Energetically Efficient Acceleration EEA-Aware
For Scientific Applications of Large-Scale On Heterogeneous Architectures

John A. García. H.1,2,4, Esteban Hernandez B.1,3, Philippe O. A. Navaux2, Carlos J. Barrios H.1

1,2,4 henao@i3s.unice.fr, 1,3 ejhernandezb@udistrital.edu.co, 2,4 navaux@inf.ufrgs.br, 1cbarrios@uis.edu.co

1 High Performance and Scientific Computing Center, SC3
Universidad Industrial de Santander, UIS - Bucaramanga, Colombia
2 Parallel and Distributed Processing Group, GPPD
Universidade Federal do Rio Grande do Sul, UFRGS - Porto Alegre, Brazil
3 Interoperability of Academic Networks Group - GIIRA,
Universidad Distrital Francisco Jose de Caldas, UDI - Bogota, Colombia
4 Laboratory of Informatics, Signaux et Systèmes de Sigüey Antipolo – ISS, Université Côte d’Azur, UMA - CNRS - Nice, France

Challenge

Heterogeneous parallel programming has two main problems on large computation systems:

- The increase of power consumption on supercomputers in proportion to the amount of computational resources used to obtain high performance.
- The underuse these resources by scientific applications with improper distribution of tasks.

Analysis of Energy Efficiency and Portability of large Scale Scientific Application:
Case Study: Ondes3D®

Minimization of Seismic Wave Propagation is a critical tool for analyzing geophysical seismic movements.

Planning workload experimental on Ondes3D®

StarPU benchmarks library for hybrid architectures.

EEA-Aware Structure

This research proposes an integrated energy-aware scheme called Energetically Efficient Acceleration (EEA) for scientific applications of large-scale on heterogeneous architectures. The EEA energy-aware scheme has a workflow of three steps as shown in EEA-Aware Scheme. In the first step, the data is captured in runtime and executed the energyPU or energyPhi monitor tool for analyzing the key factors by results of each experiment in terms of energy efficiency and estimated the power levels. In the second step, the data visualization and statistics characterization used a separate level of energyPU and energyPhi monitor tool to analyze the key factors of results of each experiment in terms of energy efficiency and estimated the power levels. In the third step, the data visualization and statistics characterization used a separate level of energyPU and energyPhi monitor tool to analyze the key factors of results of each experiment in terms of energy efficiency and estimated the power levels. Finally, the results collected by monitoring cost functions are built in the model prediction system running for obtaining the optimal computational resources in a static time to mapping parallel task granularity of scientific applications on heterogeneous architectures.

Step 1: Experimental Procedures and Results

At the first level the global parameters have chosen the workload and computational architecture according to the combination of resources that will be used. The experimental procedures were executed with a set of tests of 81 codes variants using different workload and architectural parameters on Cluster nodes GUANE. The experimental procedures were chosen following the fractional factorial design principle proposed by Raj Jain [4].

The computational resources used, 1 GUANE Node on 'A' Settings.

Step 2: Architectural Characterization

The second level uses the energyPU monitor in the post-processing for data visualization and statistical characterization of each architecture. In which the figure below shows that the accuracy using the key factors is much more accurate than just using the number of GPUs. In addition, this monitor displays information via sequence data, statistics, and tables showing results in terms of energy efficiency, for each experiment.

Minimization model with Task and Architectural Parameters

Accuray of

Step 3: EEA Prediction System

The third level uses the projected multiple regression model results to see metrics such as time, performance, power consumption, power consumption and performance per watt, which are used to execute the HPL for each combination of computational resources and calculate the best combination of computational resources, as shown in Figure.

EEA-Aware Scheme for Scalability and Portability of Large Scale Applications on Heterogeneous Architectures:

Heterogeneous Architectures

Results of running Ondes3D® with StarPU on different machines.