



WINDOWS CPU COMPARISON BENCHMARK TEST RESULTS

FOR JOYENT
Revision 4

February 10th, 2011

Scope:

This report summarizes the CPU benchmark testing performed in January of 2011 for Joyent Windows cloud servers.

References:

[1]: <http://blog.cloudharmony.com/2010/05/what-is-ecu-cpu-benchmarking-in-cloud.html>

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

[3]: Raw test data: CPU_Final_Results.xlsx

[4]: Windows test list: Windows_Tests.xlsx

Joyent CPU Benchmark Testing Report**Introduction**

The CPU 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

Amazon EC2 was used as our primary baseline benchmark for all CPU tests.

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 19 benchmarks CloudHarmony used to measure CPU performance. Of the 19, 7 had dependencies on libraries not available in Windows and were excluded from this report. Our testing report does not include:

espeak, mafft, nero2d, opstone-svd, opstone-svsp, opstone-vsp, unixbench

The following 12 tests were run on the Joyent Windows servers:

c-ray, crafty, dcraw, geekbench, graphics-magick, hmmer, john-the-ripper-blowfish, john-the-ripper-des, john-the-ripper-md5, openssl, sudokut, tscp,

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 cumulative baseline was taken from results run on the Amazon EC2 servers and calculated based on the methodology used by CloudHarmony. A cumulative baseline was taken from all Amazon results and calculated based on the methodology from CloudHarmony.

Our calculations are the same used by CloudHarmony, but exclude the seven tests which required Linux specific functionality: *espeak, mafft, nero2d, opstone-svd, opstone-svsp, opstone-vsp, unixbench.*

Test Results

The raw test data is presented below and also referenced in `CPU_Final_Results.xlsx` ref[3]. The following data can be used as basis for comparison against a future baseline.

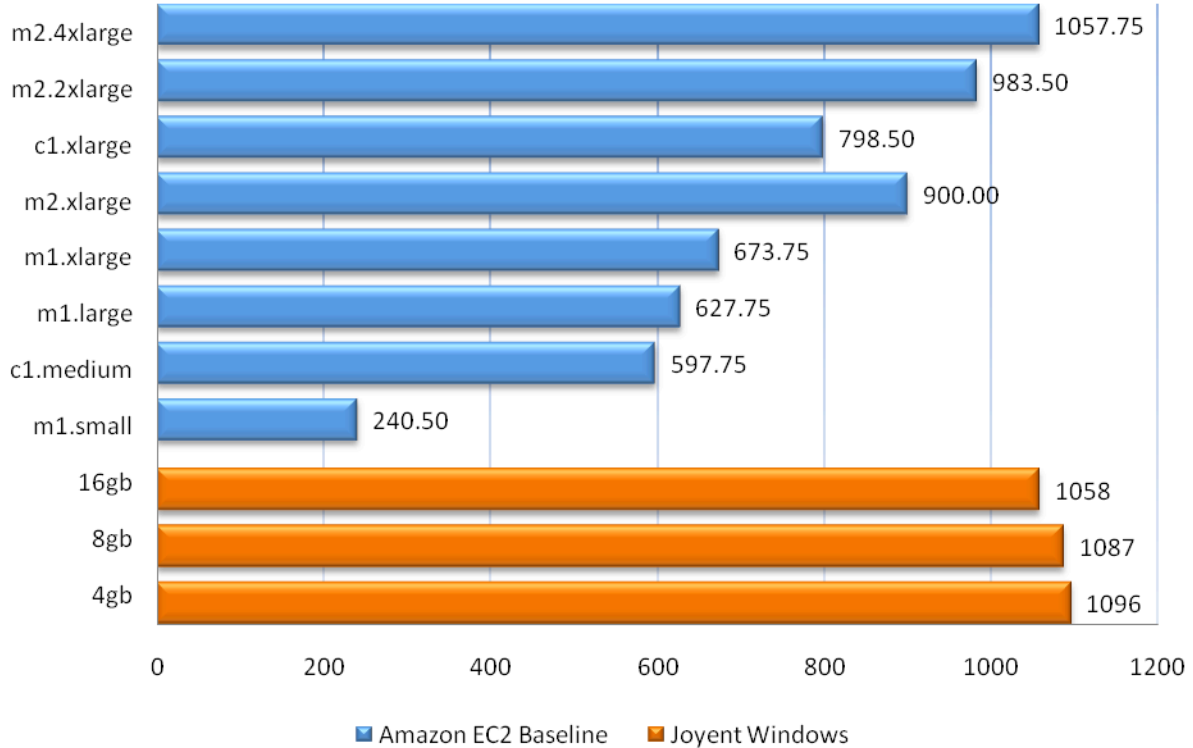
	4gb	8gb	16gb
C-Ray	106.16	107.21	109.55
dcraw	33.67	35.92	37.14
GeekBench	4662.33	4716.33	4440.33
GraphicsMagick - HWB Color Space	6.00	6.00	6.00
GraphicsMagick - Blur	6.00	6.00	6.00
GraphicsMagick - Local Adaptive Thresholding	33.11	31.67	30.22
GraphicsMagick - Resizing	15.00	14.83	14.28
GraphicsMagick - Sharpen	5.00	5.00	5.00
GraphicsMagick	65.11	63.5	61.5
Timed HMMer Search	308.58	310.17	312.54
John the Ripper - DES	284138.44	285527.94	284053.44
John the Ripper - MD5	7578.89	7533.67	7497.39
John the Ripper - Blowfish	340.61	340.50	339.78
OpenSSL	38.24	38.14	37.51
Sudokut	68.68	69.69	71.46
TSCP	228642.10	228205.57	225222.07
Crafty	117.22	116.51	119.38

The following graphs give a visual comparison between the Amazon EC2 CentOS operating system and the Joyent Windows platform.

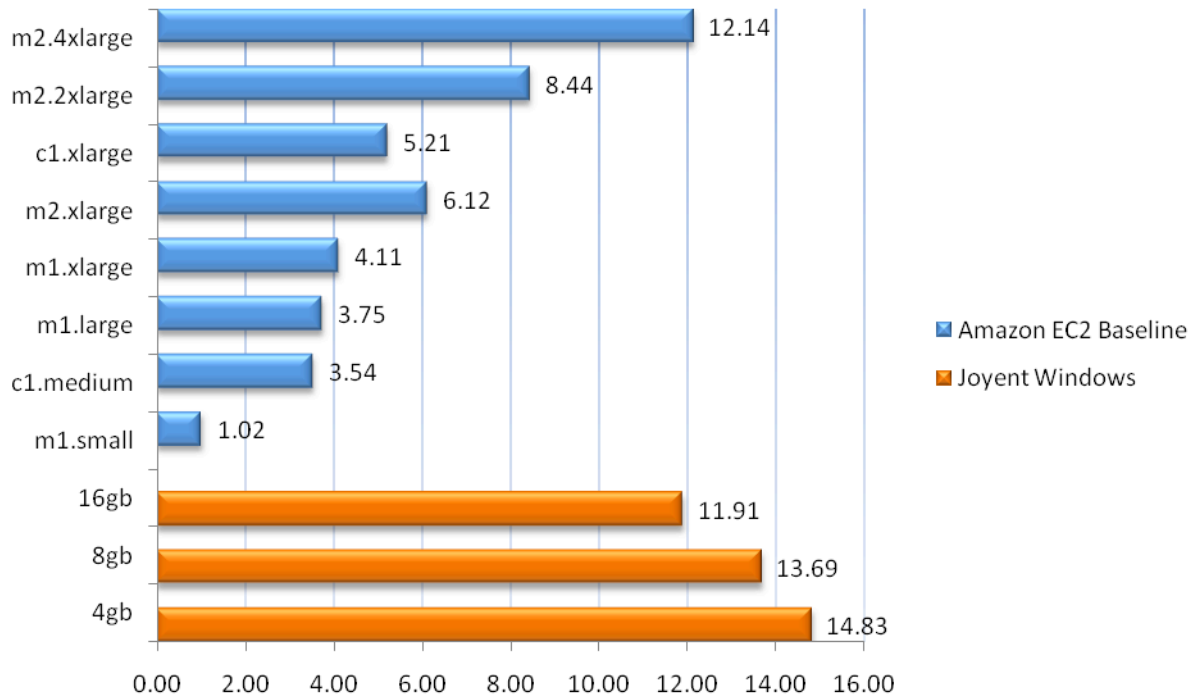
Overall the Windows systems surprisingly scored higher than that of CentOS.

These scores should not be identically compared to those on the CloudHarmony blog but can be used as a complementary set of results.

CPU Benchmark Scores - CCS [higher is better]



CPU Benchmark Scores - CCU [higher is better]



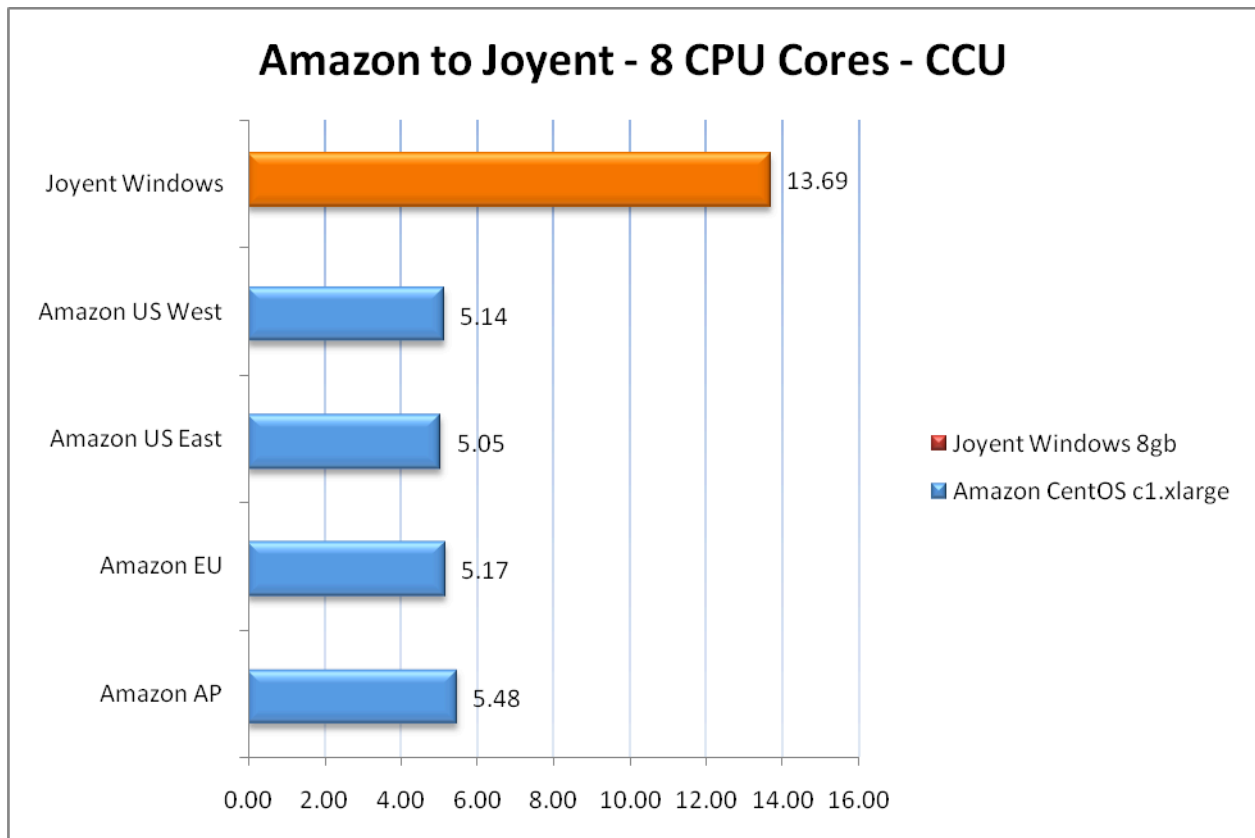
Conclusion

The Joyent Windows server instances show a high score compared to those of Amazon EC2 CentOS. Since the test binaries are of a different nature, it is still surprising that the Windows machine, which usually scores less than Linux, is actually above or close to the m2.4xlarge instance.

We can only speculate that the difference between compilers, system binaries and the underlying hardware, are the cause for this relatively high benchmark score of CCS and CCU.

The most comparable two server instances are those between Amazon EC2 c1.xlarge and the Joyent Windows 8GB server. Both machines show a total of 8 available cores and a similar 7GB to 8GB of RAM. For CPU scores, the RAM difference of 1GB should show little to no significance.

These two graphs give a visual comparison between the two similar servers:



A very large difference between the two different server operating systems.

It should be noted that these comparisons are between different operating systems which contain unique kernels and system resource handling. The benchmark tests however, do calculate similar system calls to the underlying hardware.

