Impact of Solid State Drives on Gaussian Performance

Yvonne Radskimham
University of California, San Diego
La Jolla, CA 92093
yvonne.radsmikham@gmail.com

Robert S. Sinkovits
San Diego Supercomputer Center
University of California, San Diego
La Jolla, CA 92093
sinkovit@sdsc.edu

ABSTRACT

Gaussian 09 is one of the most widely used computational chemistry programs, both worldwide and on XSEDE resources. Most Gaussian calculations generate intermediate files, thereby making the performance of the application dependent on the attributes of the local scratch file system. To determine the impact of the choice of solid-state drive (SSD) and hard drive (HDD) on run time, we benchmarked the entire Gaussian test suite, consisting of 880 problems spanning a variety of calculation types and molecule sizes, on four different scratch storage configurations. Although this may not be relevant to a particular user or research group, these problems are presumably representative of the broad computational chemistry workload that can be expected on a shared national resource. The run times on a single Sandy Bridge core range from just over a second to nearly six hours, with the majority of calculations taking on the order of seconds to minutes to complete. Scratch configurations include the Intel Lyndonville SSD, exported via the iSER protocol on the production Gordon system, and locally mounted Intel Lyndonville and Taylorsville SSDs and Western Digital Velociraptor HDD on the Gordon test cluster.

When comparing locally and iSER mounted Lyndonville drives, we find that roughly 75% of the Gaussian tests have nearly identical run times (relative difference of less than 2%) on the two configurations. A small number of the test cases were noticeably faster when using the iSER exported drives, four of which had very short run times (< 2 seconds) and two others corresponding to test problems with highly variable timings using the iSER drives. A long tail of test problems had run times that were 4-25% longer under the iSER, with number of affected jobs rapidly decreasing at longer relative run times. We considered the possibility that the results were skewed by very short running jobs, but omitting these does not significantly alter the outcome.

The results of our benchmarks comparing locally mounted SSDs and HDDs show that most of the Gaussian test problems (88%) were not measurably affected by the choice of drive. Additionally considering cases where the impact was minimal (less than 4% difference in run time) accounts for 97% of the test problems. It should be noted though that there is still a small, yet prominent, tail of jobs where the SSDs still had a clear advantage. A careful inspection of this tail also indicates that these jobs exhibit high variability in run time and that the worst-case performance on the HDDs is much worse than that for the SSDs.

XSEDE '13, July 22–25, 2013, San Diego, CA, USA.