WINDOWS MEMORY I/O BENCHMARK TEST RESULTS

FOR JOYENT
Revision 4

February 10th, 2011

Scope:

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

References:

[1]: http://blog.cloudharmony.com/2010/06/cloud-server-benchmarking-part-4-memory.html

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

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

[4]: Windows test list: Windows_Tests.xlsx

[5]: http://www.cs.virginia.edu/stream/ref.html

[6]: http://www.streambench.org

Joyent Memory IO Benchmark Testing Report

Introduction

The Memory 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 busting 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

NewServers Jumbo server configured with dual Intel E5504 quad core 2.00 GHz processors and 48GB DDR3 ECC ram was used as the baseline for all Memory IO tests. We compared both Joyent and Amazon servers to the NewServers Jumbo 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).

The Joyent Windows server benchmarks were run on Windows Server 2008 R2 Enterprise (64 bit) on 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 6 benchmarks CloudHarmony used to measure memory IO performance. Of the 6, 3 had dependencies on libraries not available in Windows and were excluded from this report. Our testing report does not include:

unixbench, redis-benchmark, cachebench

The following 3 tests were run on the Joyent Windows servers:

ramspeed, geekbench, stream

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. Geekbench was run independently from the other tests and its results were manually recorded.

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 memory IO suite of tests by running the batch file with the following arguments at the command line:

   run_benchmark.bat suite mem

   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]
4. Since the Geekbench benchmark relies on a GUI, it could not be scripted as part of the batch file and must be run independently. An installer is required to setup Geekbench (setup/Geekbench21-Setup.exe). Once installed, launch the application and select "Run Benchmarks". Save the results for future reference.

Baselines

A baseline was taken from results run on the NewServers Jumbo Server and calculated based on the methodology used by CloudHarmony. This baseline score is used to calculate a Memory IO Performance Score (MIOP) that measures relative memory IO performance between the servers under test.

Our calculations are the same used by CloudHarmony, but exclude the three tests which required Linux specific functionality: *unixbench, redis-benchmark, cachebench.*

The NewServers Jumbo server was assigned a score of 100. All other servers were assigned a score proportional to the performance of that server, where greater than 100 represents better results and less than 100 represents poorer results. For example, a server with a score of 50 scored 50% lower than the baseline server overall, while a server with a score of 125, scored 25% higher.

Test Results
The test scores of MIOP were compared to those of the NewServers baseline and a ratio was created based on a percentage of each result to that of the baseline.

Overall, tests show the Joyent Windows Servers performed as well as the NewServers baseline. These results are surprising due to the NewServers Jumbo server running without virtualization and direct access to the physical hardware.

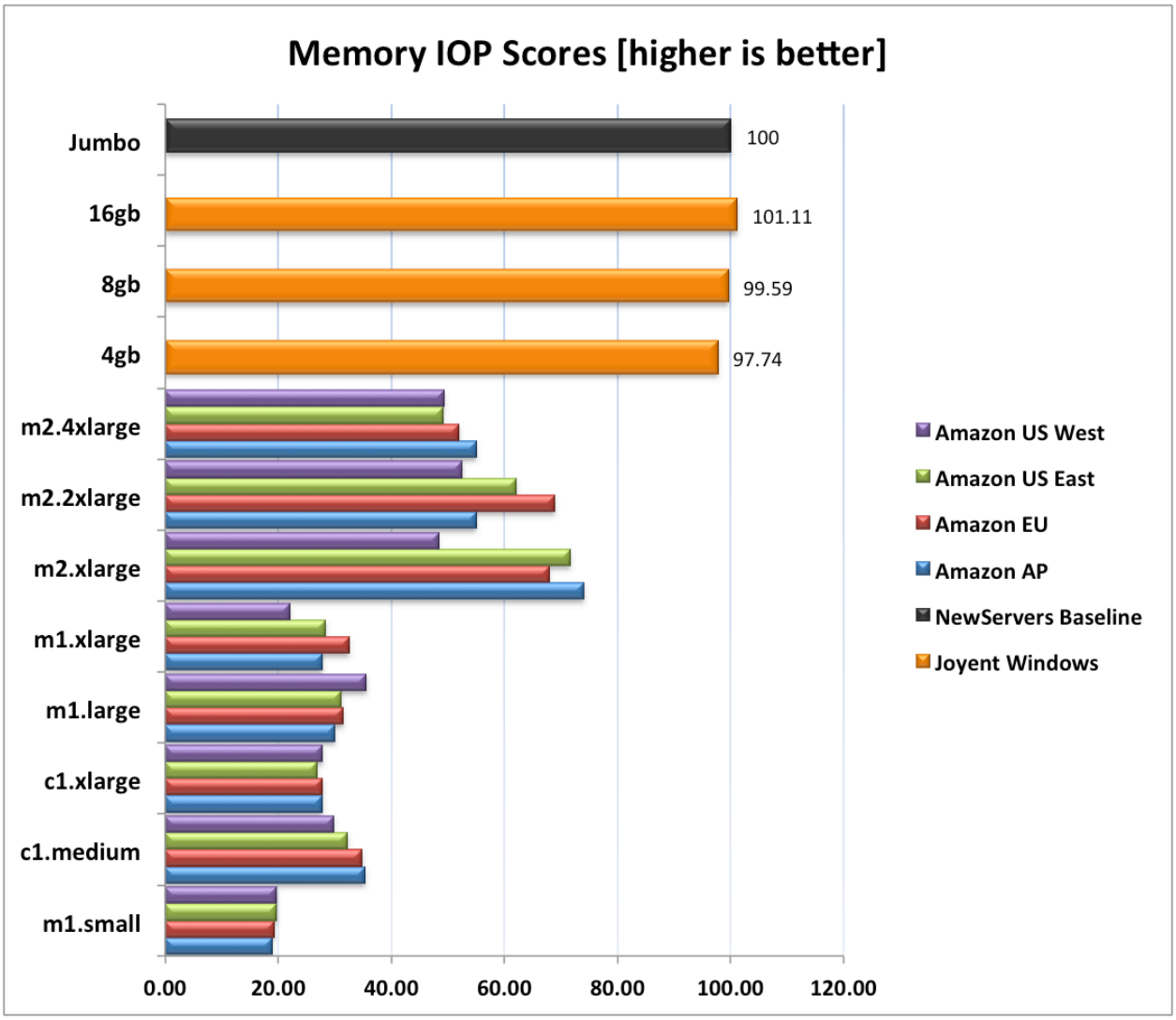
The two following graphs and data illustrate the above calculations and results:

Memory IO, calculated as MIOP ref[1], for the Joyent Windows Servers:

	m1.small	m1.large	m1.xlarge	m2.xlarge
Amazon AP	18.83	30.00	27.80	74.05
Amazon EU	19.22	31.48	32.50	67.89
Amazon US East	19.69	31.09	28.38	71.59
Amazon US West	19.67	35.59	22.09	48.33

	m2.2xlarge	m2.4xlarge	c1.medium	c1.xlarge
Amazon AP	54.96	54.98	35.31	27.80
Amazon EU	68.86	51.96	34.85	27.80
Amazon US East	62.07	49.08	32.20	26.84
Amazon US West	52.55	49.27	29.84	27.81

	4gb	8gb	16gb
Joyent Windows	97.74	99.59	101.11



Conclusion

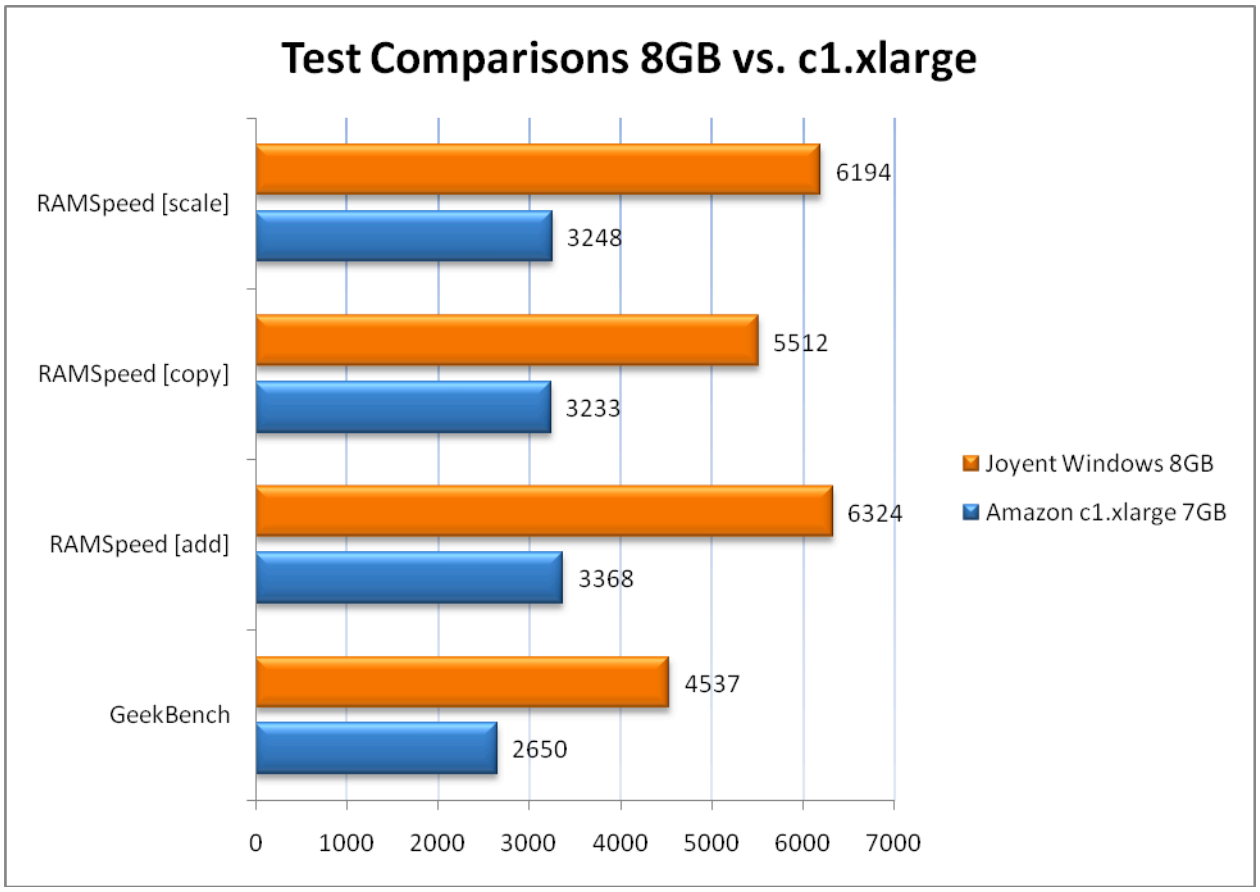
The variation in MIOP values between the 4GB and 16GB Joyent Windows servers did not show a strong linear increase in memory IO performance with the increase in server capacity. Overall, the Joyent servers outperformed all the Amazon server instances in our comparisons.

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/ NewServers test results. However, for the Memory IO tests used, we believe that the test results reflect the actual performance of the different memory IO sub-systems involved and are not skewed dramatically by CPU configurations.

While comparing memory throughput to the NewServers baseline, the Joyent 16GB server performs as well as the baseline, which is surprising due to the Jumbo server having 48GB of available memory and improved hardware.

The two most comparable server instances are the Joyent 8GB and Amazon c1.xlarge. We compared the c1.xlarge highest score from the US West region of Amazon. Both servers have an available 8 CPU cores and 7-8GB of available memory. Unfortunately, Amazon does not offer a server instance with 8GB of memory, but this small 1GB difference should not affect memory IO throughput significantly.

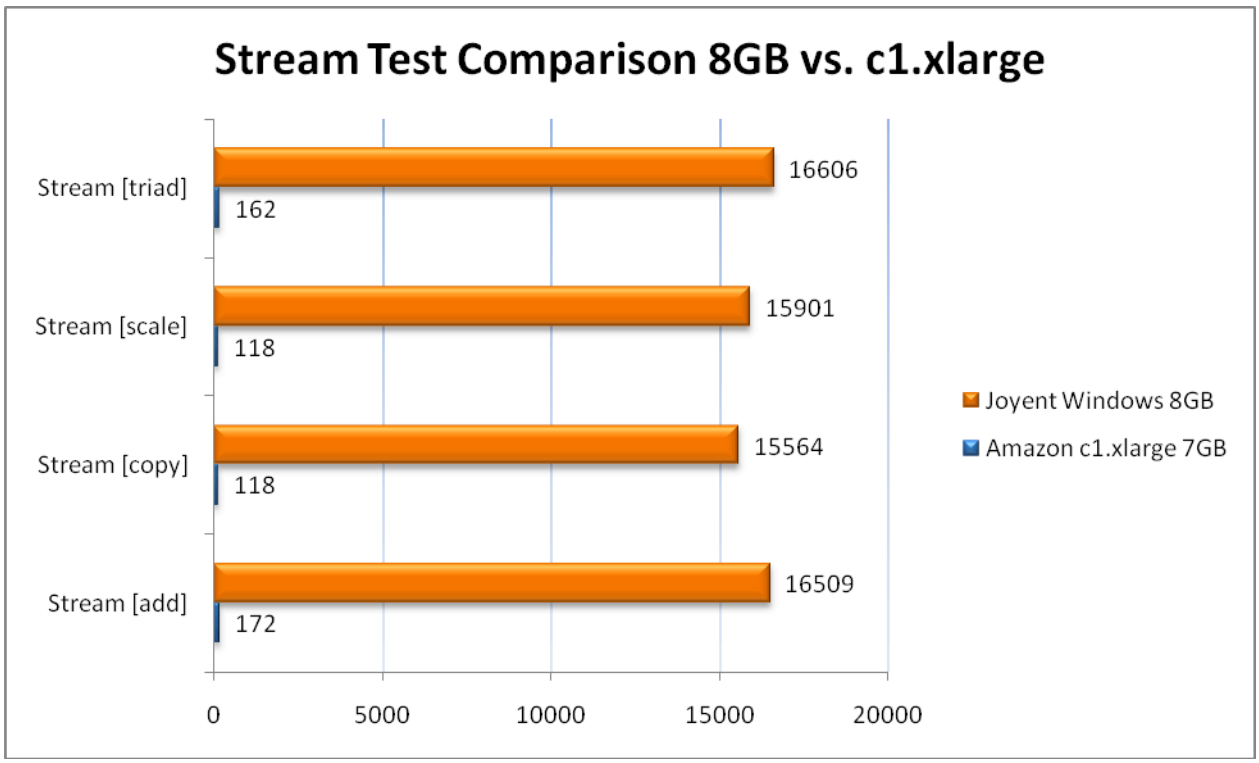
Upon analyzing the difference in scores, the Joyent Windows 8GB server outperformed the Amazon c1.xlarge. The following graphs give a visual representation of each test performed:



The most surprising result comparison is with the Stream benchmark. From the developer website, this benchmark is described as:

> *The STREAM benchmark is a simple synthetic benchmark program that measures sustainable memory bandwidth (in MB/s) and the corresponding computation rate for simple vector kernels. Ref[5].*

The STREAM Benchmark is the de facto industry standard benchmark for the measurement of computer memory bandwidth. The STREAM benchmark measures "real world" bandwidth sustainable from ordinary user programs -- not the theoretical "peak bandwidth" provided by most vendors. The STREAM benchmark archive currently has more than 1000 results, dating from 1991 to the present. Ref[6].



The difference in hardware and available operating system technologies put the Joyent Windows servers above the comparisons to Amazon EC2 default instances.

Additionally, Amazon servers showed a consistent increase in the MIOP values in the test results from small to large based on server capacity. However, individual memory tests showed variance between the servers which did not follow a linear pattern. These results are comparable to those found on the CloudHarmony blog. The Joyent Windows servers show a small variation between servers, but the throughput of memory IO is consistent with their technology.