INF/UFRGS – Prof. Lucas M. Schnorr – http://www.inf.ufrgs.br/~schnorr/
LIG/UJF – Prof. Jean-Marc Vincent – http://mescal.imag.fr/membres/jean-marc.vincent/

Scientific Master Proposal

Energy and Performance Visualization for the Analysis of Parallel Programs

Context

The performance analysis of parallel programs consists in the observation of the runtime behavior of the application, searching for performance problems and bottlenecks. This analysis is essential to correctly deploy an application in the platform, maximizing efficiently the utilization of available resources. The network topology and the processing power give a series of constraints for the application execution that must be taken into account during the performance analysis of the application. Nowadays, the performance analysis faces several problems such as the performance scalability, when applications grow larger, or the comparison of the behavior observed in several executions of the same application with different configurations.

Energy concerns appear as parallel platforms grow larger, because the amount of energy consumed by existing infrastructures could be used to power whole cities. The Green500 list ranks the top 500 parallel platforms by their energy efficiency. Even if energy is an already established objective among parallel platforms designers, few works [2, 1] tries to optimize the global energy consumption of parallel and distributed applications. Most of the existing efforts focus on micro optimizations in the sequential code of parallel applications.

Objective

The objective of this proposal is to investigate how to combine a traditional performance analysis for parallel applications that consider also energy as an important metric. The resulting methods and techniques should consider performance metrics, such as speed and resource efficiency, and also energy consumption caused by application code. We also expect that a visualization and interactive prototype should be developed in order to help analysts to identify energy bottlenecks related with the performance. The PajeNG [4, 3] free-software ecosystem[1] should be used for the development of this proposal.

Methodology

The methodology is strongly experimental, with two phases:

#1: Energy and performance data collection / State of the art. The first phase consists in the investigation of existing energy and performance data collection mechanisms, especially those where such metrics are collected along time. An exhaustive investigation of the state of the art should be done to correctly position this proposal. Novel tracing libraries should be created in the case where energy metrics are missing from traditional application tracing frameworks. Information of energy consumption from CPU and accelerators (GPU, XeonPhi, etc) have to be considered, and also the network interconnection among resources if feasible. The expected result for this phase is a trace consisting of application behavior and energy consumption metrics from the different resources used by the application.

#2 (Brazil): Development of visualization prototype. The second phase consists in the design and development of novel visualization techniques that take into account the traces collected in the first phase. Energy measurements should be represented together with the application behavior, trying to correlated energy with traditional performance analysis metrics. A discussion on how to efficiently analyze this kind of data should also be addressed, according to the state of the art in the performance analysis area.

#2 (France): Development of visualization prototype. The second phase consists in the choice of a visualization tool and a proposition of evolution of this tool to take into account energy parameters for the analysis.

[1] Available at `http://github.com/schnorr/pajeng`

Work environment

Programming:	C, C++, bash, Python, R	Report and slides:	latex
Auxiliary tools:	gdb, valgrind	Project management:	git

Required skills

In addition to the skills that can reasonably be expected from master-level students, the applicant should have a strong knowledge of system programming in C, and of modern Unix-based operating systems such as Linux.

Time schedule

Table 1 presents the expected activities for this proposal.

Table 1: Expected time schedule for the expected ativities of the proposal

Activity	**Period (units of time)**											
	1	2	3	4	5	6	7	8	9	10	11	12
1 – State of the art	•	•	•	•					•	•		
2 – Implementation		•	•	•	•					•	•	
3 – Experiments			•	•	•	•	•	•	•	•	•	
4 – Validation				•	•	•	•	•	•	•	•	•
5 – Writing the report					•				•	•	•	•
6 – Defense												•

References

[1] George Bosilca, Hatem Ltaief, and Jack Dongarra. Power profiling of cholesky and qr factorizations on distributed memory systems. *Computer Science - Research and Development*, 29(2):139–147, 2014.

[2] Rong Ge, Xizhou Feng, Shuaiwen Song, Hung-Ching Chang, Dong Li, and K.W. Cameron. Powerpack: Energy profiling and analysis of high-performance systems and applications. *Parallel and Distributed Systems, IEEE Transactions on*, 21(5):658–671, May 2010.

[3] Lucas Mello Schnorr and Arnaud Legrand. Visualizing more performance data than what would fit on your screen. In *Tools for High Performance Computing 2012*. Springer Berlin Heidelberg, 2012.

[4] Lucas Mello Schnorr, Arnaud Legrand, and Jean-Marc Vincent. Detection and analysis of resource usage anomalies in large distributed systems through multi-scale visualization. *Concurrency and Computation: Practice and Experience*, 24(15):1792–1816, 2012.