Sergi Aliaga

Northeastern University

Homework 1

Question 4

Find the current listing (58th) for the Top500 list for November 2021.Review the system architecture for each of the top 10 systems on this list. Provide a summary of trends that you find and discuss what kind of system you would design in the future so that it could be included in this impressive list. Make sure to provide a diagram of the architecture you would develop, and include details of the CPUs, memory and interconnect used.

The Top500 list is the list of the 500 most powerful commercially available non-distributed computer systems available. It is renewed 2 times each year (June and November), coinciding with the International Supercomputing Conference, in Japan, and the ACM/IEEE Supercomputing Conference, in the U.S., respectively. The ranking is based on the performance of the computer in the High Performance LINPACK (HPL) benchmark, which measures the floating-point computing power. More specifically, the benchmark measures how fast the system can solve an *NxN* system of linear equations.

From the top 10 results for November 2021 we observe the following trends:

- Strong dominance from the United States, with 5 systems in the top 10, followed by China with 3. These two are the only nations with more than one system in the top10. In addition, as an student coming from Barcelona, it felt good to see the Mare Nostrum super computer in the top 100 (74^{th}), the only one out of the top500 located in Spain.
- The following chart in Fig. 1 displays the number of cores as a function of the position in the ranking. In general, the trend is that systems with more cores report a better performance and, therefore, a better position. Computers ranked 7^{th} and 4^{th} seem to not follow this trend since they have a much larger number of cores compared to the rest but fall behind their nearest competitors. However, computers 7^{th} and 4^{th} are also the oldest systems in the ranking (2016 and 2018 respectively) which might be a sign that they are being outpaced by the more modern competitors, which use more powerful cores and/or use them more efficiently.
- The architectures used in all the top 10 are either Clustered architecture or Massive Parallel Processing (MPP) architecture, the former being the dominant with 7 systems using it.

In a Clustered architecture each node is an individual computer system working together with the rest of nodes using a shared memory. Each node is connected to each other by a variation of normal networking. In addition to increased performance, a clustered system is more fault tolerant and scalable since the loss of one node does not result in the loss of the system and it is easy to add additional nodes to the system. However, the performance of the system is highly defendant on the speed of the interconnection between nodes. On the other hand, MPP architecture consists in one unique machine with thousands of CPUs tightly interconnected. MPPs have exotic memory architectures to allow extremely high-speed exchange of intermediate results with neighboring processors. This typically results in a performance increase when it comes for certain algorithms at the expense of an increased cost in scalability and usability (programs need to be highly optimized for the specific memory architecture).

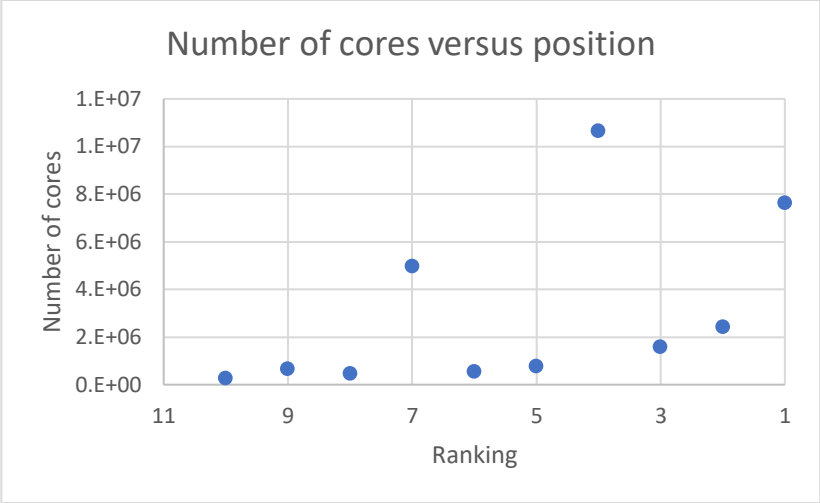

Figure 1

In the light of the advantages and drawbacks of the two most used architectures, what I would suggest as a possible improved design would be a hybrid system that included two “super-nodes”, one with MPP architecture and a Clustered one. The two nodes would be transparent to the user and an orchestrator program would oversee deriving optimized tasks to the MPP architecture and other tasks to the Clustered node. The diagram of such a system could be something like the one shown in Fig.2.

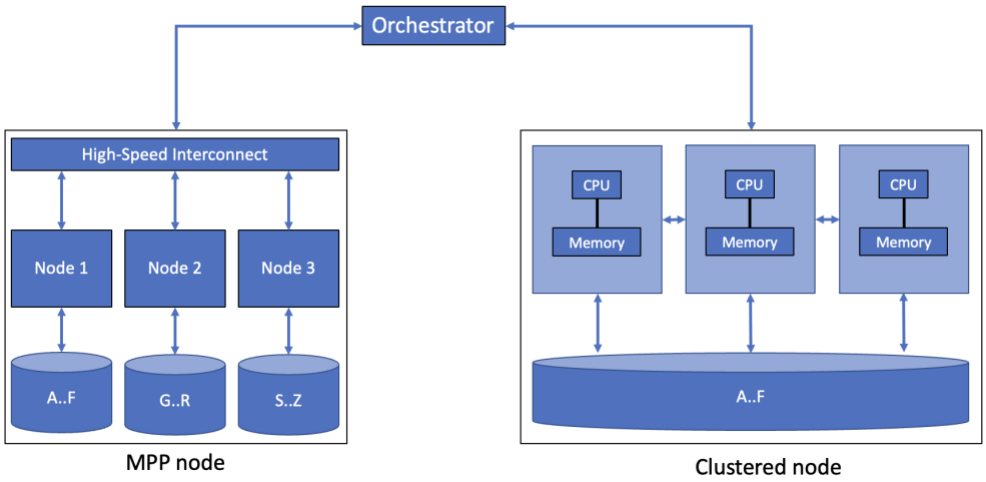

Figure 2: Possible architecture of an improved system