

Pipeline Analysis

Nikhil Shenoy

May 24, 2015

Purpose

Characterize the Pipeline pattern, and eventually all other EnsembleMD Patterns, using factors such as pilot size, number of pipeline instances, and number of steps per pipeline. Result of characterization should be a number of plots that describe the execution time of each pattern under a certain configuration of the input parameters

Procedure

1. Using the "Idle" kernel as an initial task and a walltime of 15 minutes, run the Pipeline pattern while scaling parameters with the following patterns:
 - Pilot size = [16,64,128]
 - Number of Pipeline Instances = [16,64,128]
 - Number of Steps per Pipeline = [16,64,128]
 - Additional parameters should follow the same scaling pattern
2. Iterate through the scaling parameters such that all n^3 configurations, where n is the number of factors, are tested.
3. Repeat each experiment until the results are consistent. The arithmetic mean of all execution times will be the representative value of the input configuration.
4. Using the data recorded in the `execution_profile_dictionary`, record the percentage of total execution time that is taken up by each stage of execution (stage-in, execution, stage-out).

Expected Results

1. With all other factors held constant, execution time is expected to decrease as the Pilot Size is increased.
2. With all other factors held constant, execution time is expected to stay constant as the Number of Pipeline Instances is increased.
3. With all other factors held constant, execution time is expected to increase as the Number of Steps per Pipeline is increased.