Maithily Garud
Neil Roberts
Sherry Sethi
Rohit Varma

CS 533: Spring 2015

Performance Tool Effects of SMM

Introduction

System Management Mode (SMM) is a special processor mode, intended for OS-independent features like advanced power management, that causes a complete CPU context switch on entry and exit. As new SMM use cases are proposed, previous assumptions about time spent in SMM and the performance effects may no longer hold.

We look at the performance effects of SMM on several workloads of common *NIX benchmarking tools with an eye to quantifying the effect of time spent in SMM using the coarse granularities of short and long. Our result show that short times spent in SMM have a negligible effect on performance, but the effect of long times spent in SMM can be dramatic.

System Management Mode

System Management Mode (SMM) is a special x86 processor mode that privileged software such as operating systems or hypervisors cannot access [Delgado,Karavanic]. It is entered through a System Management Interrupt (SMI). When an SMI occurs, the processor saves its current context, switches to a separate operating environment with its own address space, executes the SMI handler, and finally restores the context and resumes execution of previously running processes [Intel].

SMM was originally intended for advanced power management and other OS-independent features. However, new use cases have been proposed, such as using SMM to implement an additional layer of system security. Since these new use cases open the possibility of spending more time in SMM than was previously expected, it's worthwhile to investigate the effects of SMM on benchmark performance results.

Platform

All performance scores were collected on a Dell PowerEdge R410 1U server running Fedora 21. It is configured with an Intel Xeon E5520 quad-core (eight threads, with HT enabled) CPU clocked at 2.26GHz and 12GB DDR3 memory.

Tools

Stress-ng

Stress-ng is designed to exercise various physical subsystems of a computer as well as various operating system kernel interfaces. It features over 90 stress tests, at least 50 of which are specific to testing the CPU and at least 20 of which stress virtual memory [cking]. We focused on CPU, file I/O, and virtual memory [man stress-ng]:

1. CPU: calculates square roots of double-precision randomly generated numbers.
2. File I/O: calls the *sync()* sytem call to commit buffer cache to the disk.
3. Virtual Memory: continuously calls the mmap() and munmap() system calls and writes to the allocated memory.

SysBench

SysBench is a modular, cross-platform and multi-threaded benchmark tool for evaluating OS parameters that are important for a system running a database under intensive load [akopytov]. SysBench offers several different test modes: CPU, threads, mutex, memory, file I/O, and oltp. We focused on the CPU, threads, and mutex test modes:

1. CPU: calculates prime numbers up to a limit, specified in the *cpu-max-prime* parameter.
2. Threads: tests scheduler performance when a large number of threads are competing for a fixed number of mutexes. This test mode is modulated by the *thread-yields* parameter, which specifies the number of lock/yield/unlock loops to execute in each request.
3. Mutex: tests mutex implementation performance by emulating the behavior of all threads running concurrently while only acquiring the mutex for a short period of time. This test mode is modulated by the *mutex-locks* parameter, which specifies the number of mutex locks to acquire in each request.

Test Methodology

For each of our scenarios, we collected performance results with SMIs disabled, periodic short SMIs enabled, and periodic long SMIs enabled. SMI activity was controlled by the driver developed by [Delgado,Karavanic] and the driver was manipulated by single-purpose scripts (disable, enable short, enable long, get SMI count) writing to the */proc/smidriver* interface. Between scenarios, we disabled SMIs and waited for the counter to settle before enabling the type of SMI that we were investigating. Similarly, we confirmed that SMIs were active by waiting for the counter to resume before starting a scenario.

Individual scenarios were chosen based on our estimates of which were likely to show the effects of SMIs and their fit for our particular configuration. For example, we didn’t run OLTP scenarios because our system didn’t have the required database libraries and we lacked the expertise to configure them in a manner consistent with known best practices.

Stress-ng parameters (number of stressors) was capped at eight stressors because that was the program limit. We believe this was because our platform was capable of executing at most eight threads simultaneously. SysBench parameters were selected based on the observed runtime, with a practical cap at five minutes (the next step in the sequence was estimated to run for nearly an hour) to facilitate data collection.

Results

Our performance results from Stress-ng are much less consistent in their overall answer to our question than the performance results from SysBench. All of the raw data presented in the following graphs can be found in the appendix.

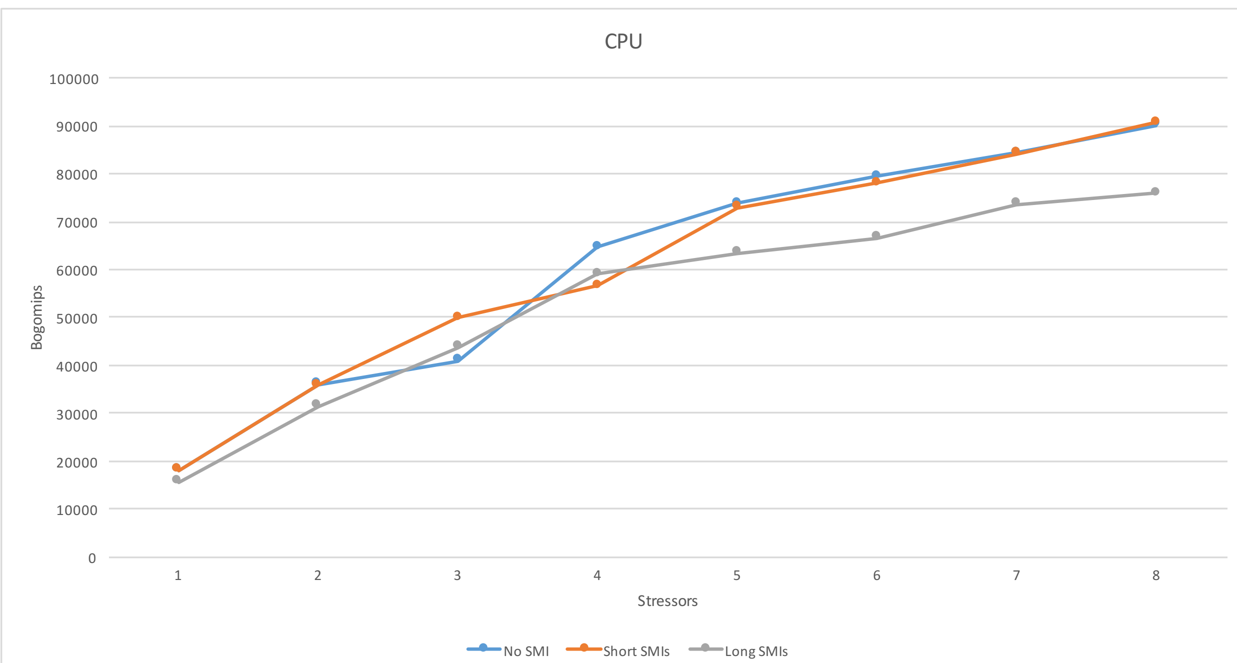

Figure 1: Stress-ng CPU

As seen in figure 1, the performance results from the Stress-ng CPU workload are basically what we might expect. Performance suffers when long

SMIs are enabled; on average, dropping about 11% compared to when SMIs are disabled. With the exception of a couple of off-the-curve points, performance is effectively unchanged when short SMIs are enabled.

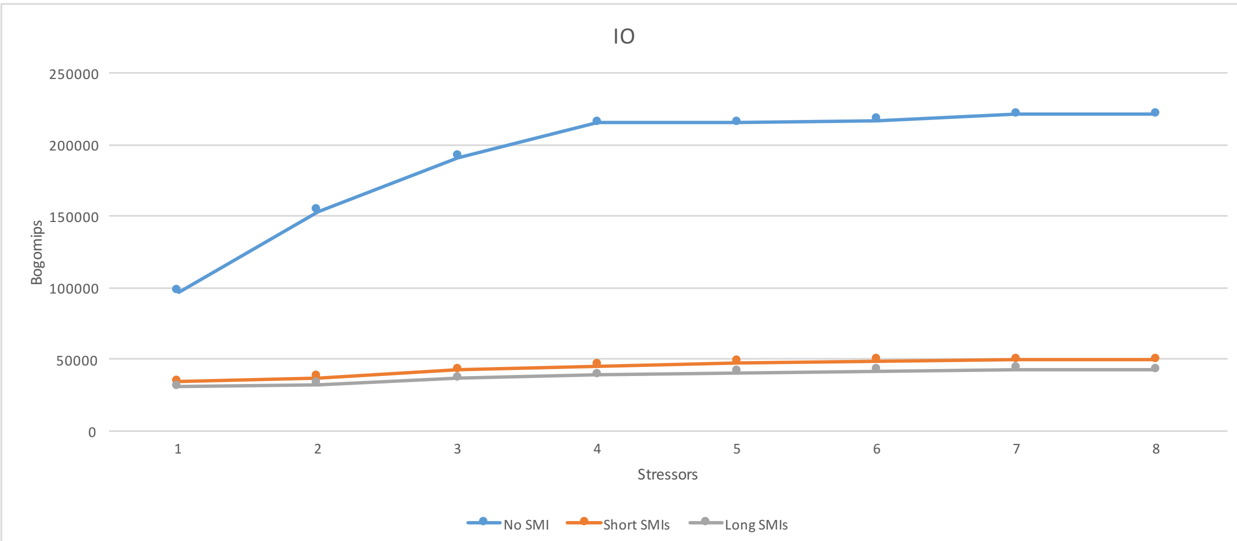

Figure 2: Stress-ng: IO

Figure 2 shows results that are not in line with any of our other results. When either short or long SMIs are enabled, performance plummets, dropping 76 – 79% on average. We believe this is because file I/O is particularly sensitive to the recurring CPU context switches.

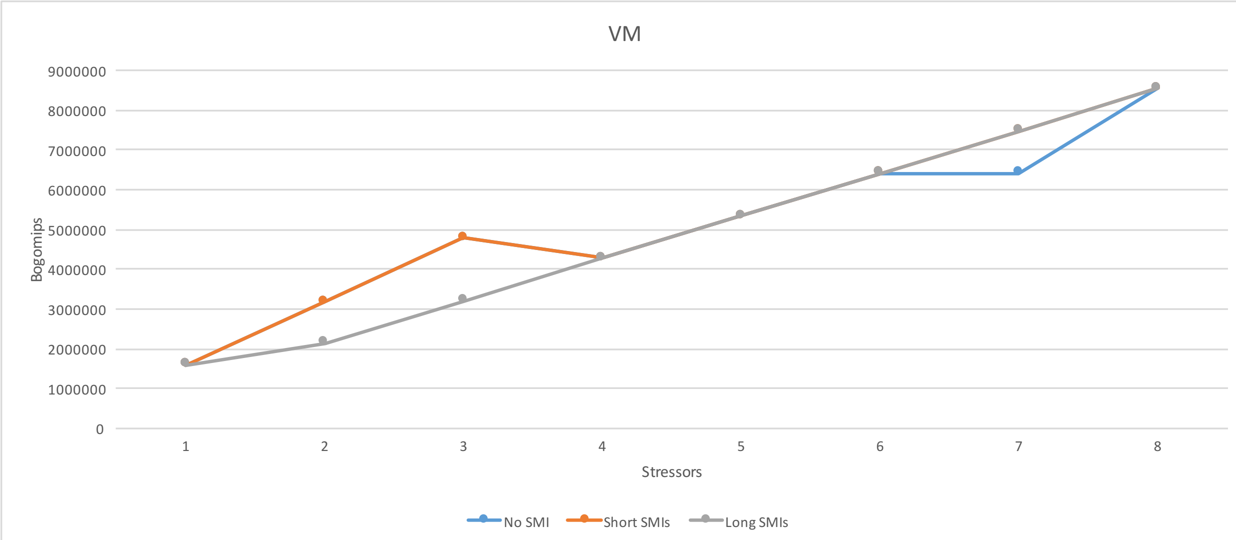

Figure 3: Stress-ng Virtual Memory

Yet another type of behavior is seen in figure 3. The performance results are more-or-less unchanged with SMIs enabled, +2 to -6% on average. This suggests that, unlike file I/O, virtual memory operations don't exhibit sensitivity to recurring CPU context switches.

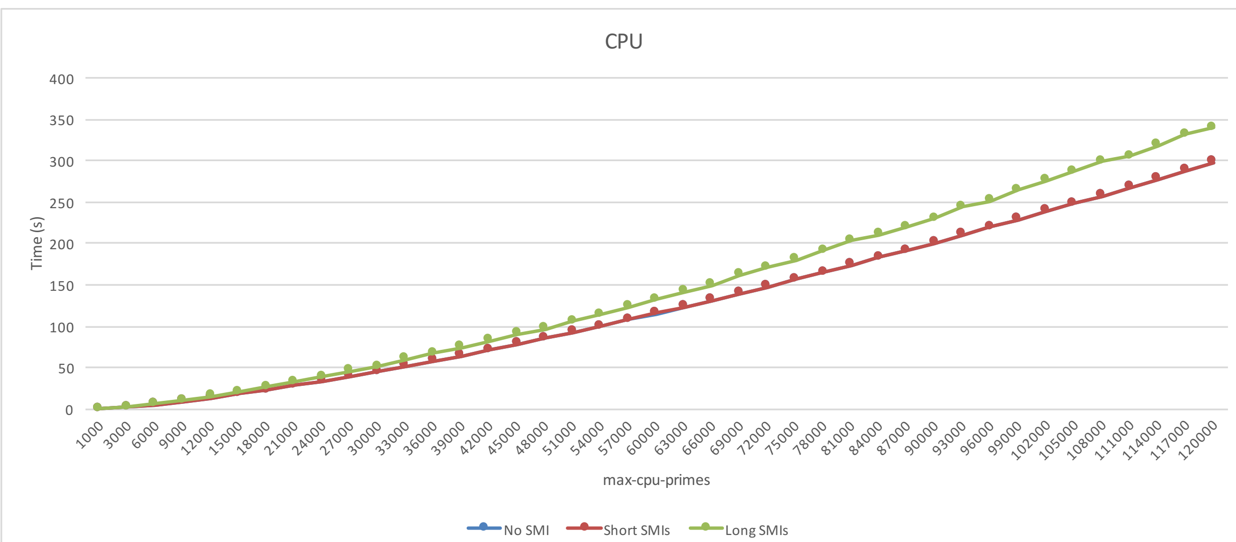

Figure 4: SysBench CPU

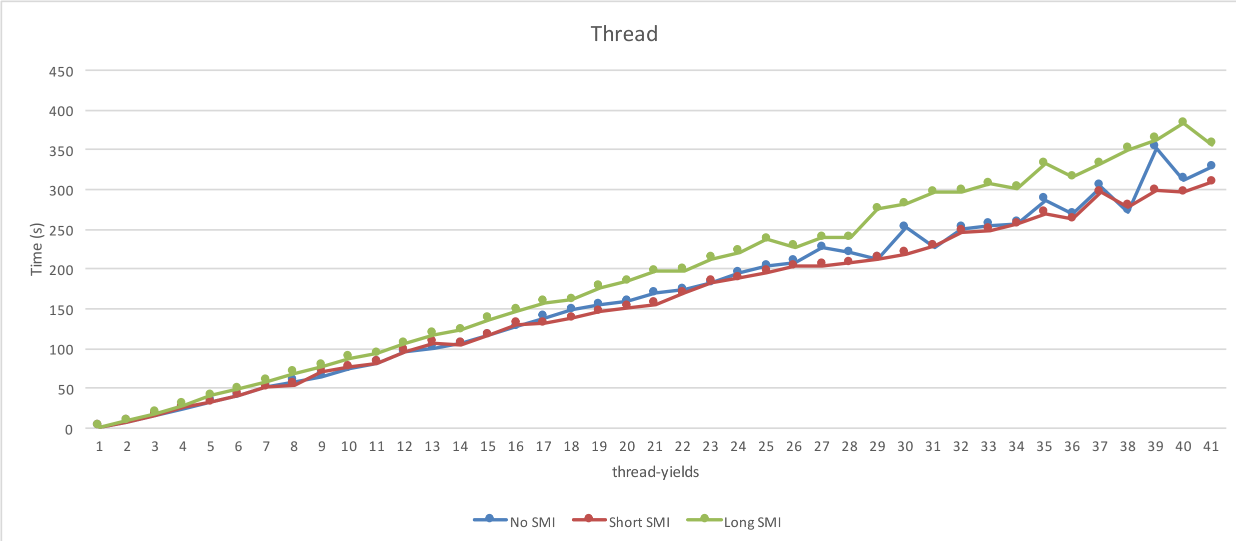

Figure 5: SysBench Threads

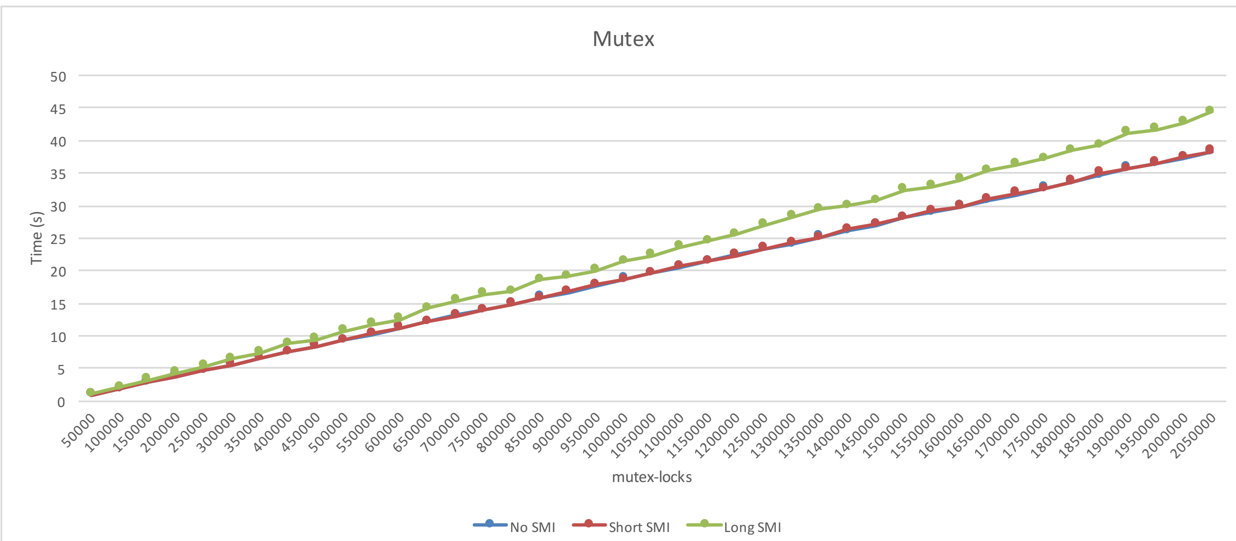

Figure 6: SysBench Mutex

Figures 4 – 6 show what appears to be the characteristic behavior of performance results under the influence of SMIs. Specifically, that there are negligible effects from short SMIs and significant effects from long SMIs as the workload is "turned up".

Summary

We looked at several workloads from a pair of popular benchmarking tools, with specific focus on how performance results are affected when SMIs are enabled. Our results demonstrate that enabling SMIs need not have a significantly negative effect on performance results, provided that short SMIs allow sufficient time for the task.

Appendix: Raw Data

Stress-ng: CPU					% Performance Change	
Num. Stressors	No SMI	Short SMIs	Long SMIs		Short:None	Long:None
1	18197	18201	15653		0.02%	-13.98%
2	35968	35894	31392		-0.21%	-12.72%
3	40914	49923	43791		22.02%	7.03%
4	64716	56559	58988		-12.60%	-8.85%
5	73862	72856	63399		-1.36%	-14.17%
6	79325	77976	66607		-1.70%	-16.03%
7	84271	84165	73686		-0.13%	-12.56%
8	90050	90580	75871		0.59%	-15.75%
				Average	0.83%	-10.88%

Table 1: Stress-ng CPU

Stress-ng: IO					% Performance Change	
Num. Stressors	No SMI	Short SMIs	Long SMIs		Short:None	Long:None
1	97158	34483	30717		-64.51%	-68.38%
2	153203	37254	32609		-75.68%	-78.72%
3	190879	42555	36561		-77.71%	-80.85%
4	215379	45433	39028		-78.91%	-81.88%
5	215176	47566	40526		-77.89%	-81.17%
6	217042	48865	41988		-77.49%	-80.65%
7	221209	49483	43084		-77.63%	-80.52%
8	221209	49427	42531		-77.66%	-80.77%
				Average	-75.93%	-79.12%

Table 2: Stress-ng IO

Stress-ng: VM					% Performance Change	
Num. Stressors	No SMI	Short SMIs	Long SMIs		Short:None	Long:None
1	1592352	1592352	1592352		0.00%	0.00%
2	3184704	3184704	2136128		0.00%	-32.93%
3	4777056	4777056	3204192		0.00%	-32.93%
4	4272256	4272256	4272256		0.00%	0.00%
5	5340320	5340320	5340320		0.00%	0.00%
6	6408384	6408384	6408384		0.00%	0.00%
7	6408384	7476448	7476448		16.67%	16.67%
8	8544512	8544512	8544512		0.00%	0.00%
				Average	2.08%	-6.15%

Table 3: Stress-ng Virtual Memory

SysBench: CPU					% Performance Change	
max-cpu-primes	No SMI	Short SMIs	Long SMIs		Short:None	Long:None
1000	0.5161	0.5163	0.5159		0.04%	-0.04%
3000	2.1456	2.1456	2.3485		0.00%	9.46%
6000	5.3285	5.3283	5.9874		0.00%	12.37%
9000	9.1005	9.1025	10.5774		0.02%	16.23%
12000	13.3487	13.3545	15.4967		0.04%	16.09%
15000	17.9725	17.9743	20.613		0.01%	14.69%
18000	22.9029	22.9105	26.6344		0.03%	16.29%
21000	28.0524	28.0597	32.2117		0.03%	14.83%
24000	33.5912	33.5952	38.6969		0.01%	15.20%
27000	39.3233	39.3357	46.0183		0.03%	17.03%
30000	45.2305	45.228	50.9565		-0.01%	12.66%
33000	51.4122	51.413	59.5447		0.00%	15.82%
36000	57.8009	57.8144	66.6961		0.02%	15.39%
39000	64.3901	64.4205	74.3162		0.05%	15.42%
42000	71.1818	71.2137	82.263		0.04%	15.57%
45000	78.1806	78.1835	90.1358		0.00%	15.29%
48000	85.2069	85.2181	96.8872		0.01%	13.71%
51000	92.4637	92.4745	106.2962		0.01%	14.96%
54000	99.9413	99.9576	113.5789		0.02%	13.65%
57000	107.6651	107.6761	123.3935		0.01%	14.61%
60000	115.456	115.4677	132.2822		0.01%	14.57%
63000	123.2234	123.2454	141.2235		0.02%	14.61%
66000	131.2645	131.2862	149.4796		0.02%	13.88%

69000	139.4171	139.4904	161.2729		0.05%	15.68%
72000	147.904	147.9666	171.0554		0.04%	15.65%
75000	156.3634	156.4449	180.374		0.05%	15.36%
78000	165.0425	165.0594	191.1802		0.01%	15.84%
81000	173.7873	173.8076	203.0747		0.01%	16.85%
84000	182.6635	182.7019	210.8265		0.02%	15.42%
87000	191.6142	191.6381	219.2985		0.01%	14.45%
90000	200.7278	200.7394	230.0381		0.01%	14.60%
93000	210.1316	210.1173	243.676		-0.01%	15.96%
96000	219.5384	219.5677	251.4824		0.01%	14.55%
99000	228.9395	229.0505	264.7467		0.05%	15.64%
102000	238.5288	238.5337	275.4706		0.00%	15.49%
105000	248.1602	248.1686	286.4501		0.00%	15.43%
108000	257.7517	257.7792	298.8337		0.01%	15.94%
111000	267.742	267.7565	305.4793		0.01%	14.09%
114000	277.5825	277.6028	318.4798		0.01%	14.73%
117000	287.7205	287.7463	331.415		0.01%	15.19%
120000	297.8586	297.8728	339.9812		0.00%	14.14%
				Average	0.02%	14.57%

Table 4: SysBench CPU test mode

SysBench: Thread					% Performance Change	
thread-yields	No SMI	Short SMIs	Long SMIs		Short:None	Long:None
1000	1.4074	1.4731	1.7241		4.67%	22.50%
5000	8.286	7.9772	9.2651		-3.73%	11.82%
10000	16.3464	16.3538	18.9079		0.05%	15.67%
15000	24.9294	25.6741	29.5294		2.99%	18.45%
20000	33.2625	32.7307	40.9459		-1.60%	23.10%
25000	40.4556	42.2787	48.8421		4.51%	20.73%
30000	53.0202	51.9565	58.7585		-2.01%	10.82%
35000	58.9725	54.4651	69.3294		-7.64%	17.56%
40000	65.2753	71.6787	77.4955		9.81%	18.72%
45000	74.7406	76.5159	88.8876		2.38%	18.93%
50000	81.9853	81.6647	93.9448		-0.39%	14.59%
55000	95.6127	96.7341	106.714		1.17%	11.61%
60000	100.965	107.8731	118.291		6.84%	17.16%
65000	106.1614	105.3801	123.1633		-0.74%	16.02%
70000	116.462	116.9267	137.2923		0.40%	17.89%
75000	127.7851	130.458	147.3558		2.09%	15.32%
80000	138.8676	131.3328	158.0746		-5.43%	13.83%

85000	148.3392	138.3733	161.6167		-6.72%	8.95%
90000	155.273	146.8471	177.167		-5.43%	14.10%
95000	159.4259	151.3982	184.883		-5.04%	15.97%
100000	169.743	156.2897	197.5901		-7.93%	16.41%
105000	173.9455	169.8797	198.3267		-2.34%	14.02%
110000	181.967	183.8722	213.0773		1.05%	17.10%
115000	194.6631	188.6286	222.005		-3.10%	14.05%
120000	203.4566	196.2944	237.2666		-3.52%	16.62%
125000	209.0953	204.0526	227.863		-2.41%	8.98%
130000	226.6594	204.8225	238.9537		-9.63%	5.42%
135000	220.869	207.5028	239.1368		-6.05%	8.27%
140000	213.3414	213.1351	275.3083		-0.10%	29.05%
145000	252.5211	219.8889	281.3788		-12.92%	11.43%
150000	227.8987	228.423	296.1052		0.23%	29.93%
155000	250.751	247.1787	297.6069		-1.42%	18.69%
160000	254.9821	248.699	306.5868		-2.46%	20.24%
165000	256.9425	256.5161	302.3064		-0.17%	17.66%
170000	287.3172	269.9463	332.3688		-6.05%	15.68%
175000	269.3486	262.7669	315.9288		-2.44%	17.29%
180000	304.0431	296.8207	332.7673		-2.38%	9.45%
185000	272.3436	278.9019	350.71		2.41%	28.77%
190000	352.9544	298.8267	363.5042		-15.34%	2.99%
195000	312.0724	296.1963	383.6368		-5.09%	22.93%
200000	328.2398	308.8409	356.714		-5.91%	8.67%
				Average	-2.18%	16.03%

Table 5: SysBench Threads test mode

SysBench: Mutex					% Performance Change	
mutex-locks	No SMI	Short SMI	Long SMI		Short:None	Long:None
50000	0.9254	0.9376	1.0269		1.32%	10.97%
100000	1.8651	1.8602	2.0646		-0.26%	10.70%
150000	2.8051	2.802	3.2613		-0.11%	16.26%
200000	3.7236	3.7249	4.29		0.03%	15.21%
250000	4.6749	4.6425	5.3527		-0.69%	14.50%
300000	5.5495	5.6083	6.4067		1.06%	15.45%
350000	6.5666	6.5546	7.4017		-0.18%	12.72%
400000	7.4725	7.4285	8.7422		-0.59%	16.99%
450000	8.3904	8.3734	9.4722		-0.20%	12.89%
500000	9.3101	9.3306	10.7633		0.22%	15.61%
550000	10.2563	10.2725	11.7446		0.16%	14.51%

600000	11.2144	11.1835	12.5794		-0.28%	12.17%
650000	12.1577	12.1129	14.1769		-0.37%	16.61%
700000	13.1926	13.097	15.3067		-0.72%	16.02%
750000	13.9347	13.9662	16.4253		0.23%	17.87%
800000	14.8965	14.882	16.7879		-0.10%	12.70%
850000	15.9327	15.7675	18.5629		-1.04%	16.51%
900000	16.6684	16.8216	19.1668		0.92%	14.99%
950000	17.7486	17.8416	20.0687		0.52%	13.07%
1000000	18.7421	18.5513	21.387		-1.02%	14.11%
1050000	19.6232	19.6335	22.3005		0.05%	13.64%
1100000	20.5687	20.6496	23.6829		0.39%	15.14%
1150000	21.4323	21.496	24.5211		0.30%	14.41%
1200000	22.4365	22.3466	25.6257		-0.40%	14.21%
1250000	23.3946	23.3893	27.0019		-0.02%	15.42%
1300000	24.072	24.2954	28.2317		0.93%	17.28%
1350000	25.1502	25.0841	29.4241		-0.26%	16.99%
1400000	26.1117	26.374	29.8954		1.00%	14.49%
1450000	27.0369	27.0879	30.7629		0.19%	13.78%
1500000	28.0801	28.1285	32.3822		0.17%	15.32%
1550000	28.9666	29.1385	32.9077		0.59%	13.61%
1600000	29.8611	29.8683	33.9802		0.02%	13.79%
1650000	30.7619	30.9226	35.415		0.52%	15.13%
1700000	31.6659	31.8771	36.2554		0.67%	14.49%
1750000	32.6161	32.5672	37.1818		-0.15%	14.00%
1800000	33.5704	33.6512	38.5236		0.24%	14.75%
1850000	34.6514	34.9281	39.2705		0.80%	13.33%
1900000	35.6974	35.6304	41.1273		-0.19%	15.21%
1950000	36.4009	36.486	41.7042		0.23%	14.57%
2000000	37.3502	37.4182	42.6987		0.18%	14.32%
2050000	38.1847	38.3677	44.4042		0.48%	16.29%
				Average	0.11%	14.64%

Table 6: SysBench Mutex test mode

References

[1] B. Delgado, K. L. Karavanic. Performance Implications of System Management Mode. In *2013 IEEE International Symposium on Workload Characterization (IISWC'13)*, Portland, OR USA, Sept. 2013.

[2] http://kernel.ubuntu.com/~cking/stress-ng/

[3] Intel, Intel® 64 and IA-32 Architectures Software Developer's Manual (Vol. 3). http://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-software-developer-system-programming-manual-325384.pdf
[4] http://www.rcollins.org/ddj/Jan97/Jan97.html
[5] https://github.com/akopytov/sysbench
[6] http://imysql.com/wp-content/uploads/2014/10/sysbench-manual.pdf
[7] http://manpages.ubuntu.com/manpages/utopic/man1/stress-ng.1.html