WINDOWS DISK IO BENCHMARK TEST RESULTS

FOR JOYENT
Revision 4

February 10th, 2011

Scope:

This report summarizes the Disk Input Output (IO) benchmark testing performed in January of 2011 for Joyent Windows cloud servers.

References:

[1]: http://blog.cloudharmony.com/2010/06/disk-io-benchmarking-in-cloud.html

[2]: Svn repository: https://svn.codespaces.com/ims/joyent-windows
username: joyent **password:** joyent

[3]: Raw test data: DISK-IO_Final_Results.xlsx

[4]: Windows test list: Windows_Tests.xlsx

Joyent Disk IO Benchmark Testing Report

Introduction

The Disk IO testing summarized in this report was performed as part of a larger benchmark effort intended to provide a basis for comparison between the Joyent Windows servers and the virtual servers offered by other cloud service providers.

Earlier in 2010, CloudHarmony engaged in an extensive benchmarking effort intended to provide "information and analysis to enable educated decisions pertaining the adoption of, and migration to cloud services". Their results and analysis are presented in a series of articles published online ref[1].

However, CloudHarmony did not include the Joyent Windows servers in their benchmark testing. Our testing procedures are intended to follow CloudHarmony's procedures as closely as possible to extend the benchmark testing to include the Joyent servers.

The CloudHarmony benchmark testing uses tools and tests that were primarily intended for a Linux-based platform, not all of which are available on the Windows platform. Thus, it should be noted that not all test versions and executables used to generate the data in this report are the same as those used by CloudHarmony due to differences in operating systems, therefore these results can not be compared side-by-side to the CloudHarmony results. Our formulas for calculating the baseline and individual server instance performance numbers are, however, identical.

Instead of trying to reproduce all of the CloudHarmony results, we focused on those outlined for the Amazon EC2 servers and Storm on Demand's Bare Metal Cloud Server that was used as the baseline for the benchmark tests ref[1]. Our tests closely approximate the methods from CloudHarmony in regards to calculations and tests used. CloudHarmony standardized on the CentOS (64 bit) operating system for the baseline tests except when unavailable. The Joyent servers run Windows Server 2008 R2 Enterprise (64 bit).

The Joyent servers provide a "bursting" capability that allows a service to use more processor resources on a temporary basis than the guaranteed minimum. This

differs from nearly all other cloud providers that provide a fixed processor configuration. While bursting capability can be a tremendous advantage to an operational system, it can complicate benchmark testing which attempts to stress the system under test to its maximum capacity. Thus the bursting capability can greatly affect the performance scores on many benchmark tests. On the Joyent Windows servers, the bursting capability allows a process on even the smallest server to potentially use nearly the entire compute capability of the underlying hardware.

Our conclusions show a comparison between Joyent's 8GB Windows server and Amazon's EC2 c1.xlarge instance. The virtual machine and underlying hardware for these systems is at least nominally similar and provides the best basis for comparison.

Benchmark Setup

Storm on Demand's Bare Metal Cloud Server was used as the baseline for all Disk IO tests. We compared both Joyent and Amazon servers to the Storm on Demand server baseline.

The EC2 servers used consist of: m1.small, c1.medium, m1.large, m1.xlarge, m2.xlarge, c1.xlarge, m2.2xlarge, m2.4xlarge. The Storm on Demand and Amazon servers – 8 servers in 4 regions – were configured identically in terms of OS, CentOS 5.4 64-bit (or 32-bit in the case of EC2 m1.small and c1.medium where 64-bit is not supported).

Joyent servers ran Windows Server 2008 R2 Enterprise (64 bit) and included the following sizes: 4GB, 8GB, 16GB.

CloudHarmony makes use of the Phoronix Test Suite, a Linux-based benchmarking tool to streamline testing procedures. The tool uses shell scripts to download, unpack, compile, install, and run benchmark tests. Because these scripts rely heavily on Linux-type system calls, the Phoronix Test was found to be incompatible with Windows.

Instead, a Windows batch (.bat) file was created and used to run the same suite of tests using the same set of parameters as the Phoronix Test suite. A native Windows port of each test has either been downloaded or recompiled for a Windows environment ref[4]. In some cases, a test was excluded because a Windows port was unavailable due to dependencies on libraries not available in Windows. The .bat file is part of the Windows benchmark zip file that is downloaded on the system to test.

Benchmark Tests

There are 8 benchmarks CloudHarmony used to measure disk IO performance. Of the 8, 6 had dependencies on libraries not available in Windows and were excluded from this report. Our testing report does not include:

Bonnie++, Blogbench, DBENCH, Flexible I/O tester, Threaded I/O tester, hdparm-read

The following tests were run on the Joyent Windows servers:

iozone-read, iozone-write

Testing Procedures

A zip file (window_benchmark.zip) ref[2] was downloaded and extracted on each server. A batch file (run_benchmark.bat) can be used to automate running of a suite of tests and gathering of results.

The following test procedure should produce similar or identical test results:

1. Extract windows_benchmark.zip into a local directory.
2. Change to the Windows benchmark directory and run the disk IO suite of tests by running the batch file with the following arguments at the command line:

   run_benchmark.bat suite disk

   Output is logged to a results directory with the current date and timestamp.

3. Once the test complete, the output for each test can be viewed in the results directory. ResultsParser.exe is a command line executable that can be used to parse the numbers from each test and average the results into a CSV file. To run the parser, run the following at the command line:

   ResultsParser.exe [ResultsDirectory]

Baselines

A baseline was taken from results run on the Storm on Demand's Bare Metal Cloud Server and calculated based on the methodology used by CloudHarmony. This baseline score is used to calculate a IO Performance Score (IOP) that measures relative disk IO performance between the servers under test.

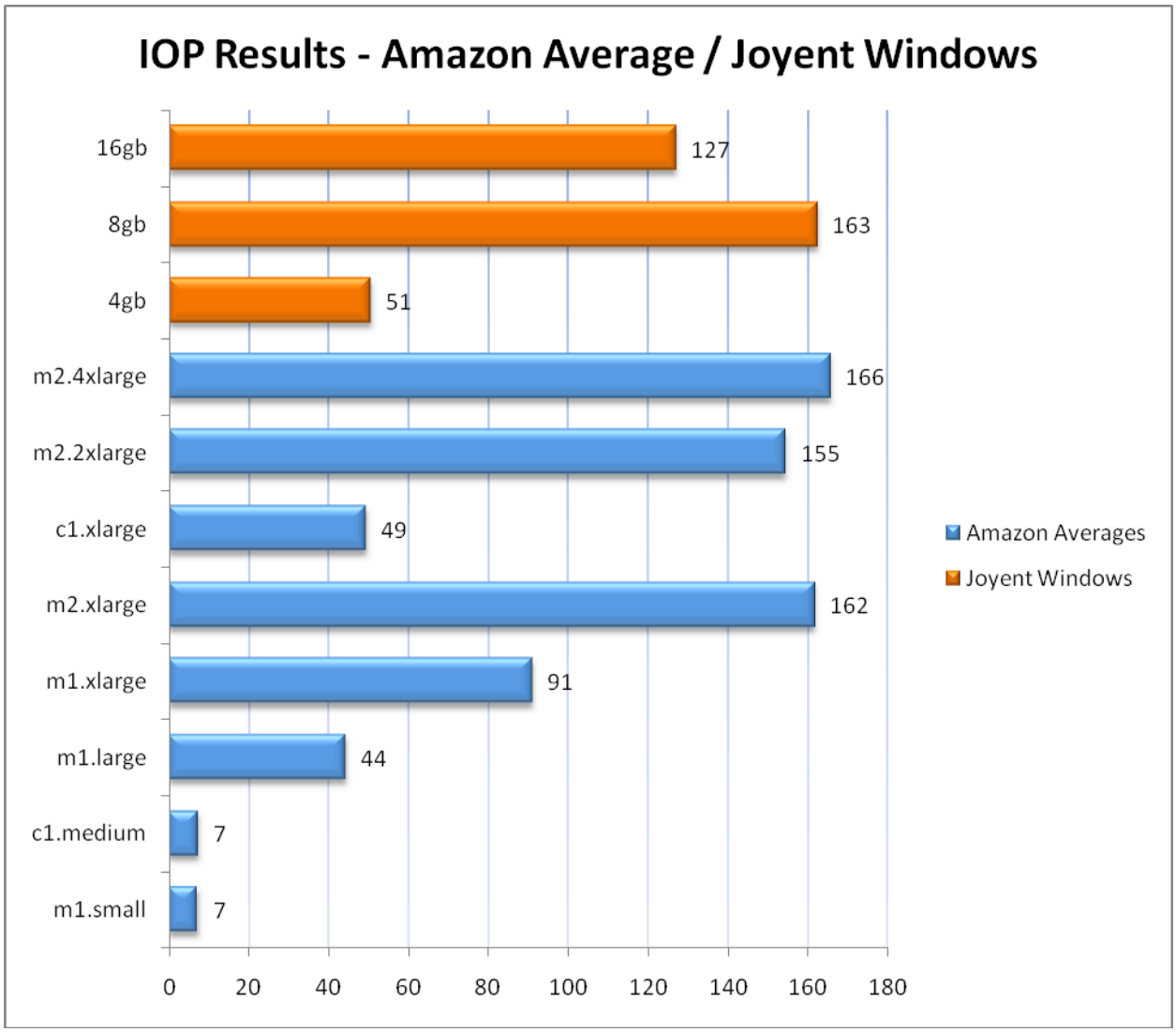
Our calculations are the same used by CloudHarmony, but exclude the six tests which required Linux specific functionality: *Bonnie++, Blogbench, DBENCH, Flexible I/O tester, Threaded I/O tester, hdparm-read.*

Test Results

Overall, the Joyent Windows servers showed a performance increase over that of Amazon EC2. As expected, neither of the virtual servers achieved the same level of disk IO performance as the baseline performance of the Storm on Demand dedicated physical hardware that scored the highest out of any server benchmarked at 692 IOP. This higher performance of the Storm on Demand server is convenient

since the IOP score calculations require a baseline higher than that of the servers under test.

The following graphs show the test results for Amazon EC2 and the Joyent Windows Servers. For accurate comparison between server resources, the 8GB and c1.xlarge have 8 CPU cores, with the only difference of 7GB of memory for c1.xlarge. It should be noted that the graph shows an average of all IOP scores from the different Amazon regions for ease of comparison.

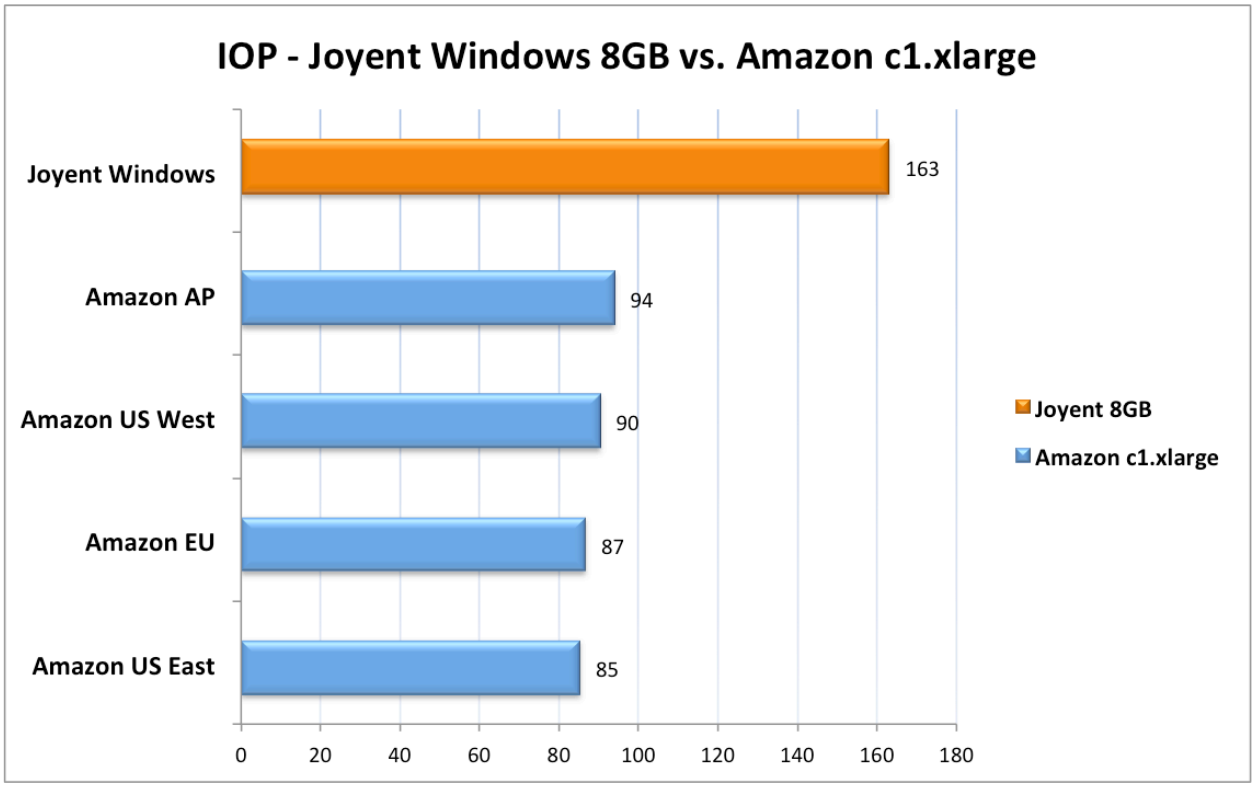


Conclusion

The results show the Joyent Windows servers outperforming Amazon's server for Disk IO in comparable machines. Amazon shows a linear increase of scores based on the server's available virtual resources. An increase is seen between Joyent's 4GB to 8GB, but not in the 8GB to 16GB server. This may reflect the expected range of variation in test data.

Note that there are many differences between the operating systems of the servers under test (CentOS vs. Windows), as well as in the test executables and versions. These differences in architecture, compilers, and libraries make direct comparison difficult between the numbers in the Joyent and Amazon/StormOnDemand test results. However, for the Disk IO tests used, we believe that the test results reflect the actual performance of the different disk IO sub-systems involved and are not skewed dramatically by CPU or memory configurations.

For the most accurate comparison, we highlight the comparison between the Amazon EC2 c1.xlarge instance and Joyent Windows 8GB servers. Both servers have similar RAM – 7GB versus 8GB – and the same available 8 CPU cores. The exact underlying disk hardware for each server is the only unknown and uncontrollable factor for these benchmarks. However, this is not a critical issues since this is an essential characteristic of the disk IO sub-systems under test and relative performance, the metric that is of most interest, will inherently take into account all the differences between the systems including physical disk hardware.



Averaging the Amazon results puts the c1.xlarge instance score at 89 IOP, while the Joyent Windows 8GB server more than doubles it at 163. The major difference between the two servers could be due to Joyent's underlying configuration of disk hardware, physical IO interface and opertiing system IO caching.