented in Python) and parallelized with COMPSs. The Cholesky factorization is commonly used for efficient linear equation solvers, Monte Carlo simulations, or to accelerate Kalman filters implemented, for example, in vehicle navigation systems to detect objects positions and compute trajectories. The COMPSs workflow, that processes a matrix of 2048×2048, generates a DAG composed of 30 COMPSs tasks. The execution time upper bound of each COMPSs tasks have been computing by selecting the maximum observed execution time of 50 executions.

Moreover, we consider a compute continuum model with two computing resources: (1) a four-core Intel(R) i7-7600U processor @ 2.80GHz and (2) a four core ARMv8 Processor rev 3 included in a NVIDIA Jetson TX2. The communication link between them is an IEEE 802.11g at 54 Mbps of bandwidth.

D.2. Results

Table 1 compares the execution time (in seconds) of the Cholesky factorization executed on our compute continuum

model, considering our ILP-based scheduling strategy and the four different baseline COMPSs schedulers, i.e., FIFO, FIFO + Data locality, LIFO and FIFO + Load Balancing strategies.

Table 1. Execution time (in sec) of the Cholesky factorization.

ILP-based Scheduler	COMPSs baseline scheduler			
	FIFO	FIFO + Data Locality	LIFO	FIFO + Load Balancing
8.16	16.36	12.53	14.68	11.92

As observed, and for this particular example, our ILP scheduler strategy clearly outperforms the COMPSs baseline schedulers, reducing the execution of time of the Cholesky factorization by two when comparing to the COMPSs FIFO schedulers presented in Section B.

E. Conclusion and Future Enhancement

This work presents a static allocation strategy based on an ILP-formulation that minimizes the execution time of a COMPSs workflow, while providing an upper bound of its response time. The next steps are to evaluate our scheduling strategy with a more complex (and realistic) computing continuum model of a smart city, and considering a workflow that implements an advanced mobility functionality with data analytics methods.

F. Acknowledgements

References

[1] S. Baruah, M. Bertogna, and G. Buttazzo. *Multiprocessor Scheduling for Real-Time Systems.* Springer, 2015.

[2] G. C. Buttazzo. *Hard real-time computing systems: predictable scheduling algorithms and applications*, volume 24. Springer Science & Business Media, 2011.

[3] F. Lordan, E. Tejedor, J. Ejarque, R. Rafanell, J. Álvarez, F. Marozzo, D. Lezzi, R. Sirvent, D. Talia, and R. M. Badia. *ServiceSs: an interoperable programming framework for the Cloud*, Journal of Grid Computing, March 2014.

[4] M. A. Serrano, A. Melani, M. Bertogna, E. Quiñones. *Response-time analysis of DAG tasks under fixed priority scheduling with limited preemptions.* In DATE, 2016.

[5] A. Melani, M. A. Serrano, M. Bertogna, I. Cerutti, E. Quiñones, G. Buttazzo. *A static scheduling approach to enable safety-critical OpenMP applications.* In ASP-DAC 2017.

[6] N. Baščelija. *Sequential and parallel algorithms for cholesky factorization of sparse matrices.* Mathematical Applications in Science and Mechanics, 2013.

Author biography

Eudald Sabaté was born in Barcelona, Spain, in 1995. He received the Bachelor degree in computer science from the Polytechnic University of Catalonia (UPC) in 2018, and is currently enrolled in the MSc in Innovation and Research in Informatics at UPC. Since April 2018, he has been working as a research student at BSC in the CAOS (Computer Architecture and Operating Systems) group.