Implementation and evaluation of Raspberry Pi clusters

Streignard Rémi 2021-2022

Cluster computing dates back to the 1970s, with the advent of time-sharing systems, designed to allow multiple users to share computing resources simultaneously.

Previously, there was a single computer that managed all tasks and processes. This single computer was called a "central processing unit" or CPU. However, with this design, if one task took too long to process, it meant waiting for other tasks to be processed.

Thus, cluster computing has changed the game in the world of computing, offering the ability to create distributed systems in which the workload is distributed among several computers connected to each other.

The objective of this study, inspired by previous research works dealing with cluster computing and benchmarking, is: on the one hand, to define an easy implementation procedure of a cluster ; on the other hand, to evaluate the performance of such a system.

The cluster implemented in this study consists of 9 RPIs that are built around a LAN with a shared storage space. Several tools are used to interconnect and automate the configurations of the 9 nodes of the cluster to make it flexible and scalable.

The performances of the cluster are then studied in a 3-phase evaluation consisting of: first, an individual evaluation phase ; second, an optimization phase ; third, a cluster evaluation phase.

The major steps of these evaluations are: 1) to collect 31 attributes on the RPIs that allow to build a default state of their performance ; 2) to evaluate the RPIs after the modification of their configuration in order to provide a visualization of the performance gains and/or losses, according to several metrics ; 3) to evaluate the RPIs in a parallel computing environment where the 9 RPIs are articulated around a common task

The results of these three phases are then presented with emphasis on: the performance similarities and/or divergences encountered within the 9 RPIs; the gains and/or losses granted by each optimization and each modification made; the processing power of the cluster expressed in GFLOPS; the gains and/or losses in GFLOPS of the cluster granted by the optimizations; and the relation of the results with previous studies.

The document is broken down as follows: **1.** *Introduction*; **2.** *Theory* ; **3.** *Literature study* ; **4.** *Motivations* ; **5.** *System description* ; **6.** *Evaluation* ; **7.** *Implementation* ; **8.** *Experimental Results* ; **9.** *Future works* ; **10.** *Conclusion*