Accelerating Atmospheric Models using GPUs

Christian Guzman-Ruiz, Mario Acosta, Oriol Jorba
Barcelona Supercomputing Center, Barcelona, Spain
E-mail: {christian.guzman, mario.acosta, oriol.jorba}@bsc.es

Keywords—Modeling and prediction, Chemistry, Parallelism and concurrency, Performance

I. Extended Abstract

Environmental models are simplified representations of an object or a process [1]. These models provide valuable information on the nature of real-world phenomena and systems [2], with many applications in science and engineering [3]. For example, environmental models play an increasingly important role in understanding the potential implications of climate change [4].

There are many types of models in the environmental sciences [5]. These models are often associated with large computational costs because of their complexity [6]. The model studied in this work, the *Multiscale Online Nonhydrostatic AtmospheRe CHemistry* model (MONARCH), is an atmospheric model that currently runs in the MareNostrum supercomputer of the Barcelona Supercomputing Center (BSC), one of the Top-500 supercomputers in the world [7] [8]. MONARCH provides regional mineral dust forecasts to the World Meteorological Organization's (WMO) Barcelona Dust Forecast Center (BDFC) and the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS). MONARCH also provides global aerosol forecasts to the International Cooperative for Aerosol Prediction (ICAP) initiative.

However, MONARCH does not currently make use of all the supercomputing resources available in the Marenostrum cluster. The MONARCH source code was originally prepared for clusters exclusively using CPUs. However, Marenostrum comprises two supercomputers with different system architectures: Marenostrum IV (MN4), a CPU-only cluster, and CTE-POWER, a GPU-accelerated cluster [9]. As MONARCH lacks GPU-specific code, as is often the case for environmental system models, it is currently only run on MN4.

For many years, the CPU-centric approach was accepted by the scientific community as individual GPUs have low computational abilities compared with CPUs, and the effective use of GPUs presents challenges to scientific software developers. However, with the advent of new, more powerful GPUs and the growth of Big Data [10], GPUs have begun assuming some of the computational responsibilities of CPUs on supercomputers for specific applications. Many studies have found improved efficiency applying GPU technology to traditional CPU-only software in supercomputers [11], and more-powerful GPU-accelerated supercomputers are appearing with increasing frequency [12].

In response to these recent developments, new approaches to GPU-based modeling have been developed. The molecular modeling community was among the earliest adopters of GPU-based computing, providing in some cases a performance of two-orders-of-magnitude faster than CPU-only implementations [13]. Atmospheric models have primarily focused on the parallelization of dynamical process calculations [14], as does the Nonhydrostatic Icosahedral Atmospheric Model (NICAM), which has been parallelized for GPUs and reports a 7–8$\times$ performance speedup compared with the CPU version [15]. Chemical solvers, another important component of atmospheric models, have also been ported to GPUs, by either fully integrating GPU usage in the solver with a reported 20$\times$ speedup [16], or porting components of the solver to GPUs to acheive a 3–4$\times$ speedup [17]. Our work follows the latter approach, as we port only part of the chemical solving to GPUs. However, previous work [17] did not take into account the number of solver iterations in the performance assessment,and the grid-cell iterations are computed sequentially for both CPU and GPU versions. Here, we are able to reduce the number of grid-cell iterations for the GPU version to one for an idealized case. Moreover, our implementation is based on the innovative *Chemistry Across Multiple Phases* (CAMP) framework, which allows multiple chemical processes to be solved simultaneously as a single system [18].

In this work, we investigate the potential of a GPU-based chemical treatment by selectively adapting computationally costly portions of the MONARCH code to GPUs, to reduce the overall execution time. We achieve this objective by performing the following set of tasks:

1) Select the most time-consuming MONARCH component by analyzing the model from a computational point of view.
2) Study the efficiency of porting the selected component to GPUs, to evaluate the potential for improvement in the execution time using a GPU implementation.
3) Propose possible GPU implementations and general optimizations for both GPU and CPU implementations.
4) Apply these proposals and study their effect on execution time, in terms of the expected and actual improvement.
5) Evaluate the heterogeneous implementation, including transfers from CPU to GPU and GPU to CPU.

A computational profiling analysis showed that chemistry-related components of MONARCH are computationally expensive. His future chemistry solver (CAMP) will take more than 48.6% of the execution time, making it an interesting objective to optimize and our main focus. To optimize this module, this work focused around a GPU implementation in

order to exploit the available parallelization of the chemical reactions and the big amount of GPU resources available at Marenostrum at the moment.

As our main optimization, we adapted CAMP to allow us to solve chemical processes simultaneously for multiple grid cells in a given spatial region as a single system, as opposed to the common practice of looping through individual grid cells. This solution avoids repeating initialization steps, memory misses, solver iterations and allows further vectorization during solving. Our adaptation only requires a single solver call per chemistry time-step, allowing the solver to compute internally the chemistry for all grid cells and achieves a factor of 12 time-speedup for the configuration tested. Furthermore, we have started to port some of the most time-consuming functions to GPUs, splitting chemical reaction computation amongst individual threads. The selected function represents 30% of the chemistry computation time. Preliminary GPU results of such functions show a speedup of x12, ending in a total module speedup of 18x from the base version. With more possible optimizations pending, our work denotes the GPU potential and the importance of developing from a concurrent point of view.

II. Acknowledgment

Special thanks to Matthew Dawson, Kim Serradell, and all the people involved from BSC.

References

[1] S. Hartmann and R. Frigg, "Scientific models," Jan. 2005.

[2] J. D. Sterman, "All models are wrong: reflections on becoming a systems scientist," *System Dynamics Review*, vol. 18, no. 4, pp. 501–531, 2002. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/sdr.261

[3] M. Poznic, "Models in Science and Engineering," Tech. Rep., Apr. 2017.

[4] M. Carmen Lemos *et al.*, "Narrowing the Climate Information Usability Gap," *Nature Climate Change*, vol. 2, Oct. 2012.

[5] E. Holzbecher, *Environmental Modeling*, May 2012.

[6] N. D Bennett *et al.*, "Performance evaluation of environmental models," Jul. 2010.

[7] "top500 Marenostrum." [Online]. Available: https://www.top500.org/system/179067

[8] O. Jorba *et al.*, "Potential significance of photoexcited NO2 on global air quality with the NMMB/BSC chemical transport model," *Journal of Geophysical Research D: Atmospheres*, vol. 117, p. D13301, Jul. 2012.

[9] "Technical Information." [Online]. Available: https://www.bsc.es/marenostrum/marenostrum/technical-information

[10] E. Mizell and R. Biery, "Introduction to GPUs for Data Analytics," p. 39.

[11] Y. Jararweh *et al.*, "Gpu-based personal supercomputing," *2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*. [Online]. Available: https://www.academia.edu/10680401/Gpu-based_personal_supercomputing

[12] "New GPU-Accelerated Supercomputers Change the Balance of Power on the TOP500 | TOP500 Supercomputer Sites." [Online]. Available: https://www.top500.org/news/new-gpu-accelerated-supercomputers-change-the-balance-of-power-on-the-top500/

[13] "GPU-accelerated molecular modeling coming of age. - PubMed - NCBI." [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20675161

[14] M. Govett *et al.*, "Parallelization and Performance of the NIM Weather Model on CPU, GPU, and MIC Processors," *Bull. Amer. Meteor. Soc.*, vol. 98, no. 10, pp. 2201–2213, Oct. 2017. [Online]. Available: http://journals.ametsoc.org/doi/10.1175/BAMS-D-15-00278.1

[15] H. Sato *et al.*, "Extreme Big Data with TSUBAME2 and Beyond," p. 20.

[16] M. Alvanos and T. Christoudias, "GPU-accelerated atmospheric chemical kinetics in the ECHAM/MESSy (EMAC) Earth system model (version 2.52)," *Geoscientific Model Development*, vol. 10, no. 10, pp. 3679–3693, Oct. 2017. [Online]. Available: https://www.geosci-model-dev.net/10/3679/2017/

[17] J. Sun *et al.*, "Computational Benefit of GPU Optimization for the Atmospheric Chemistry Modeling," *J. Adv. Model. Earth Syst.*, vol. 10, no. 8, pp. 1952–1969, Aug. 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001276

[18] M. D. Dawson, Matthew, "CAMP: A Scalable, Portable, Gas–Aerosol Chemistry Treatment for Atmospheric Models," AAAR, Portland, OR, USA, Oct. 2019.

Christian Guzman received his Bachelor's degree in Computer Engineering plus a bachelor's degree in Telecommunication Electronic Engineering, and a Master in Modelling for Science and Engineering by the Autonomous university of Barcelona (UAB). He specialized in techniques for High-Performance Computing (HPC) and is actually developing his pre-doctoral studies on the Barcelona Supercomputing Center (BSC), working on the development of Chemistry Across Multiple Phases (CAMP) module alongside the Multiscale Online Nonhydrostatic Atmosphere Chemistry Model (MONARCH), contributing to the most-computational and logic part from a performance point of view and integrating multitudinal ways of GPU computation in search of speeding-up the system.