

## PERFORMANCE COMPARISON BETWEEN THREE PROCESSING MODES ON INTEL NEHALEM I5-460M

Department of electronic computers, Belarusian State University of Informatics and Radio-electronics  
Minsk, Belarus

Askari M.

Ivanov N. N. Associate Professor

In modern computer systems, one of the main aims is to improve system performance. One of the ways to improve the performance of system is to improve Cycle per Instruction. Due to the increasing popularity of multi-core systems, understanding the impact of various working modes of applications on these processors will allow developers of operating systems to control the status of execution of certain classes of applications on processor and thus help to improve system performance. Applications on a multi-core system can be executed on speed and rate modes, and in the case of rate mode the applications can be run on Hyper-threading and non Hyper-threading modes. In this paper, Intel Nehalem i5-460M has been studied. Results show that in 11 categories of 12 categories from SPEC CPU2000 benchmarks, average of Cycle per Instruction on speed mode is less than non Hyper-threading mode and average of Cycle per Instruction on non Hyper-threading mode is less than Hyper-threading mode. Cycle per Instruction is calculated by Intel-vtune2013 by collecting the values associated with system events. Analysis of the results is performed using ANOVA Multiple Comparison method in SPSS.

Multi-processor systems have a particular popularity among users. They are able to run applications in on speed and rate modes. On speed mode if the compiler used is capable to parallel execution, one copy of a benchmark is executed on multiple cores as distributed and otherwise it uses only one core. In the case non Hyper-threading mode, physical cores will run applications and on Hyper-threading mode physical and virtual cores will run applications.

The system under test with architecture Nehalem is Intel Core i5-460M with 2 cores. Intel Core i5-460M micro-architecture has two cores and three levels of caches, where L1 and L2 are exclusive and L3 is an inclusive cache with respect to L1 and L2.

The L1 cache is divided for instruction and data parts, they are allocated to each core separately, L2 cache is also allocated to each core, instructions and data are stored in L2 cache together. L3 cache is shared between the cores.

TLB design is performed in hardware mode on this processor. It has two levels, the first level of the buffer allocated for each core, and then it is divided for instruction and data. Instruction TLB is divided into two modes: 4Kbyte pages size and 2 (or 4) Megabyte pages size. 4Kbyte mode has 4-way set associative structure and 64 entries line in cache. 4Mbyte mode has fully associative structure with 7 entries line. Data TLB is divided into two modes: 4Kbyte pages size and 2 (or 4) Megabyte pages size. 4Kbyte mode has 4-way set associative structure and 64 entries line in cache. 4Mbyte mode has 4-way set associative structure with 32 entries line. The second level TLB (STLB) allocated for each core separately. If during the execution multi-threading is enabling, the STLB will be shared between two thread of each core. In this case, each core of 460M is capable of executing two threads simultaneously of a total 4 threads. Figure 1 shows a diagram of an Intel i5-460M.

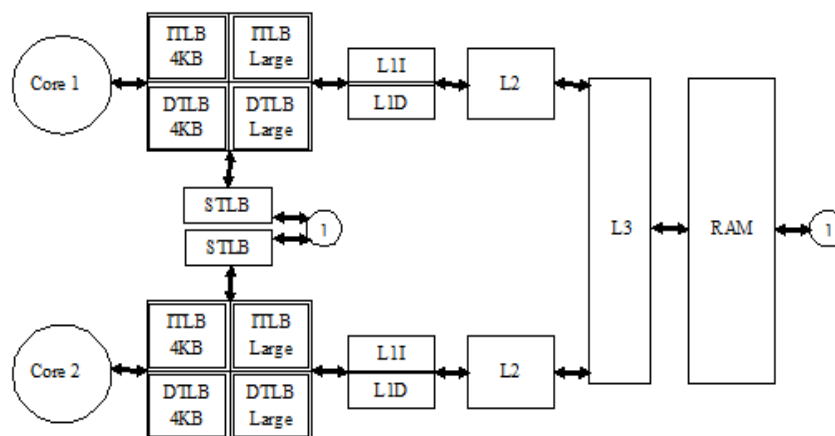


Figure 1—A diagram of memory hierarchy in Intel Nehalem i5-460M

Memory hierarchy is respected in this system. If the address is not found in the first level of TLB, then second level of TLB will be to search and if it is not found again, then memory controller addresses to Random Access Memory (RAM). If the address is not found then there is a miss situation with a complex and time-consuming search for the virtual page that contains the desired address.

The purpose of this study is to compare the Cycle per Instruction (CPI) of this architecture on a three modes including speed, non Hyper-threading and Hyper-threading. For this purpose, all 12 categories of SPEC CINT CPU2000 package have been implemented on these 3 modes.

Experiments performed are based on 64-bit Intel environment which use features of Performance Monitoring Unit (PMU) to measure various events using Intel-Vtune2013. PMU is a hardware part that builds inside a processor in all modern systems to count the performance parameters like instruction cycles, cache hits, cache misses and etc. Each batch was run 50 times, and each run was performed three times randomly. A total of 1800 runs for each case is performed and for three modes in total 5400 times. Events used to calculate CPI are counted by Intel-vtune2013 and are included: CPU\_CLK\_UNHALTED.THREAD that is total cycles of execution for the program under test, and INST\_RETIRED.ANY that is the number of instructions that retired execution.

Intel-vtune amplifier is an application that works on 32 and 64 bit x86 based systems to count the events related to system performance. Windows 7 is the operating system (OS) used on 460M that can run parallel programs on all CPU cores. Programs that have been used to run and to count their events by Intel-vtune2013 are SPEC CPU2000 benchmarks.

To calculate the CPI has been used Equation 1:

$$CPI = CPU\_CLK\_UNHALTED.THREAD / INST\_RETIRED.ANY (1)$$

The results of these experiments are compared using ANOVA Multiple Comparison method in SPSS. In this context, a null hypothesis is defined as assume there is no significant relationship between the measured parameters considering Equation 2:

$$H_0: \text{Average of speed mode} = \text{Average of non Hyperthreading mode} = \text{Average of Hyperthreading mode} (2)$$

The concept of a null hypothesis is used in two approaches: the significance testing approach of Ronald Fisher and the hypothesis testing approach of Jerzy Neyman and Egon Pearson. Both methods are considered the certain error rates. If researcher can bring a strong indication from a statistical standpoint to reject the null hypothesis, then the null hypothesis is rejected. The minimum confidence interval for rejecting the null hypothesis is equal to 95%. In fact, the null hypothesis is rejected if the significant factor is less than an assumption value 0.05. The results of this analysis are given in Table 1. In this table, 2 Core modes are equal by non Hyper-threading mode and 4 Core modes is equal by Hyper-threading mode.

The table 1 includes the 12 benchmarks of CINT CPU2000 of SPEC Corporation. The Standard Performance Evaluation Corporation (SPEC) is a corporation to make a standardized set of relevant benchmarks to use to evaluate system performance [SPEC]. In this paper, all 12 benchmarks of package CINT2000 (fix point operations) are used: 164.gzip, 175.vpr, 176.gcc, 181.mcf, 186.crafty, 197.parser, 252.eon, 253.perlbnk, 254.gap, 255.vortex, 256.bzip2, and 300.twolf. The column titled "Average" is the average of CPIs in each benchmark and the column titled "Std. Dev." is standard deviation of CPIs. The significant F between each two modes of experiments is mentioned in column titled "Significant F.". If this value is less than 0.05 then the null hypothesis is rejected in this case.

Table 1—Result of analysis of the three methods on Intel i5–460M

Benchmark	CPI on Speed mode		CPI on 2 Core mode		CPI on 4 Core mode		Significant F		
	Average	Std. Dev.	Average	Std. Dev.	Average	Std. Dev.	Speed & 2 Core	Speed & 4 Core	2 Core & 4 Core
Eon	1.5060	0.1474	1.3073	0.0536	1.4611	0.1248	3.70e-14	0.1633	1.69e-9
Gzip	0.9043	0.5248	1.0772	0.1003	1.2314	0.0377	1.57e-24	1.12e-51	6.11e-21
Parser	1.0128	0.0812	1.2557	0.0946	1.4561	0.0370	5.91e-34	3.80e-63	1.33e-26
Vortex	0.7766	0.0341	0.9956	0.0347	1.2223	0.0251	4.73e-72	1.16e-114	5.26e-74
Twolf	1.4021	0.1851	1.6927	0.1937	2.5513	0.2555	7.52e-10	4.02e-58	1.70e-43
Gap	0.8176	0.0357	0.9848	0.0244	1.1707	0.0408	6.37e-53	2.57e-95	1.95e-58
Vpr	1.3929	0.2039	1.6518	0.1156	2.2086	0.1245	8.06e-14	1.35e-57	6.38e-39
Mcf	3.1481	0.6948	7.4458	0.4379	10.9857	1.4514	1.38e-48	1.93e-81	1.87e-39
Bzip2	1.0451	0.0688	1.2330	0.0478	1.4267	0.0475	9.32e-36	1.48e-71	4.40e-37
Crafty	0.7176	0.0676	0.9184	0.0259	1.1249	0.0301	2.87e-48	2.39e-87	1.18e-49
Perlbnk	0.6477	0.0422	0.7964	0.0218	0.9694	0.0183	7.85e-55	5.06e-99	7.16e-63
Gcc	1.0020	0.0521	1.0174	0.0506	1.3647	0.0442	0.3609	9.6359	2.93e-73

According to Table 1 it is seen that, in all cases null hypothesis is rejected, except Eon and Gcc. In the benchmark Eon in the case Speed and 4 Core, the null hypothesis is not rejected and it is not a significant relationship between them. Also, in the benchmark Gcc in the case Speed and 2 core and in the case Speed and 4 Cores it is not a significant relationship between them.

In addition, in the all cases average of CPI on speed mode is less than average of CPI on non Hyper-threading and it is less than Hyper-threading mode. There is only one exception in the case of Eon. In this case, average of speed mode is more than Hyper-threading mode and it is more than non Hyper-threading mode.

These data indicate that although the dual-core and quad-core modes, the total execution time can be down due to the parallel execution, but the average of CPI is increased. The authors proposed that this amount be improved by studying on parts that are affecting on CPI like memory hierarchy miss rates.

#### Main references:

1. Hennessy, J. Computer Architecture / J. Hennessy, D. Patterson // Philadelphia: Elsevier, 2011.
2. Akhter, S. Multi-Core Programming/ S. Akhter, J. Roberts // USA, Intel Press, 2006.
3. Reinders, J. Intel VTune Performance Analyzer Essentials/ J. Reinders // USA, Intel Press, 2005.