A high-performance model for shallow-water simulations in distributed and heterogeneous architectures

Daniel Conde, Ricardo B. Canelas, and Rui M. L. Ferreira
Instituto Superior Técnico, CERIS, Lisbon, Portugal (daniel.conde@tecnico.ulisboa.pt)

One of the most common challenges in hydrodynamic modelling is the trade off one must make between highly resolved simulations and the time required for their computation. In the particular case of urban floods, modelers are often forced to simplify the complex geometries of the problem, or to implicitly include some of its hydrodynamic effects, due to the typically very large spatial scales involved and limited computational resources.

At CERIS – Instituto Superior Técnico, Universidade de Lisboa – the STA V-2D shallow-water model, particularly suited for strong transient flows in complex and dynamic geometries, has been under development for the past recent years (Canelas et al., 2013 & Conde et al., 2013). The model is based on an explicit, first-order 2DH finite-volume discretization scheme for unstructured triangular meshes, in which a flux-splitting technique is paired with a reviewed Roe-Riemann solver, yielding a model applicable to discontinuous flows over time-evolving geometries. STA V-2D features solid transport in both Euleran and Lagrangian forms, with the first aiming at describing the transport of fine natural sediments and the latter aimed at large individual debris. The model has been validated with theoretical solutions and laboratory experiments (Canelas et al., 2013 & Conde et al., 2015).

This work presents our most recent effort in STA V-2D: the re-design of the code in a modern Object-Oriented parallel framework for heterogeneous computations in CPUs and GPUs. The programming language of choice for this re-design was C++, due to its wide support of established and emerging parallel programming interfaces. The current implementation of STA V-2D provides two different levels of parallel granularity: inter-node and intra-node. Inter-node parallelism is achieved by distributing a simulation across a set of worker nodes, with communication between nodes being explicitly managed through MPI. At this level, the main difficulty is associated with the unstructured nature of the mesh topology with the corresponding employed solution, based on space-filling curves, being analyzed and discussed. Intra-node parallelism is achieved through OpenMP for CPUs and CUDA for GPUs, depending on which kind of device the process is running. Here the main difficulty is associated with the Object-Oriented approach, where the presence of complex data structures can degrade model performance considerably.

STA V-2D now supports fully distributed and heterogeneous simulations where multiple different devices can be used to accelerate computation time. The advantages, short-comings and specific solutions for the employed unified Object-Oriented approach, where the source code for CPU and GPU has the same compilation units (no device specific branches like seen in available models), are discussed and quantified with a thorough scalability and performance analysis.

The assembled parallel model is expected to achieve faster than real-time simulations for high resolutions (from meters to sub-meter) in large scaled problems (from cities to watersheds), effectively bridging the gap between detailed and timely simulation results.

Acknowledgements

This research as partially supported by Portuguese and European funds, within programs COMPETE2020 and PORL-FEDER, through project PTDC/ECM-HID/6387/2014 and Doctoral Grant SFRH/BD/97933/2013 granted by the National Foundation for Science and Technology (FCT).
References