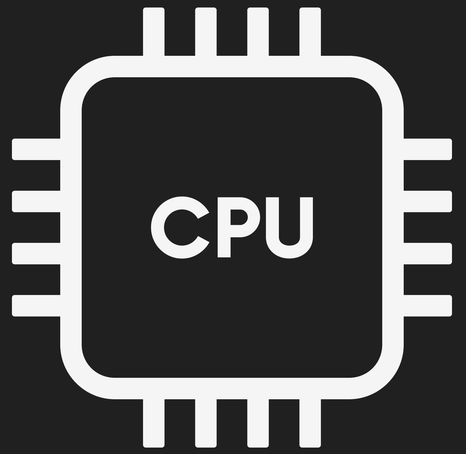
The Spectre Attack Explained In One Minute-ish With Minimal Code

Benjamin Stokman

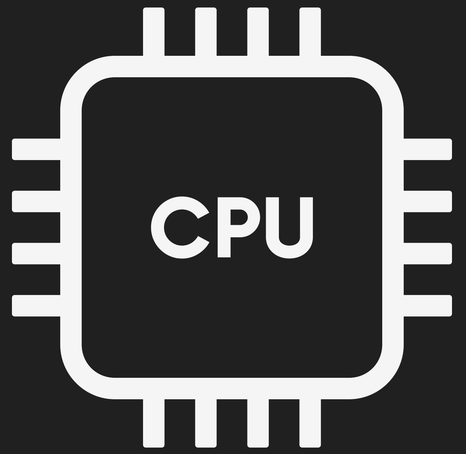
Goal & Design of Attack

- Attacker wants to read memory from another program
- CPU won't let it do that!
- Must trick CPU into giving up data
- Attacker program uses cache to leak data

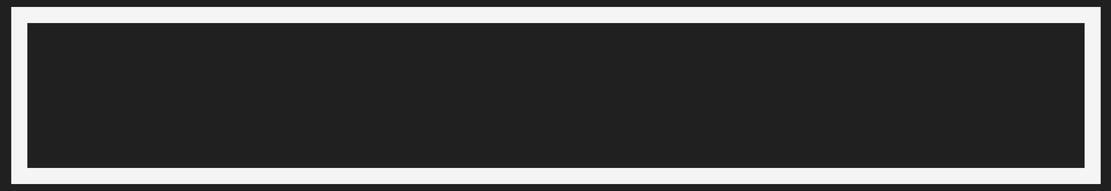
Example



Example



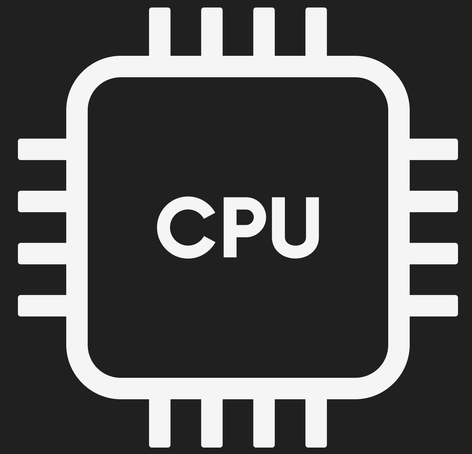
1. Attacker program creates an array in memory with 256 cache block sized items
2. Attacker program makes sure this array is cleared from the cache



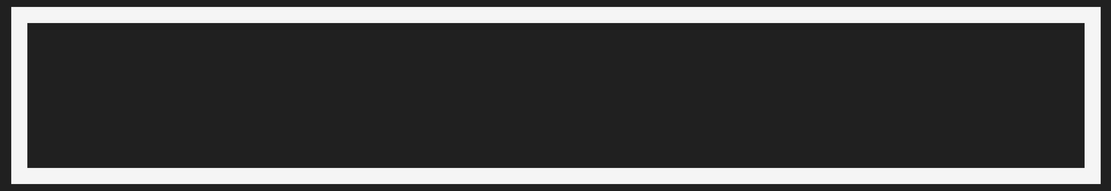
array of 256 cache-sized blocks

Example

I want to read
0x 5000001!



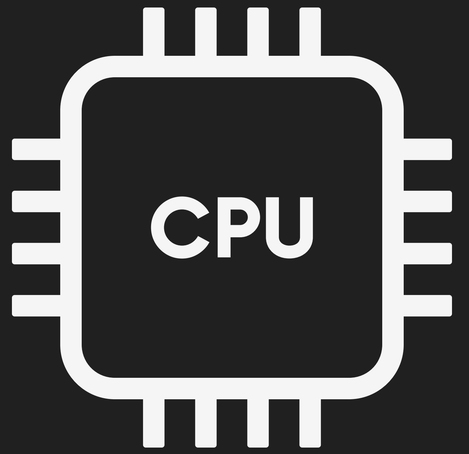
3. Attacker program chooses what memory address to read from



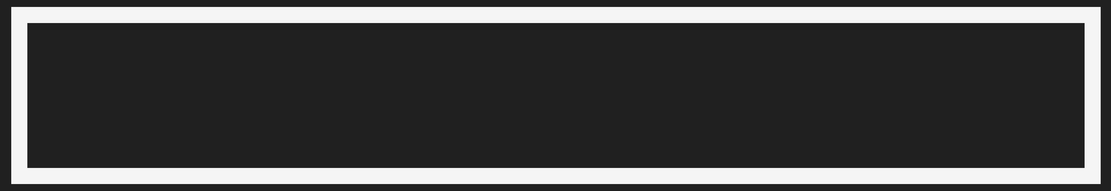
array of 256 cache-sized blocks

Example

secret = [0x5000001]
array[secret]



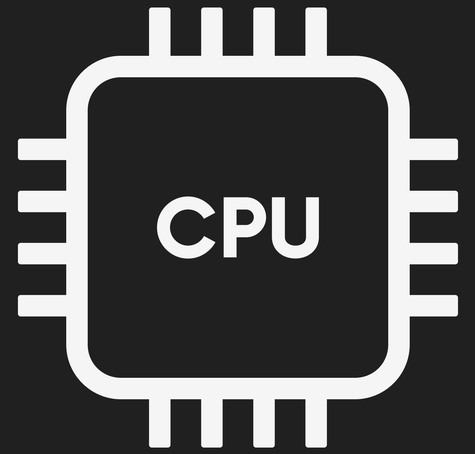
4. Attacker program will tell the CPU to read a value from the array coresponding to what the memory address contains



array of 256 cache-sized blocks

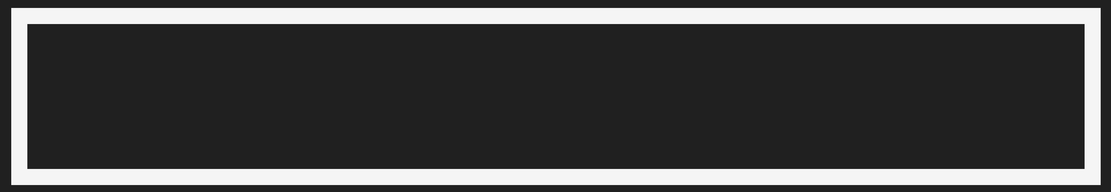
Example

secret = [0x5000001]
array[secret]



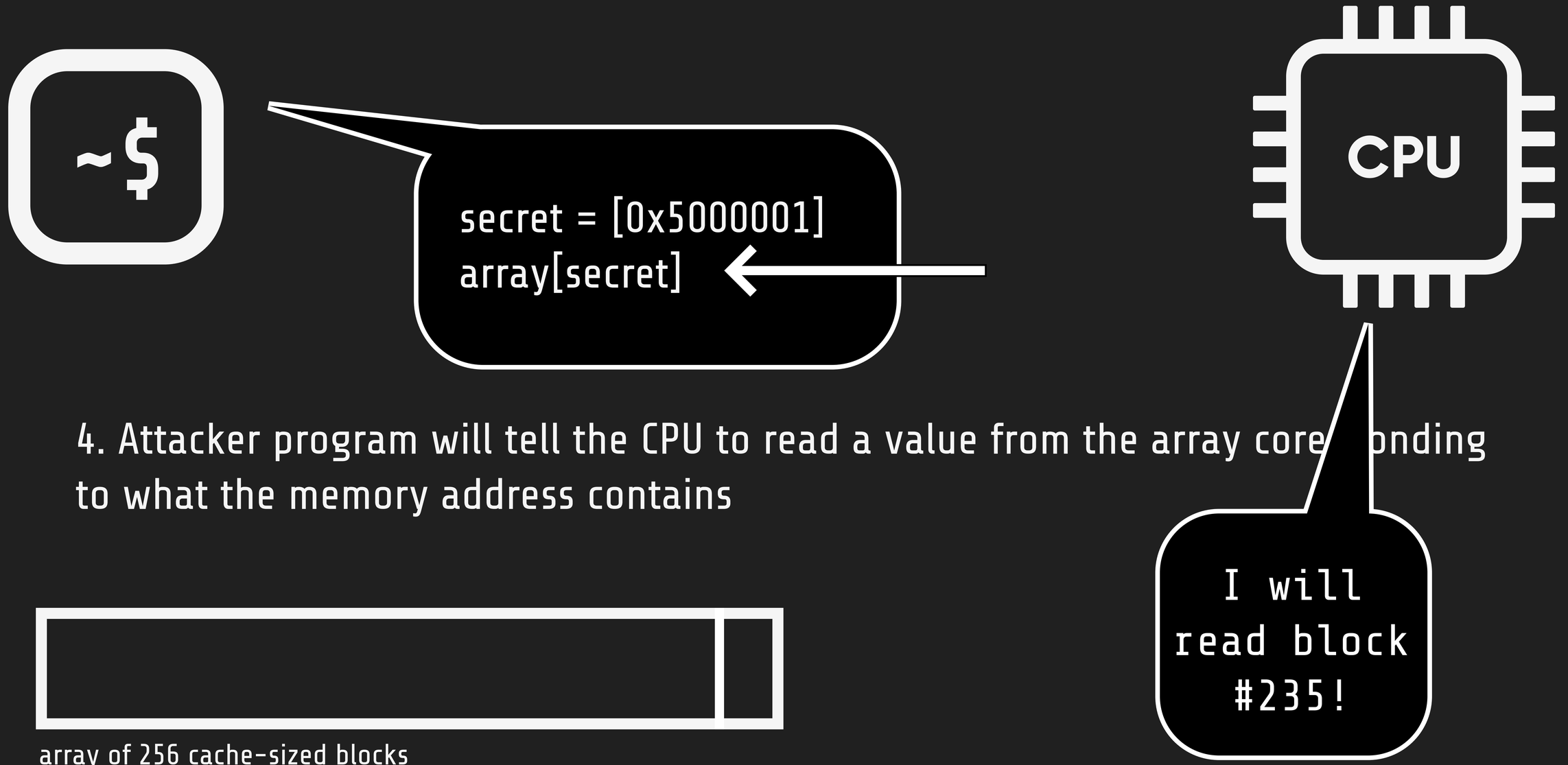
4. Attacker program will tell the CPU to read a value from the array corresponding to what the memory address contains

Value at that address is 235!



array of 256 cache-sized blocks

Example



4. Attacker program will tell the CPU to read a value from the array corresponding to what the memory address contains

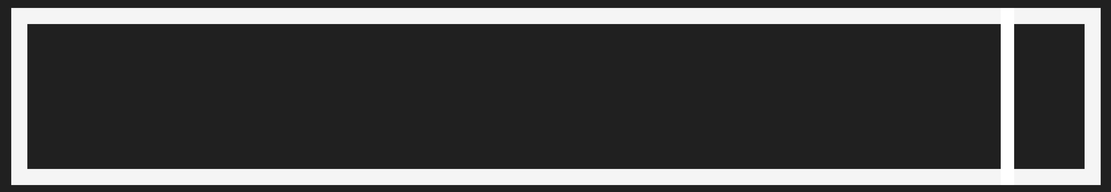
array of 256 cache-sized blocks

Example

4. Attacker program sees which block was loaded into cache

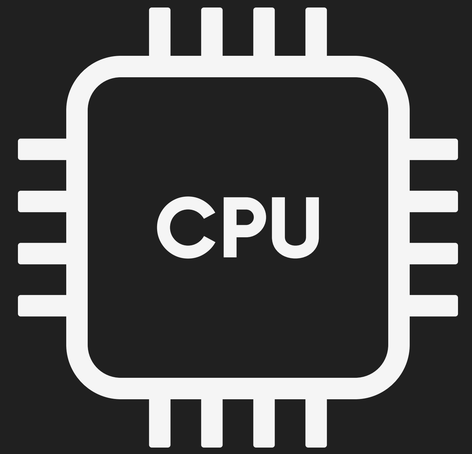
Blocks in cache take about 1/1000th of the time to read than blocks in memory

The CPU does not evict the block from the cache!



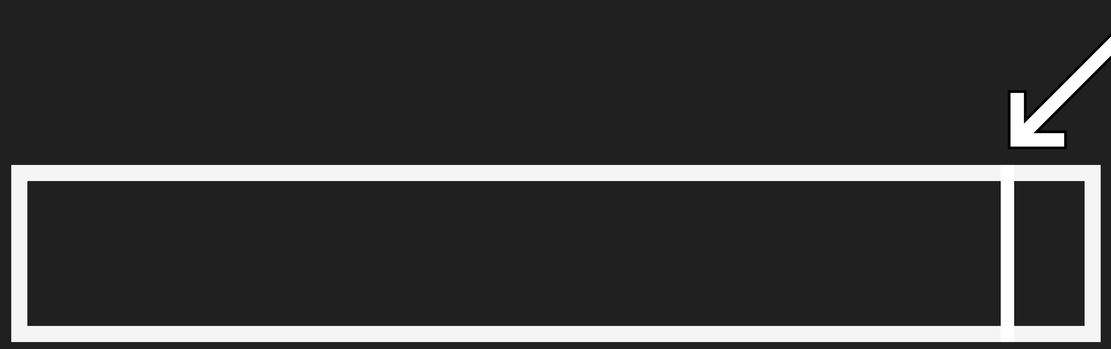
array of 256 cache-sized blocks

Example



4. Attacker program sees which block was loaded into cache

Blocks in cache take about 1/1000th of the time to read than blocks in memory



array of 256 cache-sized blocks

Other Notes

The attacker program uses a high precision timer to measure the read delay.

The attacker program can repeat this process indefinitely.

The code which tricks the CPU into loading the value the program isn't allowed to read is able to run due to branch prediction, which is a mechanism in the CPU to guess if conditionals run or not.