instaclustr

The impact of Spectre and Meltdown

Ben Bromhead, CTO, Instaclustr

January 2018

Agenda

instaclustr

- An introduction to Speculative Execution, Spectre and Meltdown
- Implemented fixes
- Some graphs on the impact

Intro to Spectre and Meltdown

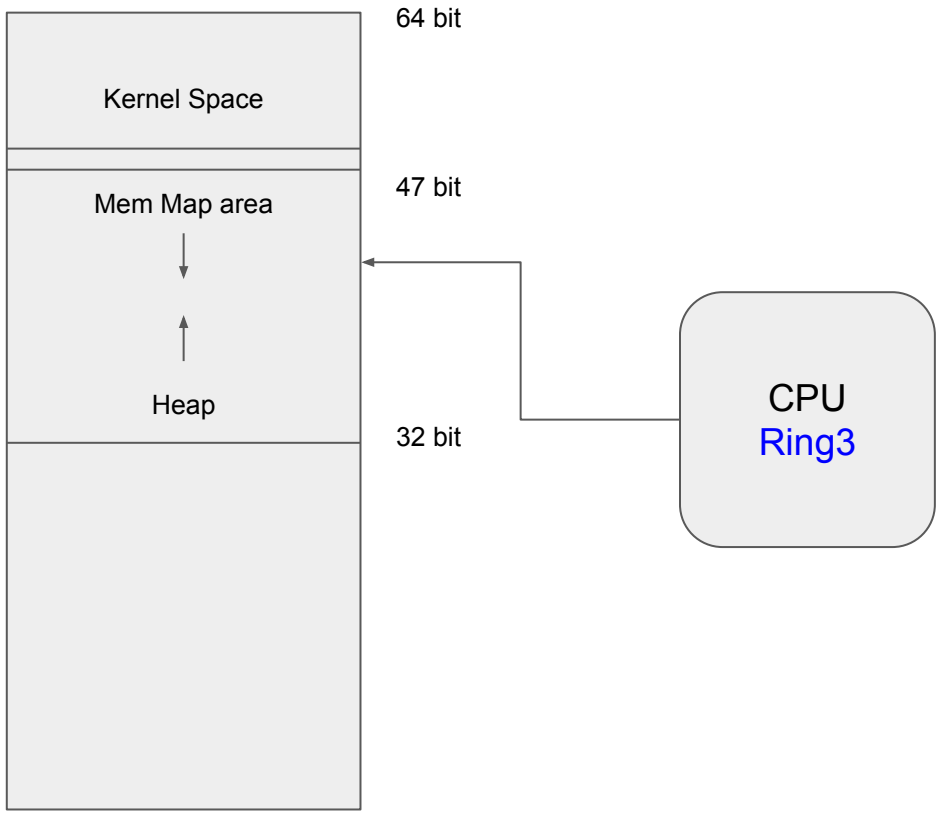
instaclustr

Intro to Spectre and Meltdown

instaclustr

- Three exploits
  - Target information leakage in speculative execution by CPUs.
  - One is vendor specific (Intel).
  - The other two impact most CPUs that implement speculative execution.
- Despite existing for 20+ years, 10+ years of theorised existence, bugs were discovered independently by multiple teams within a 6 month window.
- Fundamentally the concepts exploit three related elements of CPU and OS design:
  - Kernel Memory Mapping
  - CPU L1, L2 and L3 Cache behaviour
  - CPU behavior to minimise stall time (Speculative execution and branch prediction).

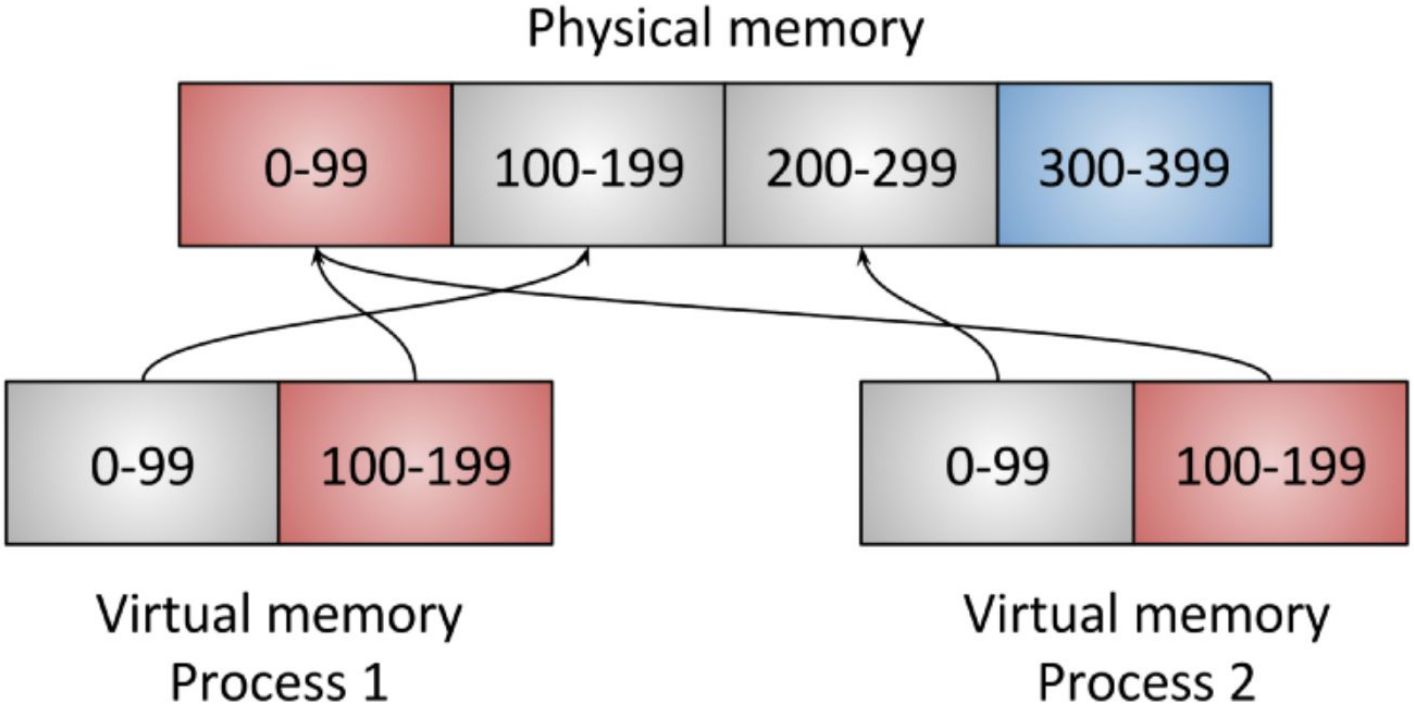
Memory Protection



Virtual Memory space (not including ASLR offsets)

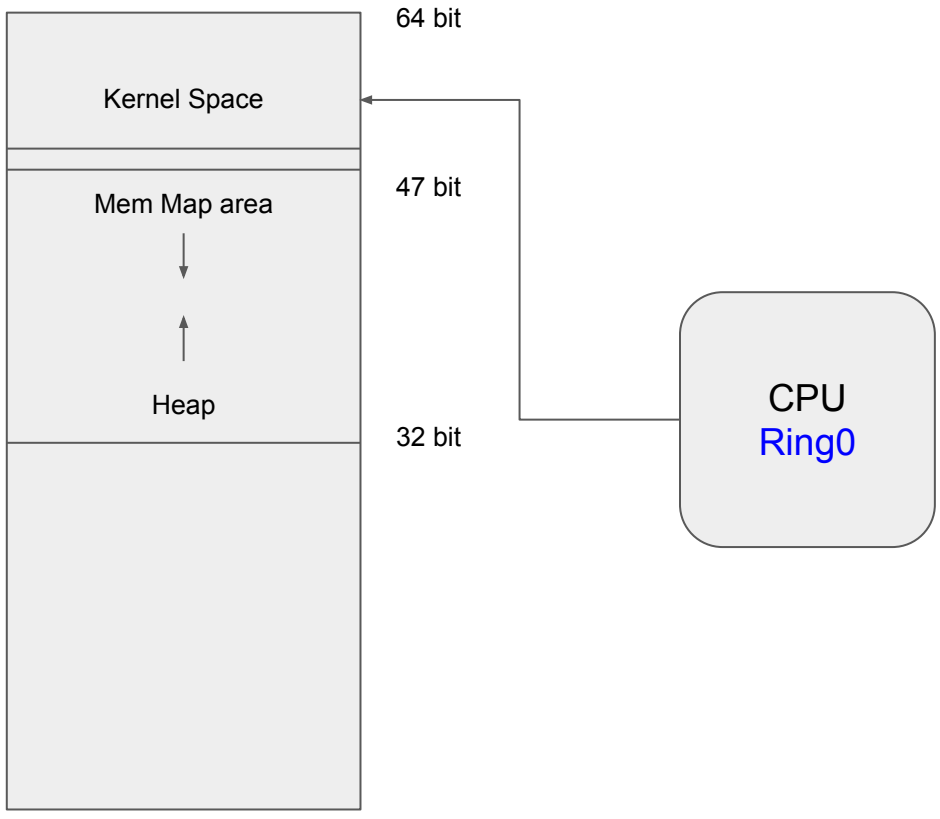
Memory Protection

instaclustr



Memory Protection

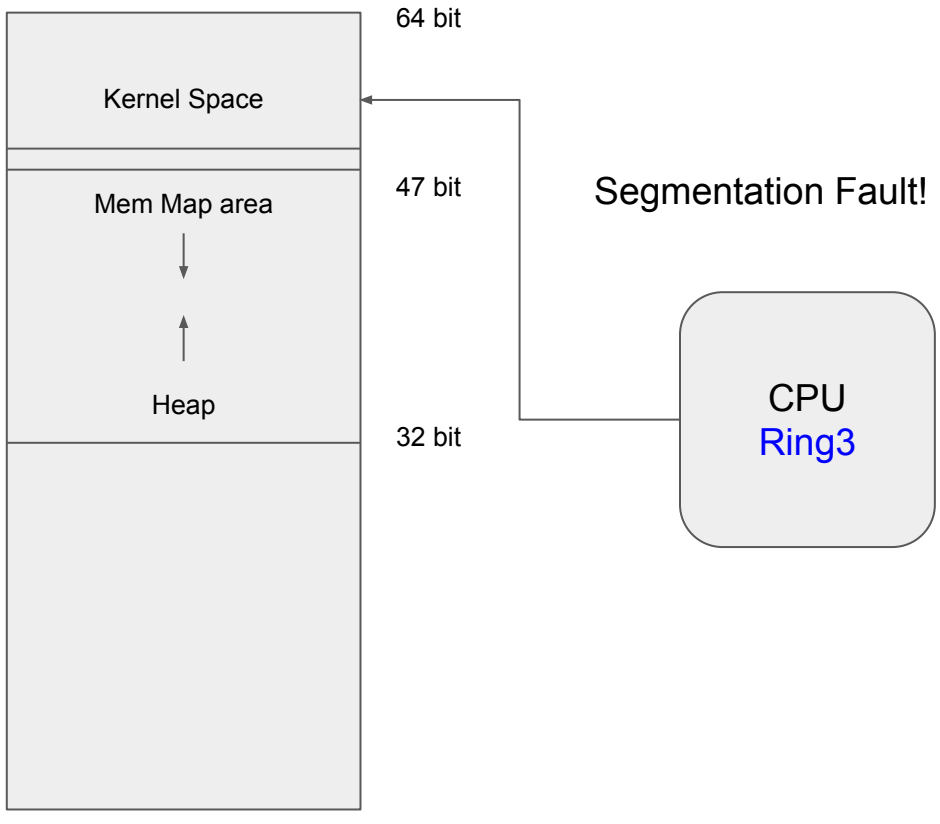
instaclustr



Process Virtual Memory space (not including ASLR offsets)

Memory Protection

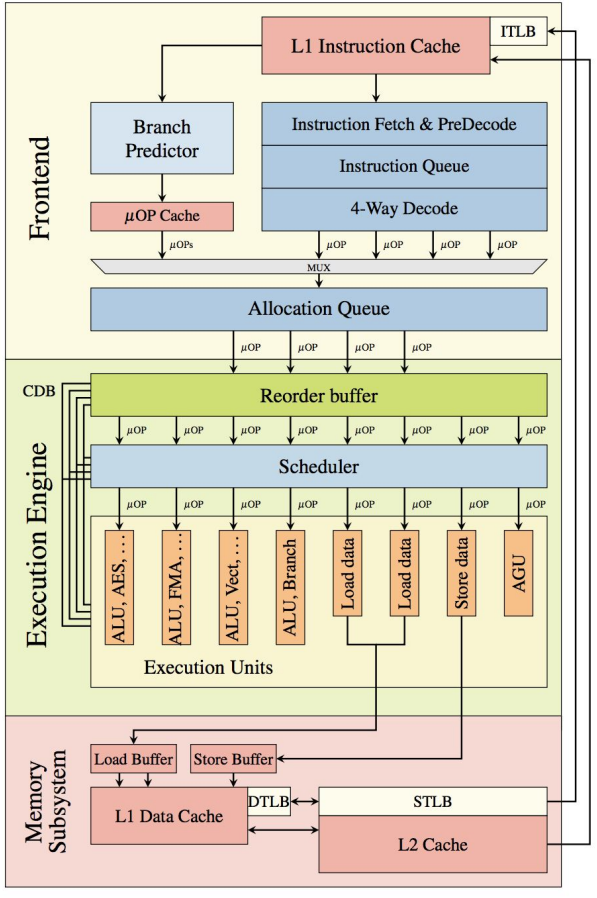
instaclustr



Virtual Memory space (not including ASLR offsets)

Out of order execution

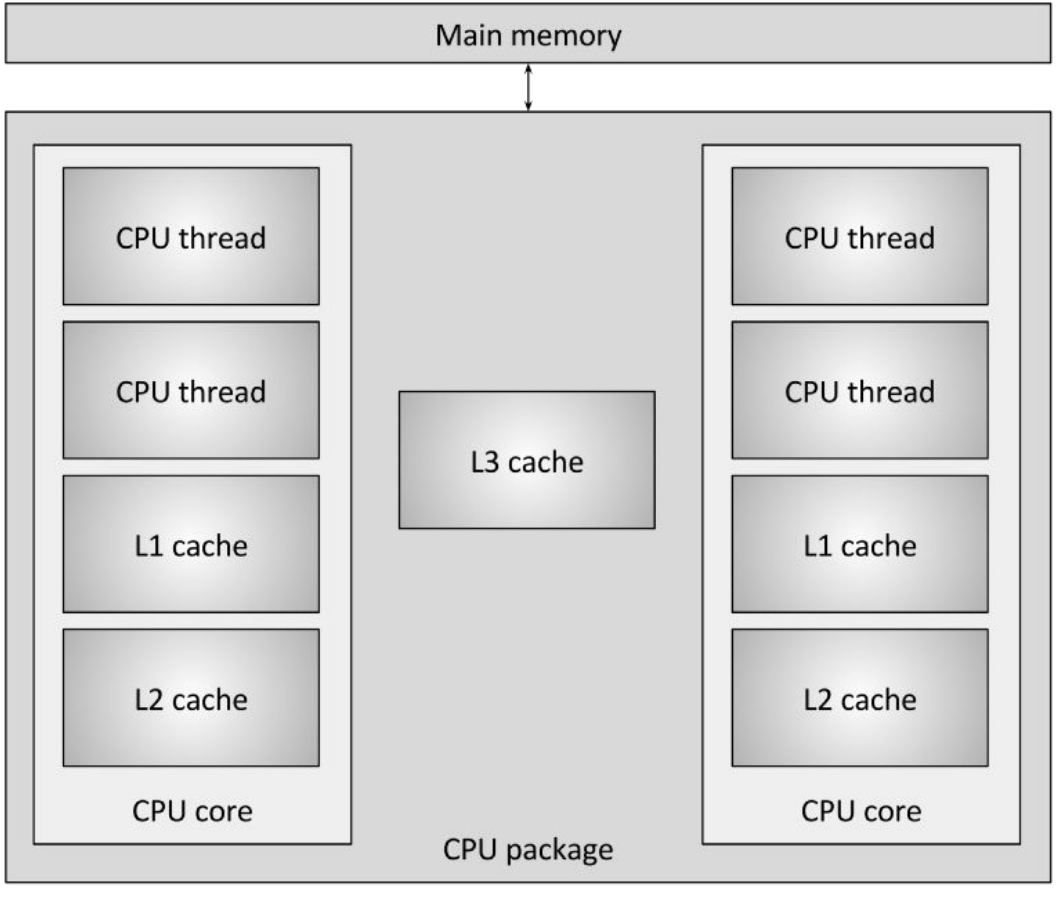
instaclustr



Courtesy - https://meltdownattack.com/meltdown.pdf

Cache behavior

instaclustr

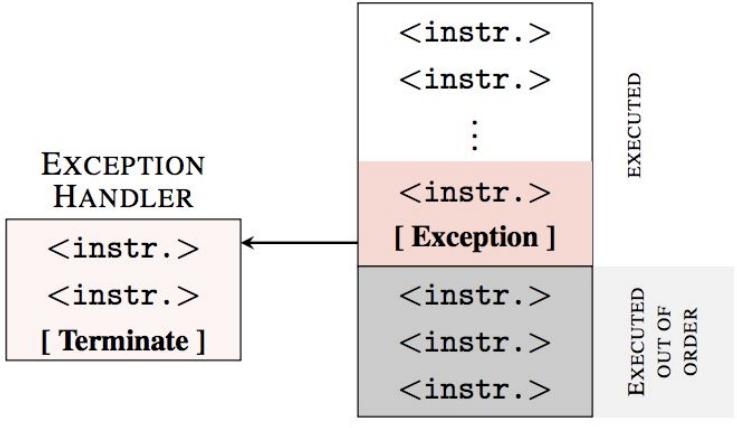


Courtesy - https://medium.com/@mattklein123

Putting it all together - Meltdown

instaclustr

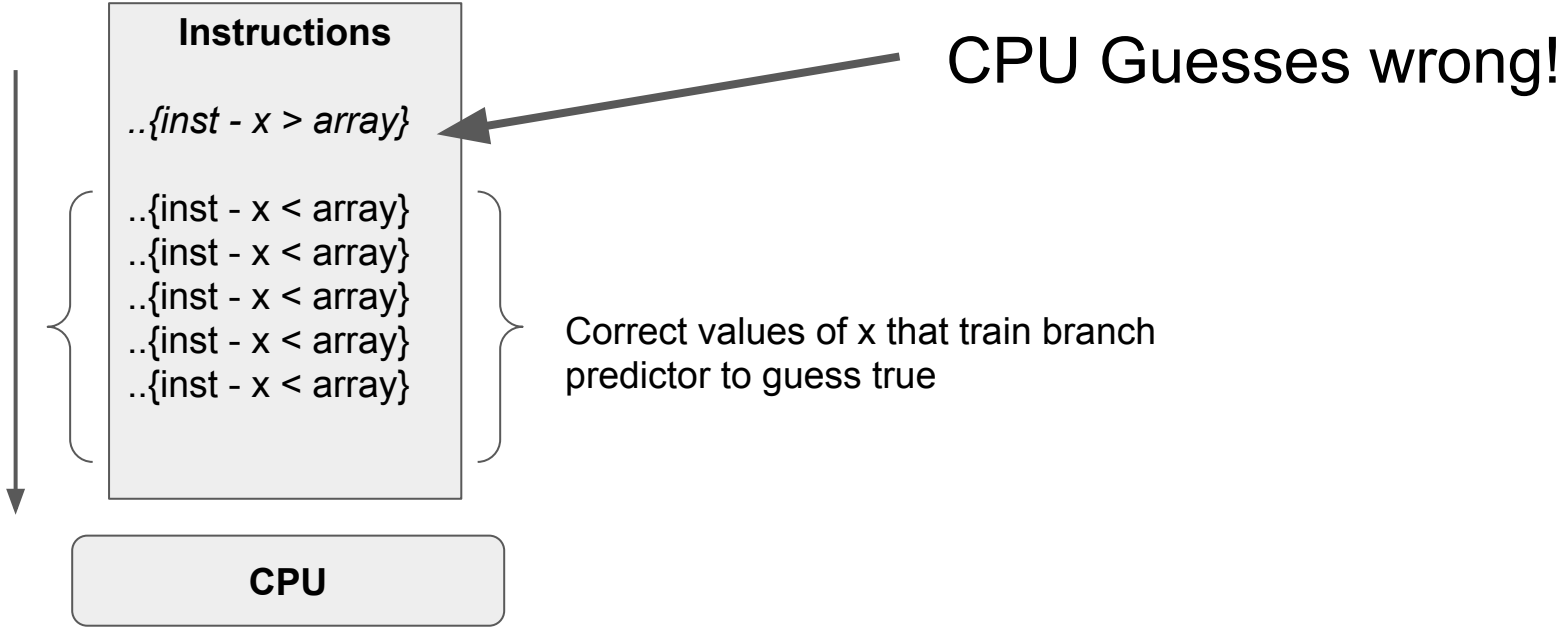
```
1|   raise_exception();
2|   // the line below is never reached
3|   access(probe_array[data * 4096]);
```



Putting it all together - Spectre

instaclustr

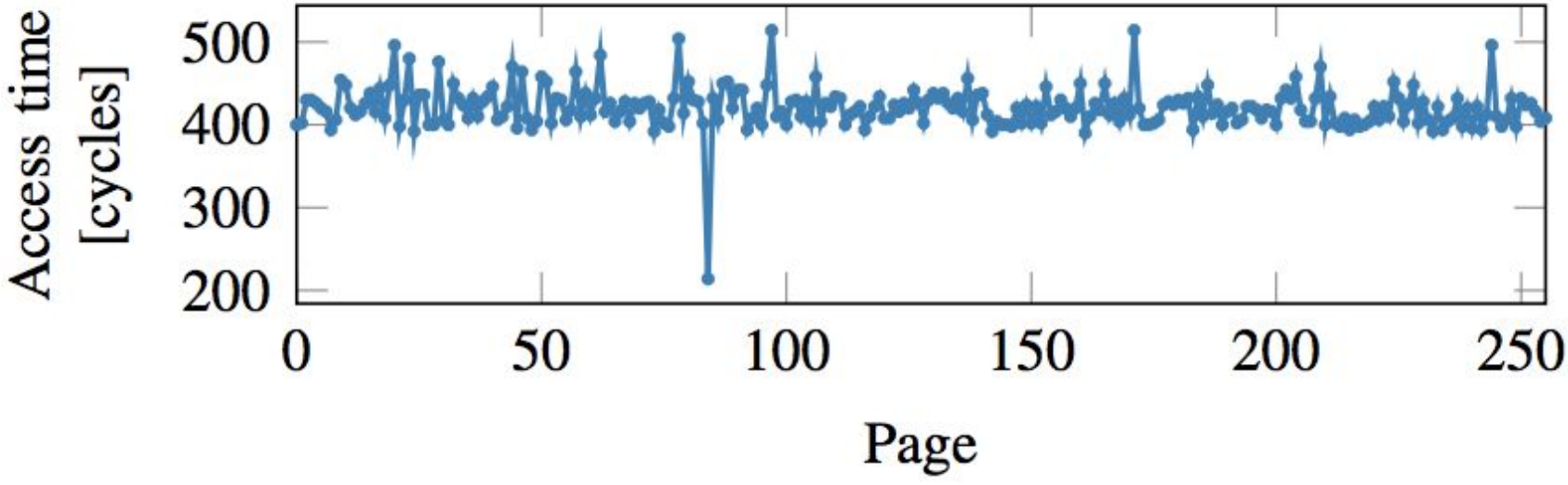
```
1|   if (x < array1_size)
2|       y = array2[array1[x] * 256];
```



Putting it all together

instaclustr

```
1|   raise_exception();
2|   // the line below is never reached
3|   access(probe_array[data * 4096]);
```



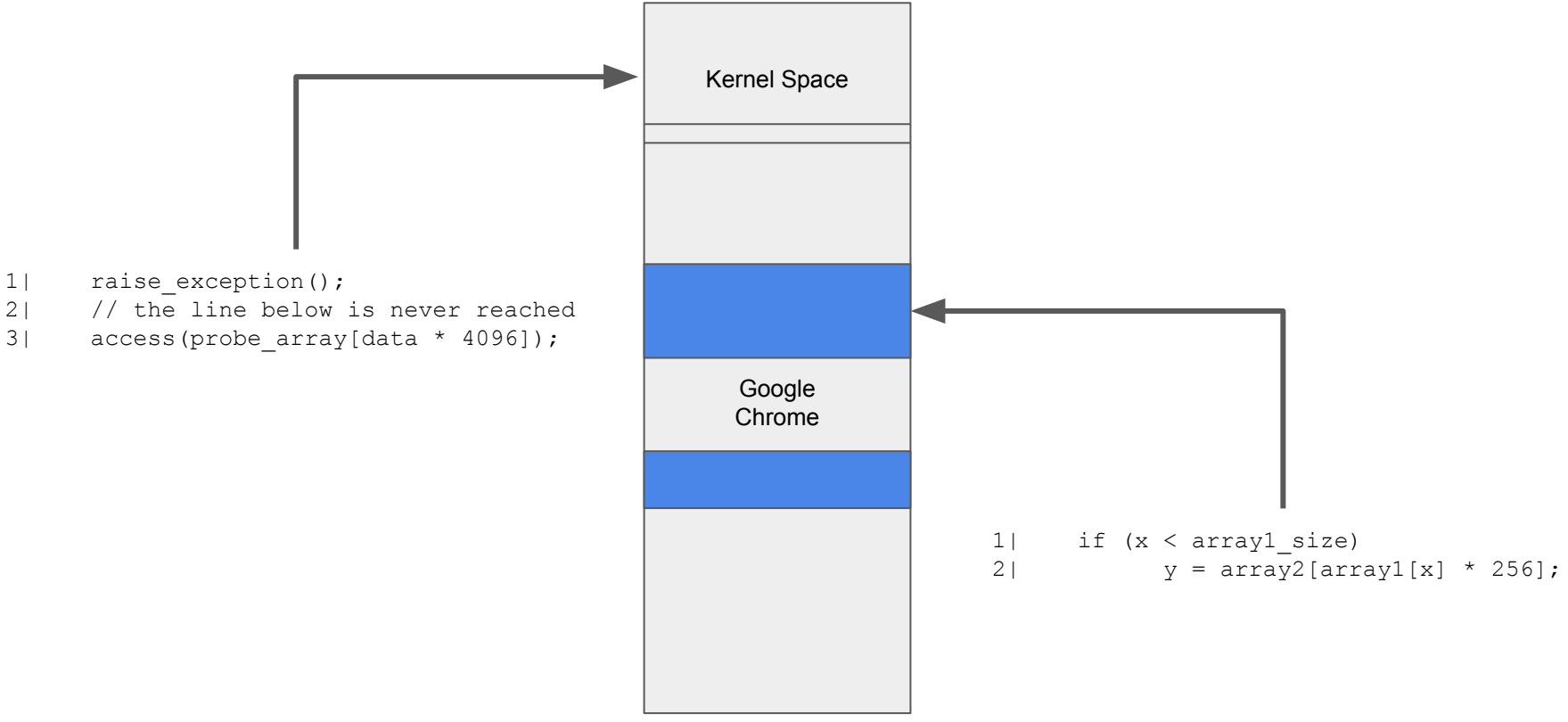
Meltdown vs Spectre

instaclustr

- Both leverage CPU speculative execution
  - Meltdown - out of order, non dependent execution
  - Spectre - branch prediction
- Both exfiltrate data via timing of cache access
- Meltdown can access protected memory (including that of a hypervisor as well as kernel) due to Intel not checking the privilege bit upon instructions executed speculatively

Meltdown vs Spectre

instaclustr



Context Switching (Kernel mapped into all user process spaces)

The fixes

instaclustr

Spectre variant 1 - Recompile everything

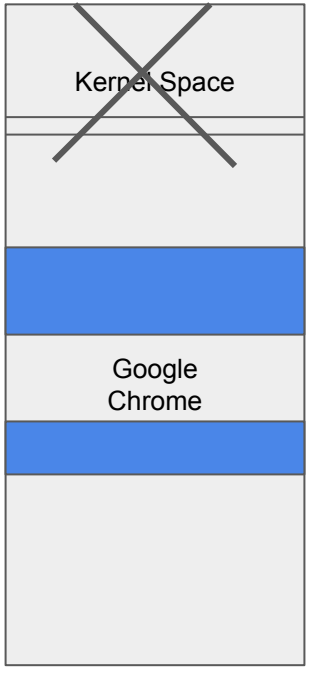
Spectre variant 2 - CPU microcode updates OR Google Retpoline

Meltdown - Linux patchset call KPTI (Kernel Page Isolation)

KPTI

instaclustr

Isolates Kernel page table from user space (still in physical memory, just not process page table)



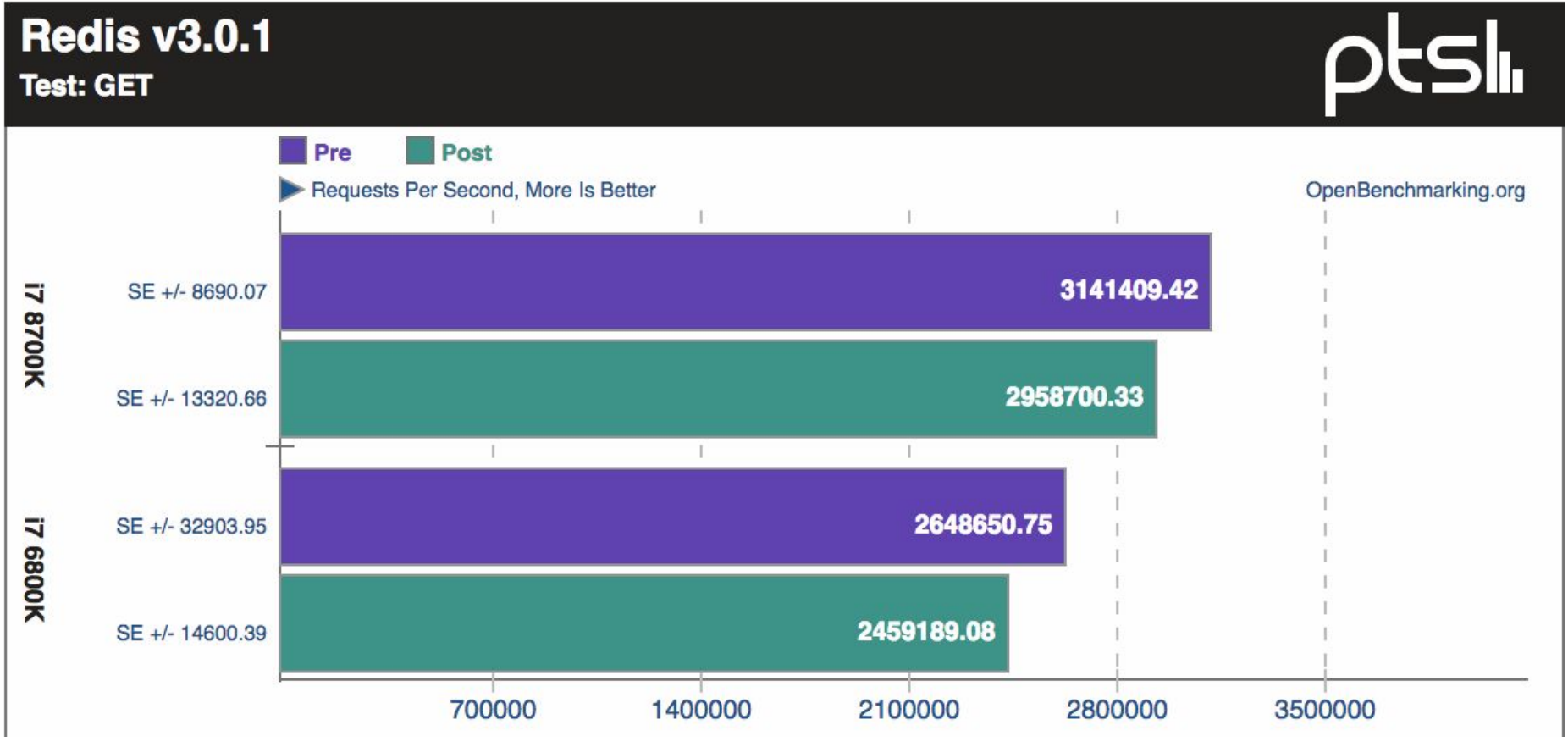
KPTI Performance impact

instaclustr

- KPTI unmaps the majority of the kernel virtual memory while user code is on CPU.
- Performance impact of this swap estimated to be between 5 - 30% depending on the number of system calls a process makes.
- How does this impact Cassandra?

KPTI Performance impact

instaclustr



KPTI Performance impact

instaclustr

Meltdown's Impact on Cassandra Latency

10 Jan 2018

What impact on latency should you expect from applying the kernel patches for the Meltdown security vulnerability?

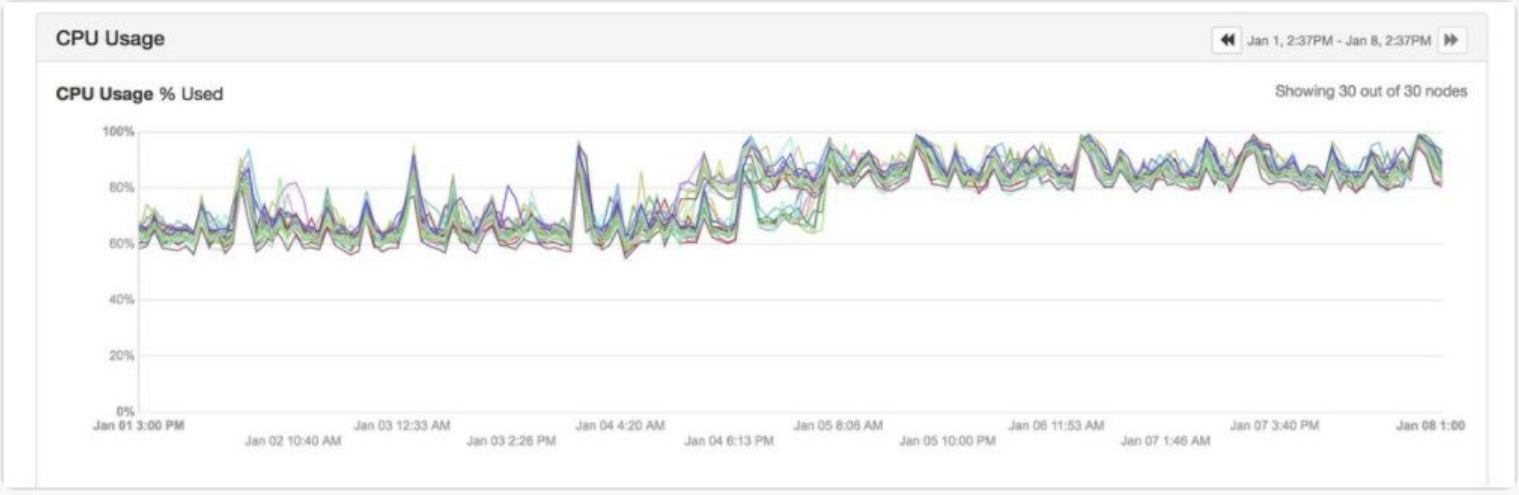
TL;DR expect a latency increase of at least 20% for both reads and writes.

KPTI Performance impact

instaclustr

Ben Bromhead @BenBromhead · Jan 8

CPU utilisation hit on one of our own Cassandra clusters due to #Meltdown #spectre. AWS impact only. No patches on the guest OS.



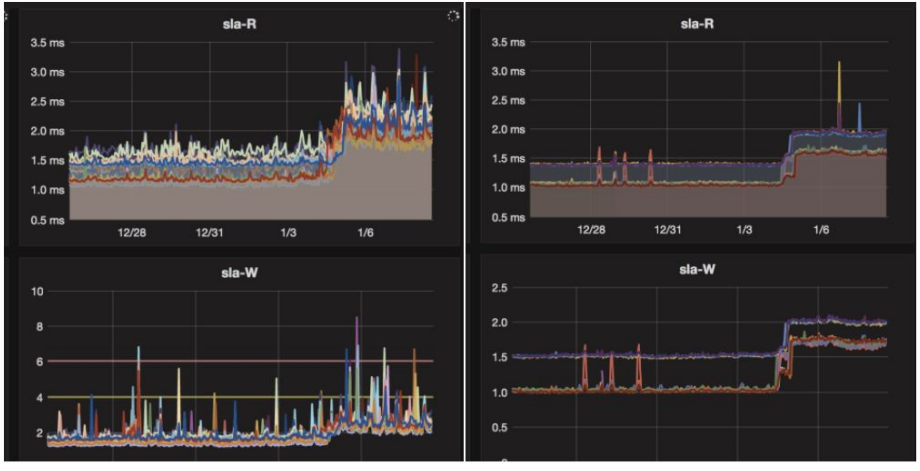
6 51 54

KPTI Performance impact

instaclustr

Ben Bromhead
@BenBromhead

#Meltdown #spectre Impact on p50 Cassandra latency for two arbitrary production clusters. Taken from synthetic transaction monitoring (a simple canary we run)



8:35 AM - 9 Jan 2018

Wait what

instaclustr

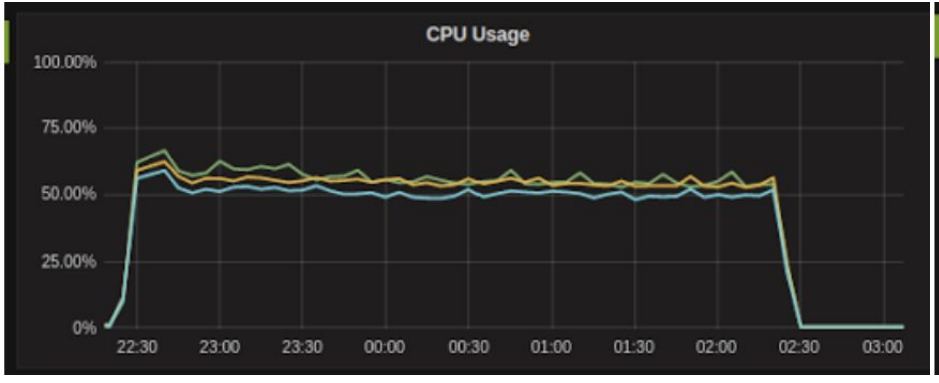
- During the first round of testing we noticed a slight increase in performance AFTER upgrading.
- We observed a similar increase in performance across our fleet of managed nodes and concluded this was due to a patch made to the underlying AWS hypervisors and not related to the guest OS upgrades - the tests were repeated to confirm results.

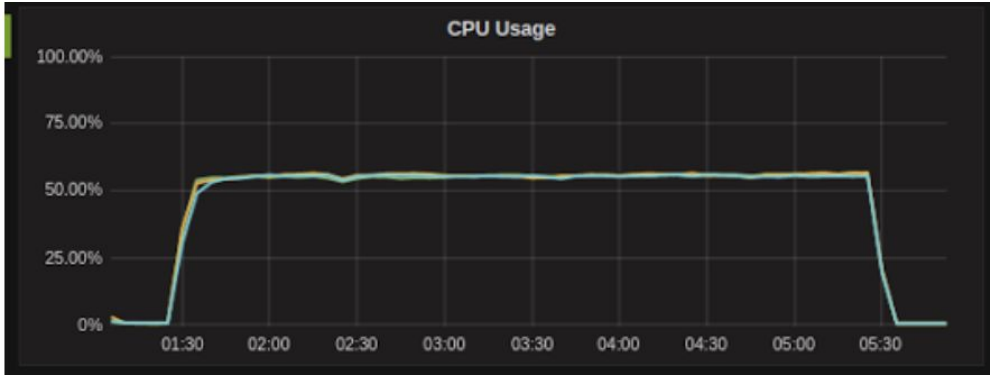
Impact after resolved AWS patches

instaclustr

Impact after resolved AWS patches

instaclustr





Real world impact after patches

instaclustr

Over our AWS fleet of 1500 nodes (m4.xl, r4.xl and i3.2xl) serving real world production workloads

- Slight increase to no increase in CPU utilisation (especially in the cloud)
- Minimal impact on production environments
- Slight increase in base line latency
- Most of this stuff appears to be within a certain margin of error anyway

What about the other benchmarks?

instaclustr

A number of companies conducted Cassandra benchmarking including:

- The Last Pickle
- Datastax
- Scylla

All showed an impact at a throughput maximum. E.g. a cluster pushed to its absolute limit.

instaclustr

Questions?

Ben Bromhead
CTO
ben@instaclustr.com