Offloading computational tasks of hybrid MPI + OpenMP/OmpSs-2 applications to GPUs

Table of contents:

Quick Overview

NBody Benchmark

Users can clone or download this examples from the ?https://pm.bsc.es/gitlab/DEEP-EST/apps/NBody repository and transfer it to a DEEP working directory.

Description

An NBody simulation numerically approximates the evolution of a system of bodies in which each body continuously interacts with every other body. A familiar example is an astrophysical simulation in which each body represents a galaxy or an individual star, and the bodies attract each other through the gravitational force.

N-body simulation arises in many other computational science problems as well. For example, protein folding is studied using N-body simulation to calculate electrostatic and *Van der Waals* forces. Turbulent fluid flow simulation and global illumination computation in computer graphics are other examples of problems that use NBody simulation.

Requirements

The requirements of this application are shown in the following lists. The main requirements are:

- **GNU Compiler Collection**.
- **OmpSs-2**: OmpSs-2 is the second generation of the **OmpSs** programming model. It is a task-based programming model originated from the ideas of the OpenMP and StarSs programming models. The specification and user-guide are available at ?https://pm.bsc.es/ompss-2-docs/spec/ and ?https://pm.bsc.es/ompss-2-docs/user-guide/, respectively. OmpSs-2 requires both **Mercurium** and **Nanos6** tools. Mercurium is a source-to-source compiler which provides the necessary support for transforming the high-level directives into a parallelized version of the application. The Nanos6 runtime system library provides the services to manage all the parallelism in the application (e.g., task creation, synchronization, scheduling, etc). Downloads at ?https://github.com/bsc-pm.
- **Clang + LLVM OpenMP** (derived):
- **MPI**: This application requires an MPI library supporting the multi-threading level of thread support.

In addition, there are some optional tools which enable the building of other application versions:

- **CUDA** and NVIDIA **Unified Memory** devices: This application has CUDA variants in which some of the N-body kernels are executed on the available GPU devices.
- **Task-Aware MPI (TAMPI)**: The Task-Aware MPI library provides the interoperability mechanism for MPI and OpenMP/OmpSs-2. Downloads and more information at ?https://github.com/bsc-pm/tampi.

Versions

The NBody application has several versions which are compiled in different binaries, by executing the `make` command. All of them divide the particle space into smaller blocks. MPI processes are divided into two groups: GPU processes and CPU processes. GPU processes are responsible for computing the forces between each pair of particles blocks, and then, these forces are sent to the CPU processes, where each process updates its particles blocks using the received forces. The particles and forces blocks are equally distributed amongst each MPI process in each group. Thus, each MPI process is in charge of computing the forces or updating the particles of a consecutive chunk of blocks.

The available versions are:

- **nbody.mpi.${BS}bs.bin**: Parallel version using MPI.
- **nbody.mpi.ompss2.${BS}bs.bin**: Parallel version using MPI + OmpSs-2 tasks. Both computation and communication phases are taskified, however, communication tasks are serialized by declaring an artificial dependency on a sentinel variable. This is to prevent deadlocks between processes, since communication tasks perform blocking MPI calls.

- **nbody.mpi.ompss2.cuda.${BS}bs.bin**: The same as the previous version but using CUDA tasks to execute the most compute-instensive parts of the application at the available GPUs.
- **nbody.tampi.ompss2.${BS}bs.bin**: Parallel version using MPI + OmpSs-2 tasks + TAMPI library. This version disables the artificial dependencies on the sentinel variable, so communication tasks can run in parallel. The TAMPI library is in charge of managing the blocking MPI calls to avoid the blocking of the underlying execution resources.
- **nbody.tampi.ompss2.cuda.${BS}bs.bin**: The same as the previous version but using CUDA tasks to execute the most compute-instensive parts of the application at the available GPUs.
- **nbody.mpi.omp.${BS}bs.bin**:
- **nbody.mpi.omptarget.${BS}bs.bin**:
- **nbody.tampi.omp.${BS}bs.bin**:
- **nbody.tampi.omptarget.${BS}bs.bin**: