Multi-CPU and multi-GPU hybrid computations of multi-scale scalar transport

Jean-Matthieu Etancelin, Georges-Henri Cottet, Franck Pérignon, Christophe Picard

To cite this version:

HAL Id: hal-00982986
https://hal.archives-ouvertes.fr/hal-00982986
Submitted on 24 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
MULTI-CPU AND MULTI-GPU HYBRID COMPUTATIONS OF MULTI-SCALE SCALAR TRANSPORT

JEAN-MATTHIEU ETANCELIN*, GEORGES-HENRI COTTET*, FRANCK PÉRIGNON* AND CHRISTOPHE PICARD*

* Univ. Grenoble Alpes and CNRS
LJK, F-38000 Grenoble, France
e-mail: Jean-Matthieu.Etancelin@imag.fr

Key words: Multi-scale method, Hybrid computing, Remeshed particle method.

Abstract.

The transport of a scalar in a turbulent flow is a central issue in many fields such as combustion, environmental flows or multiphase flows. The problem to solve consists in the coupling of the incompressible Navier-Stokes equations with an advection-diffusion equation. The ratio of the flow viscosity to the scalar diffusivity, called the Schmidt number, characterizes the dynamics of this model. For Schmidt numbers higher than one, the smallest turbulence length scale of the flow motion is larger than the one of the scalar fluctuations [1]. In these situations, the use of a multi-scale hybrid method is mandatory to avoid solving the flow at the smallest scale.

In our calculations, both the flow and scalar solvers rely on remeshed particle methods. These semi-Lagrangian methods were originally designed for advection dominated problems. Particle remeshing enables to preserve a regular particle distribution and to control the accuracy of the method [2]. High order remeshing kernels combined with dimensional splitting allows to derive methods which combine stability, even for large time-steps, high order of accuracy and affordable cost [3].

In practice, the two parts of the problem (flow and scalar solvers) are computed simultaneously on different processes. The hybrid computing strategy consists in defining a subset of the processes that only solves the flow using the parallel CPU advection solver from [1] in addition to finite differences and FFT-based Poisson solver. Meanwhile, the other processes compute the scalar transport on several GPUs by extending the library introduced in [3]. The proposed implementation involves three levels of parallelism: tasks split among all processes, multi-CPU solvers within the tasks, and multi-core computations using GPUs. These levels are implemented using the MPI programming model and the OpenCL framework on the GPU devices.

The benchmark shown on Figure 1 concerns the transport of a passive scalar in a turbulent plane jet in a periodic unit box. Velocity and scalar fields are initialized with hyperbolic tangent profiles complemented by a small random perturbations (see [4] for details). The Schmidt number is equal to 10 and the Reynolds number, based on the jet width, is equal to to $10^3$. The flow and scalar meshes use $128^3$ and $512^3$ grid points, respectively.
Figure 1: Side and top views of the contours of a passive scalar in a turbulent plane jet in a hybrid CPU/GPU computations with $128^3$ and $512^3$ grid points for the velocity and the scalar, respectively.

The computational times per iteration are given on Table 1 for the above simulation. The overall time to solution is driven by the flow solver when using a single GPU. In the multi-GPU case, we observe on preliminary results a large difference between the total time and the computational parts. This is explained by the amount of communications, from and to GPU devices and between hosts, which are not yet fully optimized. However, in each part, we obtain a speedup close to two when doubling the amount of resources.

Table 1: Computational times (in seconds per iteration) for a plane jet simulation on a system with two hexacore Xeon E5-2640 and two Tesla K20m.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Total time</th>
<th>Flow</th>
<th>Passive Scalar</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 CPU + 1 GPU</td>
<td>2.56</td>
<td>2.49</td>
<td>0.87</td>
</tr>
<tr>
<td>9 CPU + 1 GPU</td>
<td>1.43</td>
<td>1.36</td>
<td>0.86</td>
</tr>
</tbody>
</table>

The optimization of the necessary communications between the GPUs for the scalar transport is the subject of ongoing work.

REFERENCES


