CPU performance monitoring using the Time-Stamp Counter register

This laboratory work introduces basic information on the Time-Stamp Counter CPU register, which is used for performance monitoring. The focus is on how to use this feature correctly and how to obtain accurate results when measuring the computing performance.

1. Introduction

The Time-Stamp Counter (TSC) is a 64-bit model specific register (MSR) that accurately counts the cycles that occur on the processor. It is present on all x86 processors. The TSC is incremented every clock cycle and is set to zero every time the processor is reset. This register is accessible to the programmer since the Pentium processor.

To access the TSC, the programmer has to call the RDTSC (read time-stamp counter) instruction from assembly language. The RDTSC instruction loads the EDX:EAX with the content of the TSC register. The EDX will contain the high-order 32 bits and the EAX will contain the low-order 32 bits.

The TSC counts the CPU cycles, so the value returned by the RDTSC instruction will be the number of cycles counted from the last processor reset to the point RDTSC was called. To obtain time in seconds, the value provided by the TSC has to be divided with the processor frequency (in Hz) as shown below:

#seconds = #cycles / frequency

This method for performance monitoring is very useful for measuring the cycle count for small sections of code. For example when trying to compare the performance of sections of code that have the same result but use different instructions. Another case in which this method can be of use is to obtain the average execution time for a function or section of code.

2. How to use RDTSC instruction

To obtain accurate results when measuring the performance with RDTSC instruction, the programmer has to be aware of the main issues that affect the cycle count and how to work-around these issues. The main issues that affect cycle count are:

- Out-of-order execution: the order of instruction execution is not as in the source code, this may cause the RDTSC instruction to return a cycle count that is less or greater than the actual cycle count for the measured sequence of code.
- Data cache and instruction cache: if the code or data are not in the cache, the cycle count is much larger.
- Context switches: if they occur during measurement, the result will be biased.
- Frequency changes: results are not accurate if there are frequency changes during measurement.
- Multi-core processors: the cycle counters on the cores are not synchronized. If the process migrates during measurement, the result will be wrong.

The solutions for these issues are discussed as follows.

Basic measurement method from assembly language

To make a basic cycle measurement for an instruction, the programmer has to use the following code from assembly language:

```
rdtsc
mov time_high, edx
mov time_low, eax
add var, ecx
rdtsc
sub eax, time_low
sbb edx, time_high
mov time_high, edx
mov time_low, eax
```

The previous code section measures the cycles for the execution of an ADD instruction between a value found in the memory and the value inside ECX register. It is important to use the 64-bit value provided by the TSC. If we use only the low-order 32 bits, the counter overflows (turns zero) in about 11 seconds. Any code with execution time larger than 11 seconds will not be measured correctly. The ADD instruction can be replaced by any sequence of instructions or any function/procedure.

Out-of-order execution

In case of out-of-order execution the instructions do not execute necessarily in the order indicated by the source code, so the RDTSC instruction can be executed before or after the location indicated in the source code. This causes measurements that are not accurate. To prevent the RDTSC instruction from being executed out-of-order, a serializing instruction has to be executed before it. The CPUID instruction causes all the instructions preceding it to be executed, before its execution. If CPUID instruction is placed before, the RDTSC instruction will be executed in-order. To make an exact measurement, the overhead of CPUID can be measured and subtracted from the cycle count obtained for the measured code. The CPUID instruction takes longer to execute the first two times it is called. It is better to measure the execution of its third call, and use this value for all future measurements.

Measurement method in C/C++

There are some important considerations for measuring the cycles from code written in C/C++ language. Not all compilers recognize the RDTSC and CPUID instructions. The work-around for these cases is to use emit statements that replace the instructions with the corresponding opcodes in the “obj”.

```
#define rdtsc __asm __emit 0fh __asm __emit 031h
#define cpuid __asm __emit 0fh __asm __emit 0a2h
```

If the compiler does not recognize the RTDSC instruction, it will not be aware of it modifying the EDX and EAX registers. For this reason is better to save the general purpose registers before using this instruction. Use the RDTSC instruction from C/C++ as below:

```
unsigned cycles_high1=0, cycles_low1=0, cupid_time=0;
unsigned cycles_high2=0, cycles_low2=0;
unsigned __int64 temp_cycles1=0, temp_cycles2=0;
__int64 total_cycles=0;


//compute the CPUID overhead
__asm {
```

```
   pushad
   CPUID
   RDTSC
   mov cycles_high1, edx
   mov cycles_low1, eax
   popad
   pushad
   CPUID
   RDTSC
   popad

   pushad
   CPUID
   RDTSC
   mov cycles_high1, edx
   mov cycles_low1, eax
   popad
   pushad
   CPUID
   RDTSC
   popad

   pushad
   CPUID
   RDTSC
   mov cycles_high1, edx
   mov cycles_low1, eax
   popad
   pushad
   CPUID
   RDTSC
   sub eax, cycles_low1
   mov cupid_time, eax
   popad
}

cycles_high1=0;
cycles_low1=0;

//Measure the code sequence

__asm {
    pushad
    CPUID
    RDTSC
    mov cycles_high1, edx
    mov cycles_low1, eax
```

```
    popad
}

//Section of code to be measured

__asm {
    pushad
    CPUID
    RDTSC
    mov cycles_high2, edx
    mov cycles_low2, eax
    popad
}
temp_cycles1 = ((unsigned __int64)cycles_high1 << 32) | cycles_low1;
temp_cycles2 = ((unsigned __int64)cycles_high2 << 32) | cycles_low2;
total_cycles = temp_cycles2 - temp_cycles1 - cpuid_time;
```

Data cache and instruction cache

The repeated measurement of the same section of code can produce different results. This can be caused by the cache effects. If the instructions or data are not found in the L1 (instruction or data) cache the cycle count is larger because it takes longer to bring the instructions/data from the main memory. In some cases, the cache effects have to be taken into consideration, for example, when trying to obtain an average cycle count for some section of code. In other cases, the focus is to measure the exact cycle count for some instructions and to obtain a value which is repeatable.

When leaving out the cache interference on cycle count, we have to be sure that the instructions of the sequence we want to measure, and the data we access in that sequence are in the L1 cache. This is done only for small sections of code and small data sets (less than 1KB). The access to the entire data set brings it to the cache (e.g. read every element from the data set). To bring the measured code in the cache, put the code in a function call the function twice and measure only the second call. To be sure that cache effects are taken out, repeat the measurements at least 5 times.

Context switches

Context switches can bias the measurements done with the RDTSC instruction (processes that interrupt during measurement and context switch time will be accounted for). The probability of the measured code to be interrupted by other processes increases with its length. In case of small sequences of code with short execution time, the probability is very small. For large sequences of code, set the process priority to the highest priority. In this way other processes will not interrupt during measurement.

Multi-core processors

The TSCs on the processor's cores are not synchronized. So it is not sure that if a process migrates during execution from one core to another, the measurement will not be affected. To avoid this problem, the measured process's affinity has to be set to just one core, to prevent process migration.

Frequency changes

If the processors support technologies that make possible to change the frequency while functioning (for power management), the execution time measurement is affected. The solution for this problem is to prevent

frequency changes through the settings in the operating system. If the measurement is made in clock cycles, not in seconds, the measurement will be independent of the processor frequency.

3. Exercises

1. Using the RDTSC instruction, measure the following instructions (find repeatable values):
    a. ADD (both values in local registers);
    b. ADD (value in variable);
    c. MUL
    d. FDIV
    e. FSUB
2. Write a function that sorts an array of unsigned integer values. The array is declared first static, then dynamic (two versions).
    a. Measure the average execution time in cycles and seconds (for static/dynamic array) for the sort function.
    b. Measure the exact execution time for the sort function. Find a value that is repeatable.
    c. Which part/instructions of the sort function takes longer to execute? Try to optimize your sort function.

Bibliography

1. Intel, "Using the RDTSC Instruction for Performance Monitoring", http://www.ccsl.carleton.ca/~jamuir/rdtscpm1.pdf
2. Peter Kankowski, "Performance measurements with RDTSC", http://www.strchr.com/performance_measurements_with_rdtsc
3. John Lyon-Smith, "Getting accurate per thread timing on Windows", http://web.archive.org/web/20090510073435/http://lyon-smith.org/blogs/code-o-rama/archive/2007/07/17/timing-code-on-windows-with-the-rdtsc-instruction.aspx