Abstract:
Modern high performance systems are becoming increasingly complex and powerful due to advancements in processor and memory architecture. In order to keep up with this increasing complexity, applications have to be augmented with certain capabilities to fully exploit such systems. These may be at the application level, such as static or dynamic adaptations or at the system level, like having strategies in place to override some of the default operating system policies, the main objective being to improve computational performance of the application. The current work proposes two such capabilities with respect to multi-threaded scientific applications, in particular a large scale physics application computing ab-initio nuclear structure. The first involves using a middleware tool to invoke dynamic adaptations in the application, so as to be able to adjust to the changing computational resource availability at run-time. The second involves a strategy for effective placement of data in main memory, to optimize memory access latencies and bandwidth. These capabilities when included were found to have a significant impact on the application performance, resulting in average speedups of as much as two to four times.