Microarchitecture Analysis of Profile 1
eSTREAM Ciphers on Intel Core2 Duo
Naiyi Hu
Department of Electrical and Computer Engineering
University of New Haven

Abstraction

In this study, we presented a detailed performance analysis of a set of stream ciphers on an Intel Core 2 microprocessor based personal computer. The performance data was collected with Intel VTune Amplifier XE 2011 to reflect CPI, cache hit rate, branch misprediction rate and exploited instruction-level parallelism (ILP). It was observed that all tested workloads yielded a CPI value from 0.44 to 0.68. We noticed a high correlation between the exploited ILP and the CPI value. This indicated that code performance could greatly benefit from the ILP on a superscaler CPU. The stream ciphers used in this paper does not have large amounts of branches. However, performance data showed that even a small branch misprediction rate could affect the CPI to some extent. It was also observed a fairly high cache hit rate (greater than 99%) for all tested workloads, so they were considered having little impact on CPI value in this case.

Experimental Environment

Testing Platform

The test was performed on a dual-core CPU computer system. The testing platform configuration is listed in table 1.

<table>
<thead>
<tr>
<th>Processor</th>
<th>Intel Core 2 Duo CPU E6500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock Frequency</td>
<td>2.8 GHz</td>
</tr>
<tr>
<td>Number of cores</td>
<td>2</td>
</tr>
<tr>
<td>L1 Data cache</td>
<td>32K bytes, 8-way set associative, 64-byte line size</td>
</tr>
<tr>
<td>L1 Instruction cache</td>
<td>32K bytes, 8-way set associative, 64-byte line size</td>
</tr>
<tr>
<td>L2 Cache</td>
<td>2MB, 8-way set associative, 64-byte line size</td>
</tr>
<tr>
<td>Memory Type</td>
<td>DDR2</td>
</tr>
<tr>
<td>Memory Frequency</td>
<td>533 MHz</td>
</tr>
</tbody>
</table>
| Operating System | Microsoft Windows XP Professional 
                  | Veronica 2004, SP2 |

Table 1: Testing platform configuration used in this study

Workloads

The stream ciphers used in this study were chosen from the eSTREAM portfolio software. They were HC-128[1], Rabbit[2], Salsa20[3] and SODENAK[4]. The workloads were compiled with Microsoft C/C++ compiler version 15 using O2 optimization level.

Analysis Tools

Intel VTune Amplifier XE 2011 was used to collect performance data in this study. It provides event-based sampling, which has only little impact on the workloads but can accurately reflect the actual software performance.

General Analysis

Number of Retired Instructions

The retired instruction refers to those instructions that are executed completely and update the machine state. This does not include those instructions that are partially executed and discarded due to branch misprediction (see branch misprediction section for wasted work). Fig. 7 shows the number of retired instructions of the four stream ciphers used in this study.

Average Cycles Per Instruction (CPI)

CPI is one of the most important metrics that reflect overall software performance. Fig. 2 shows the CPI values of the workloads.

Cache Hit Rate

The miss in cache can result in long latency operations in pipeline. Especially when there are other operations depending on the results of such operations, significant stalls could occur. Fig. 3 shows the cache hit rate of L1 and L2 cache.

Branch Misprediction

Branch prediction could significantly lower the run time performance of software. In general, the penalty of a branch misprediction is decomposed into cycle stalls and wasted work.

Instruction-level Parallelism

The software performance can be greatly improved on a superscalar CPU by exploiting the parallelism within it. In this study, the ILP is defined as the average number of micro-op being dispatched for execution per cycle.

The processor used in this study has 6 dispatch ports, so theoretically it can reach an ILP of 6. However, in practical, the amount of exploited parallelism depends greatly on the structure of the software. Data dependency could significantly limit the ILP for the ILP we defined above, control dependency does not lower the ILP, but the work done in parallel will be counted as wasted work if the branch is mispredicted. Fig. 4 shows the ILP calculated for the four stream ciphers in this study.

Analysis of the Correlation with CPI

In Table 3, all performance metrics are put together in order to find out the correlation with CPI.

Table 3: Correlation between CPI and other metrics

Following conclusions can be observed from the above table:

1. In this study, the amount of exploited ILP shows high negative correlation with CPI.
2. The branch misprediction rate seems to have the strongest correlation with CPI. SODENAK and Salsa20 are similar in cache hit rate and ILP. However, they have a difference of as large as 0.082 in CPI and it is most likely attributed to the disparity (0.8%) in branch misprediction rate. Table 2 also shows an indirect proof to this conclusion. A large amount (18%) of wasted work was done due to the 0.8% misprediction.
3. All the four stream ciphers seem to have a small working set, so they can all benefit a lot from the cache. Cache hit rate has little impacts to CPI in this case.

References


Acknowledgement

Thanks Dr. Christopher Martinez for his advice and direction in this study.

Fig. 1: Number of retired instructions of the workloads

As we can see from the chart:
• Number of retired instructions varied from 3.8 billion to 22 billion.
• The number of instructions varied due to structure of cipher.
• Salaa20 executed the most number of retired instructions. It uses a large amount of rotations to eliminate the correlation between the secret key and the output[5].
• HC-128 uses two large S-boxes[6] so it can achieve the required security level without much branches.

Fig. 2: CPI values of the workloads

It was observed that:
• CPI varied between 0.45 and 0.67.
• All CPI values were under 1 as expected on a superscalar CPU.
• Disparity in CPI is due to differences in processor resource efficiency (i.e. Cache, Branch, ILP).

Fig. 3: L1 and L2 cache hit rates of the workloads

We can conclude from the above chart:
• The VStream Keplrer did not capture any L1 instruction cache miss events during the test. This indicates that all stream ciphers have less than 20% of cache size for L1 instruction cache.
• All the four stream ciphers experienced a high L1 data and L2 cache hit rate of over 99%, which indicates that they all have small working sets.
• The cache hit rates of the workloads were very close, so cache hit rate was not the major factor that caused the disparity in CPI values in this case.

Branch Misprediction

Branch cycle stalls happen from the time when misprediction is detected till the branch and all older micro-ops are retired. The cycle stall is represented as a ratio of cycle stall time over total execution time.

The wasted work refers to the number of cycles wasted in executing the wrong branch path. All the micro-ops that were dispatched and executed must be flushed from the CPI.

The branch misprediction rate (misprediction per instruction) of the workloads is shown in Table 2.

Table 2: Branch misprediction rate of the workloads

Following conclusions can be observed from the above table:
• All workloads experienced a low branch misprediction rate. SODENAK had the highest misprediction rate of 0.8%.
• Even low misprediction rate could cause great penalty. SODENAK had a cycle stall ratio of 3% and it approximates 18% wasted work due to 0.8% branch misprediction. HC-128 had a cycle stall ratio of 2% due to 0.01% branch misprediction.

Fig. 4: The exploited ILP in the workloads

It was observed that the exploited ILP in Salaa20 and SODENAK was higher. This indicates that less data dependencies was involved in those two ciphers. In fact, the high parallelism in Salaa20 is obvious. The Salaa20 cipher uses integer counter modulo (ICM) to generate keystream instead of using output feedback mode (OFB). So each internal stage does not depend on previous stages and can be computed independently. The other three ciphers all use OFB mode.