Breaking and Fixing Speculative Load Hardening

Zhiyuan Zhang University of Adelaide Supervisors: Yuval Yarom and Chitchanok Chuengsatiansup

Zhiyuan Zhang, Gilles Barthe, Chitchanok Chuengsatiansup, Peter Schwabe, Yuval Yarom, "Breaking and Fixing Speculative Load Hardening", eprint <u>https://eprint.iacr.org/2022/715</u>.

Zhiyuan Zhang, Gilles Barthe, Chitchanok Chuengsatiansup, Peter Schwabe, Yuval Yarom, "Breaking and Fixing Speculative Load Hardening", eprint <u>https://eprint.iacr.org/2022/715</u>.





Semantic Proof

Zhiyuan Zhang, Gilles Barthe, Chitchanok Chuengsatiansup, Peter Schwabe, Yuval Yarom, "Breaking and Fixing Speculative Load Hardening", eprint <u>https://eprint.iacr.org/2022/715</u>.



Zhiyuan Zhang, Gilles Barthe, Chitchanok Chuengsatiansup, Peter Schwabe, Yuval Yarom, "Breaking and Fixing Speculative Load Hardening", eprint <u>https://eprint.iacr.org/2022/715</u>.

- Speculative Execution
- Spectre Attack and Speculative Load Hardening (SLH)
- Break SLH via control flow transfer and variant-time executions
- Fix SLH \rightarrow Ultimate SLH and performance evaluation

```
Speculative Execution
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

Resolve condition takes time

- Compute condition
- Fetch value from main memory

Resolve condition takes time

- Compute condition
- Fetch value from main memory



Let's predict the branch condition











```
Speculative Execution

if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}

After a while, the condition</pre>
Check branch history
Condition used to be true
The condition may also be true this time
```

is resolved to be false



Let's squash these executions, start executing on the correct path Exit

Speculative Execution Check branch history if (index < arrayLen) {</pre> Condition used to be true x = array[index]; y = array2[x * 4096];The condition may also be true this time } d After a while, the condition is resolved to be false Exit Let's squash these executions, Don't want misprediction next time, update start executing on the correct branch history

path

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

I want to leak entire virtual space memory. It would be good if I can execute the code with an out-of-bound index

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```



I want to leak entire virtual space memory. It would be good if I can execute the code with an out-of-bound index

The speculative execution is based on branch history. I think I could poison the branch history...

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

Keep feeding in-bound index



```
Spectre Attack
```

```
if (index < arrayLen) {</pre>
  x = array[index];
  y = array2[x * 4096];
}
```

```
ÂĊ
```

Keep feeding in-bound index



```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

Keep feeding	in-bound index
Flush arrayLer	1
Flush arrav2	

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

/	$\overline{}$

Кеер	feeding	in-bound	index

Flush arrayLen

Flush array2

Feed an out-of-bound index

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```

Keep feeding in-bound index
Flush arrayLen
Flush array2
Feed an out-of-bound index



According to branch history, I predict index is in-bound

```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```



```
Spectre Attack
```

```
if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}</pre>
```



It's a misprediction. Not a big deal. I rollback all executions.

Keep feeding in-bound index Flush arrayLen Flush array2 Feed an out-of-bound index According to branch history, I predict index is in-bound Read array[index] and access array2

26

```
Spectre Attack

if (index < arrayLen) {</td>

x = array[index];

y = array2[x * 4096];

}

It's a misprediction. Not a big deal.

I rollback all executions.

Keep feeding in-bound index

Flush arrayLen

Flush array2

Feed an out-of-bound index
```

The cache status has been changed



predict index is in-bound

Read array[index] and access array2

```
Spectre Attack

if (index < arrayLen) {
    x = array[index];
    y = array2[x * 4096];
}
</pre>
Keep feeding in-bound index

Flush arrayLen

Flush array2

Feed an out-of-bound index
```



It's a misprediction. Not a big deal. I rollback all executions.





You leak architecture changes

- Cache status
- Execution port contention
- Power consumptions
-



Read array[index] and access array2





Implemented in LLVM









Implemented in LLVM

Track speculative state

Poison loaded value / address



if (index < arrayLen) {
 x = array[index];
 y = array2[x * 4096];
}</pre>

Track speculative state

Poison loaded value / address



Track speculative state

Poison loaded value / address

mask = 0;

if (index < arrayLen) {</pre>

$$y = array2[x * 4096];$$



Track speculative state

Poison loaded value / address

mask = 0; if (index < arrayLen) { mask = index < arrayLen ? mask : -1; x = array[index]; y = array2[x * 4096];



Track speculative state

Poison loaded value / address

mask = 0; if (index < arrayLen) { mask = index < arrayLen ? mask : -1; x = array[index]; y = array2[x * 4096];

Conditional update cannot be speculated



Track speculative state

Poison loaded value / address

mask = 0; if (index < arrayLen) { mask = index < arrayLen ? mask : -1; x = array[index] | mask; y = array2[x * 4096];

Conditional update cannot be speculated
Speculative Load Hardening (SLH)



Track speculative state

Poison loaded value / address

mask = 0; if (index < arrayLen) { mask = index < arrayLen ? mask : -1; x = array[index] | mask; y = array2[x * 4096];

Conditional update cannot be speculated

Poison the value with -1 to avoid race condition between flushing the pipeline and loading memory



• SLH hardens memory reading



• SLH only hardens memory reading



- SLH only hardens memory reading
- Leaks could from
 - Control flow transfer
 - Speculative store
 - Limited execution resources
 - Execution Ports
 - Reservation Station

•



- SLH only hardens memory reading
- Leaks could from
 - Control flow transfer
 - Speculative store
 - Limited execution resources
 - Execution Ports
 - Reservation Station
 - •

• We have interest in crypto code

- We have interest in crypto code
- Crypto code in constant-time
 - No secret-relevant memory access
 - No control flow transfer based on secret
 - No variant-time executions

- We have interest in crypto code
- Crypto code in constant-time
 - No secret-relevant memory access
 - No control flow transfer based on secret
 - No variant-time executions
- However, constant-time computing is not efficient

- We have interest in crypto code
- Crypto code in constant time
 - No secret-relevant memory access
 - No control flow transfer based on secret
 - No variant-time executions
- However, constant-time computing is not efficient
- Only use constant-time computing on secret value

- We have interest in crypto code
- Crypto code in constant time
 - No secret-relevant memory access
 - No control flow transfer based on secret
 - No variant-time executions
- However, constant-time computing is not efficient
- Only use constant-time computing on secret value

```
victim(int value, int isPublic) {
    if (isPublic) {
        //Leaky code
    }
}
```

```
// Boundary Check
if (isPublic) {
    if (value == 0) {
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```

}





Flush the outer branch

• Brings > 150 cycles speculation window

```
// Boundary Check
if (isPublic) {
    if (value == 0) {
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```



Flush the outer branch

• Brings > 150 cycles speculation window

The inner branch is resolved much faster

```
// Boundary Check
if (isPublic) {
    if (value == 0) {
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```



Flush the outer branch

• Brings > 150 cycles speculation window

The inner branch is resolved much faster



Instructions are sent to various execution ports I measure which port is being used.

```
// Boundary Check
if (isPublic) {
  if (value == 0) {
    a2 = a1 | a2;
    a3 = a2 | a3;
    . . .
  } else {
    a1 = crc32(a1, a1);
    a2 = crc32(a2, a2);
    . . .
  }
```



```
// Boundary Check
if (isPublic) {
    if (value == 0) {
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```



```
// Boundary Check
if (isPublic) {
  if (value == 0) {
    a2 = a1 | a2;
    a3 = a2 | a3;
    . . .
  } else {
    a1 = crc32(a1, a1);
    a2 = crc32(a2, a2);
    . . .
  }
```



// Boundary Check
if (isPublic) {
 if (value == 0) {
 a2 = a1 | a2;
 a3 = a2 | a3;
 ...
 } else {
 a1 = crc32(a1, a1);
 a2 = crc32(a2, a2);
 ...
 }





```
// Boundary Check
if (isPublic) {
    if (value == 0) {
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```

```
// Boundary Check
mask = 0;
if (isPublic) {
    mask = isPublic ? mask : -1;
    if (value == 0) {
        mask = value == 0 ? mask : -1;
        a2 = a1 | a2;
        a3 = a2 | a3;
        ...
    } else {
        a1 = crc32(a1, a1);
        a2 = crc32(a2, a2);
        ...
    }
}
```





```
// Boundary Check
mask = 0;
isPublic |= mask;
if (isPublic) {
  mask = isPublic ? mask : -1;
  value |= mask;
  if (value == 0) {
    mask = value == 0 ? mask : -1;
    a2 = a1 | a2;
    a3 = a2 | a3;
    . . .
  } else {
    a1 = crc32(a1, a1);
    a2 = crc32(a2, a2);
    . . .
  }
```



```
// Boundary Check
mask = 0;
isPublic |= mask
if (isPublic) { 🗸
  mask = isPublic ? mask : -1;
  value |= mask;
 if (value == 0) {
   mask = value == 0 ? mask : -1;
   a2 = a1 | a2;
   a3 = a2 | a3;
    . . .
  } else {
   a1 = crc32(a1, a1);
   a2 = crc32(a2, a2);
    . . .
```





Mitigation Result



• DIV Instruction

- DIV Instruction
- REPEAT instruction (REP MOV RAX, RBX)
 - The number of iteration depends on ECX

- DIV Instruction
- REPEAT instruction (REP MOV RAX, RBX)
 - The number of iteration depends on ECX
- Floating point instructions

- DIV Instruction
- REPEAT instruction (REP MOV RAX, RBX)
 - The number of iteration depends on ECX
- Floating point instructions

value = sqrtsd(value); value = mulsd(value, value);



value = sqrtsd(value); value = mulsd(value, value);







value = sqrtsd(value);

value = mulsd(value, value);

On i7-10710U:

Executing a pair of SQRTSD and MULSD:

- 65536: 5 cycles
- 2.34e-308: 7 cycles







Okay, constant-time computing is slow. I use non-constant-time computing for public data



Okay, constant-time computing is slow. I use non-constant-time computing for public data

// Boundary Check

```
if (isPublic) {
   value = sqrtsd(value);
   value = mulsd(value, value);
}
```
Variant-time Instructions

// Boundary Check

- if (isPublic) {
 - value = sqrtsd(value);
 - value = mulsd(value, value);



Okay, constant-time computing is slow. I use non-constant-time computing for public data



The code can be speculatively executed. Measure the execution time.



}

Variant-time Instructions Boundary Check if (isPublic) { value = sqrtsd(value); value = mulsd(value, value); } I wait until the branch is

resolved

Okay, constant-time computing is slow. I use non-constant-time computing for public data

The code can be speculatively executed. Measure the execution time.

Variant-time Instructions Boundary Check if (isPublic) { value = sqrtsd(value); value = mulsd(value, value); } Constant-time under speculative execution. No spectre attacks! l wait until the branch is resolved

Okay, constant-time computing is slow. I use non-constant-time computing for public data

The code can be speculatively executed. Measure the execution time.

Variant-time Instructions Okay, constant-time computing is slow. I use non-constant-time computing for public data Boundary Check if (isPublic) { The code can be speculatively executed. value = sqrtsd(value); Measure the execution time. value = mulsd(value, value); } Constant-time under speculative execution. No spectre attacks! wait until the branch is resolved I doubt it.

76

Review pipeline stages

How is an instruction handled by the processer



Front-end



Review pipeline stages

How is an instruction handled by the processer









Review pipeline stages

























Execution Engine

































Limitation of resources



Limitation of resources



```
// Boundary Check
if (isPublic) {
  value = sqrtsd(value);
  value = mulsd(value, value);
  ...
  value = sqrtsd(value);
  value = mulsd(value, value);
  memory_access(adrs);
}
```

}

```
mask = 0;
// Boundary Check
if (isPublic) {
  mask = isPublic ? mask : -1;
  value = sqrtsd(value);
  value = mulsd(value, value);
  ...
  value = sqrtsd(value);
  value = mulsd(value, value);
  memory_access(adrs);
}
```

Compiled by SLH



```
mask = 0;
// Boundary Check
if (isPublic) {
  mask = isPublic ? mask : -1;
  value |= mask;
  value = sqrtsd(value);
  value = mulsd(value, value);
  ...
  value = sqrtsd(value, value);
  value = mulsd(value, value);
  memory_access(adrs);
}
```

Compiled by USLH





Reservation Station (RS)

Performance



Benchmark with SPEC2017
Summary

- SLH is good, but it is not perfect
- Variable-time executions also leak secret in speculative execution
- Accessing a secret-independent memory may also leak information
- USLH costs more, but it is still better than disabling speculative execution