Xu, Qiang, “Automatic Construction of Coordinated Performance Skeletons”

Advisor: Dr. Jaspal Subhlok

The performance skeleton of a parallel application is a short running program whose performance in any scenario reflects the performance of the application it represents. Specifically, the execution time of the performance skeleton is a small fixed fraction of the execution time of the corresponding application in any execution environment. This work presents a framework for automatic construction of performance skeletons. The approach is based on capturing the execution behavior of a parallel application from traces and automatically generating a synthetic skeleton program that reflects that execution behavior. This research introduces ***coordinated performance skeletons*** that consist of a single SPMD program that executes on each node of the system.

Coordinated skeleton construction proceeds as follows. First, we convert all the physical traces into a single representative logical trace by identifying the underlying communication topology employed by the application and converting the point-to-point communication between processes in the representative physical trace into logical communication that represents similar communication across all processes executing the application. Then we compress the single logical trace into a compact coordinated execution signature by capturing the repeating patterns and discovering the underlying loop structure inside the logical trace. The compression employs a novel greedy version of Crochemore’s algorithm borrowed from bioinformatics. Finally, we use this compact coordinated execution signature as the basis for building a synthetic executable parallel skeleton program.

Experiments have been carried out to construct coordinated performance skeletons for NAS Class C benchmarks. The results show that we can accurately and efficiently identify the communication topology and logicalize the execution traces. The longest physical trace of 23.5 million lines was logicalized in 134 seconds. Greedy compression of the logical traces, averaging 71695 events, took 16.5 seconds for a compression ratio of 1815, on average. The prediction accuracy of synthetic coordinated performance skeletons was tested on 1) the same cluster with different CPU and network sharing and 2) different clusters with different hardware architectures. The results demonstrate that a coordinated performance skeleton can predict the application execution time fairly accurately. The measured average prediction error was 4% on the same cluster and 9.1% across different clusters.