Programming Assignment 4 Solution

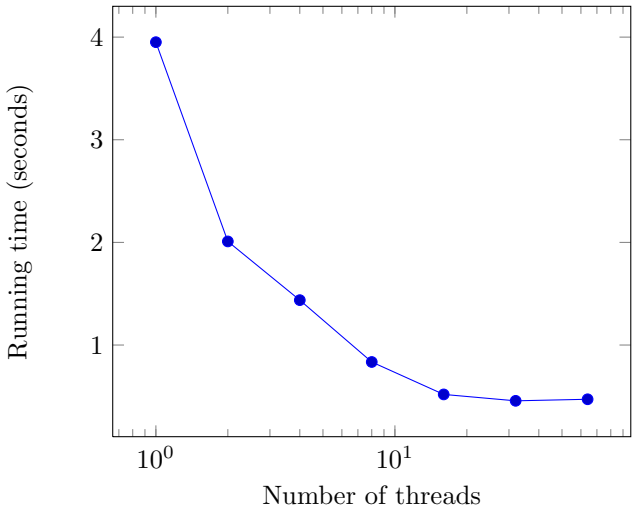
May 7, 2012

Code

A sample implementation is available on the Piazza homepage.

Analysis

1. Memory bound! The computation is essentiall neglible.
2. Scales with number of processors until about 16 threads, then plateaus.



3. Radix sort parallel actually improves bandwidth (super linear scaling) over serial implementation, because the large dataset fits into the distributed caches of the program but does not fit into a single cache for the serial implementation.

Bandwidth (MB/s)	Small	Large
Serial	69.3	52.6
Parallel	150.22	237.453

4. This question also touches on the idea of memory being distributed across caches.
   - It is helpful as long as we have threads waiting on a memory request. If our dataset is so large that it does not even fit into main memory, than it may be helpful to generate many threads, so that something is always running while disk requests are being serviced.
   - It is not helpful when our dataset is small enough to fit into local memory, and will hurt our performance by generating cost in creating threads and context switching.