L1TF in KVM

About Me

- Alexander Graf
- KVM and QEMU developer for SUSE
  - Server class PowerPC KVM port
  - Nested SVM
- Founding member of SUSE ARM team
- U-Boot UEFI support

Speculation

Speculation

```
static long last_add;

long read_add(long *data, long add)
{
   last_add += add;
   return *data + add;
}
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

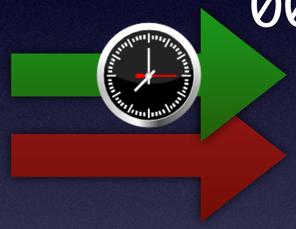
Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation



```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

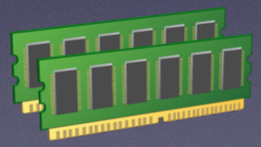
Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Speculation

```
0000000000000000 <read_add>:
   0:48 01 35 00 00 00 00   add    %rsi,last_add(%rip)
   7:48 89 f0               mov    %rsi,%rax
   a:48 03 07               add    (%rdi),%rax
   d:c3                     retq
```

Caches



xx GB

~250 cycles

`last_add(%rip)`

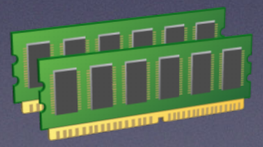
Latency numbers from https://www.7-cpu.com/cpu/Skylake.html

Caches

L3

8 MB

42 cycles



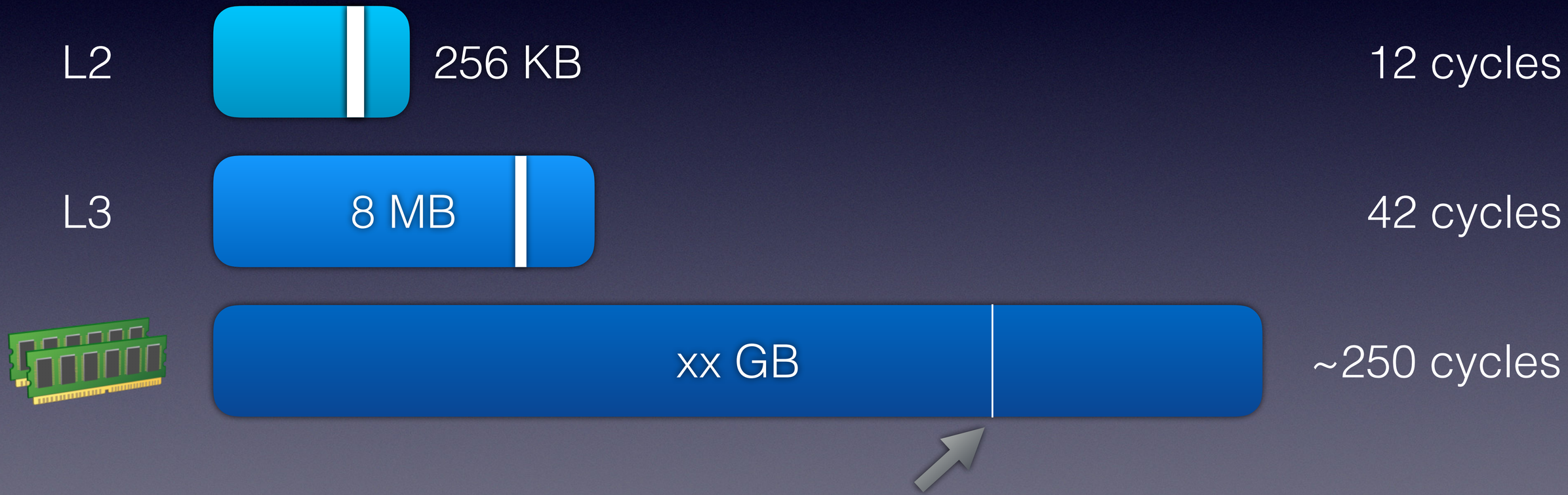
xx GB

~250 cycles

`last_add(%rip)`

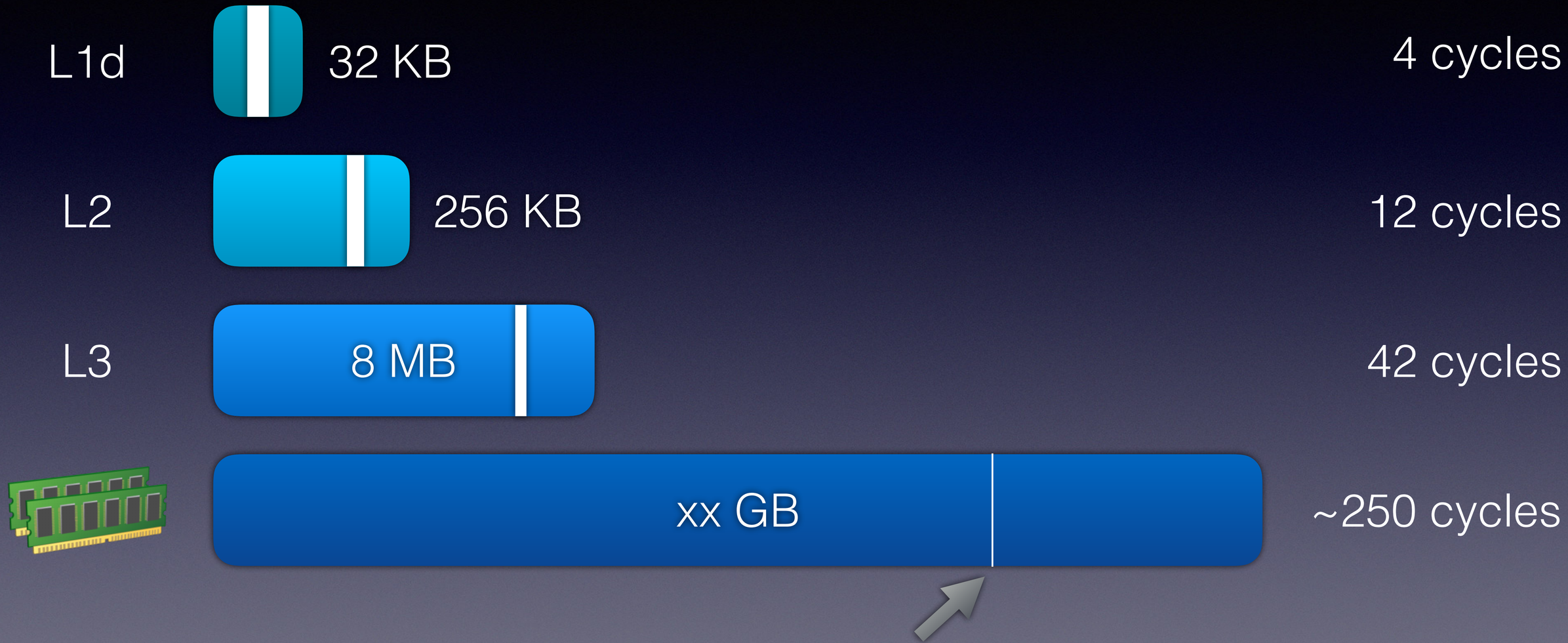
Latency numbers from https://www.7-cpu.com/cpu/Skylake.html

Caches



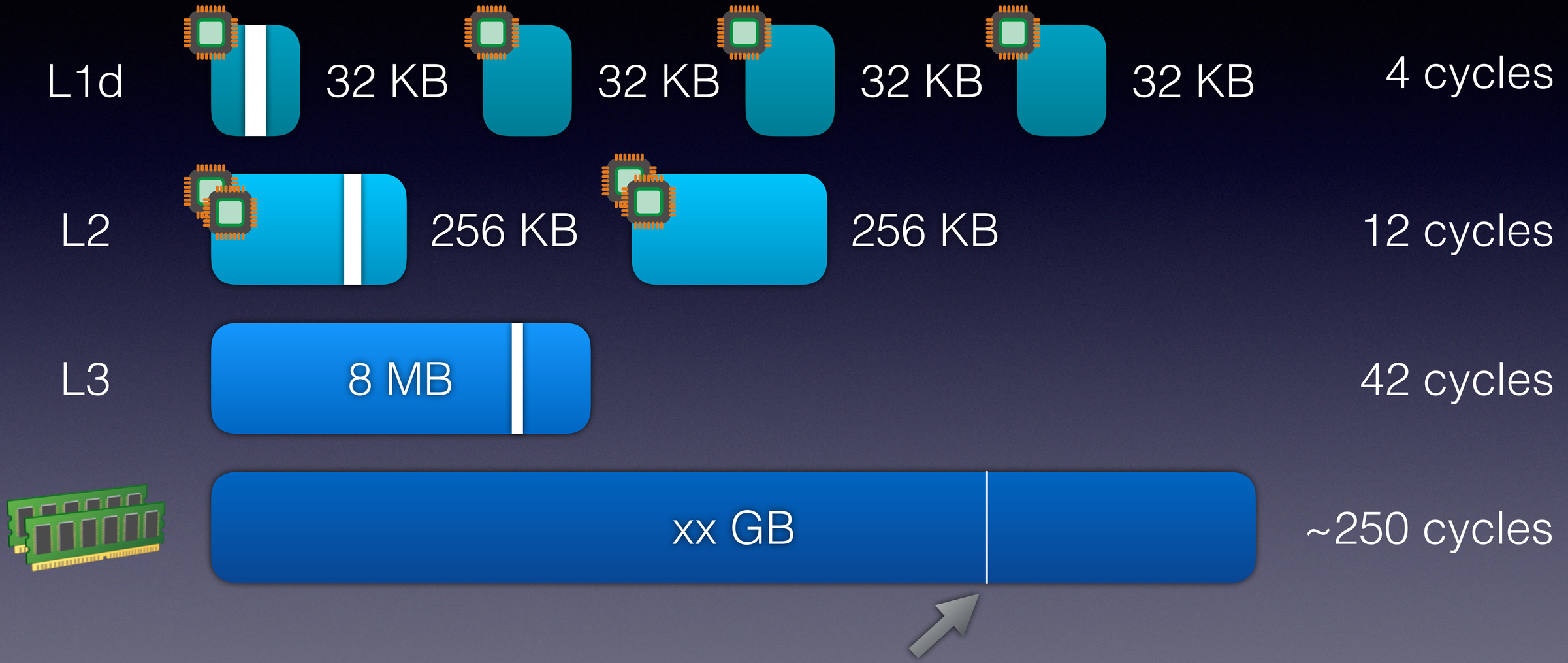
Latency numbers from https://www.7-cpu.com/cpu/Skylake.html

Caches



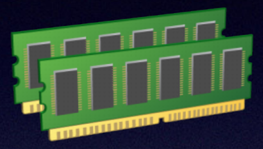
Latency numbers from https://www.7-cpu.com/cpu/Skylake.html

Caches

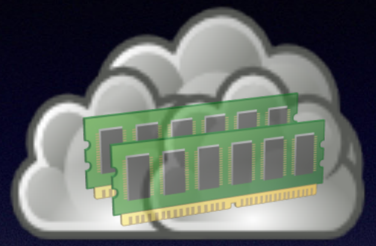


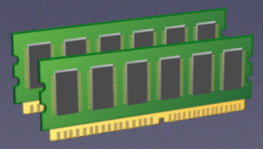
Latency numbers from https://www.7-cpu.com/cpu/Skylake.html

Paging



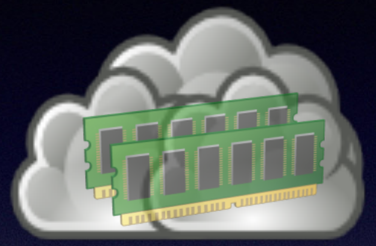
Paging



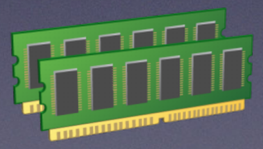


xx GB

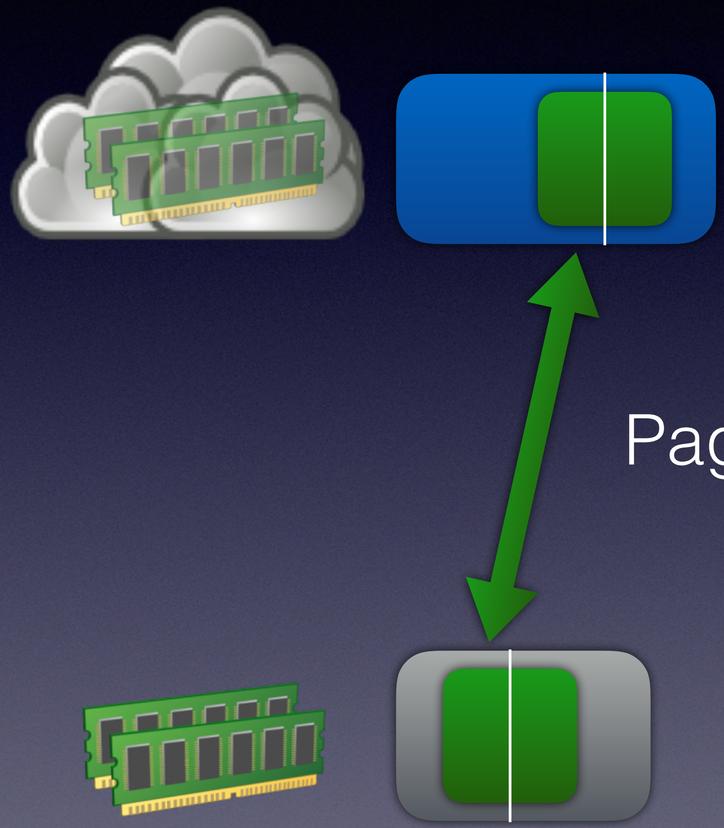
Paging



Page Table Entry



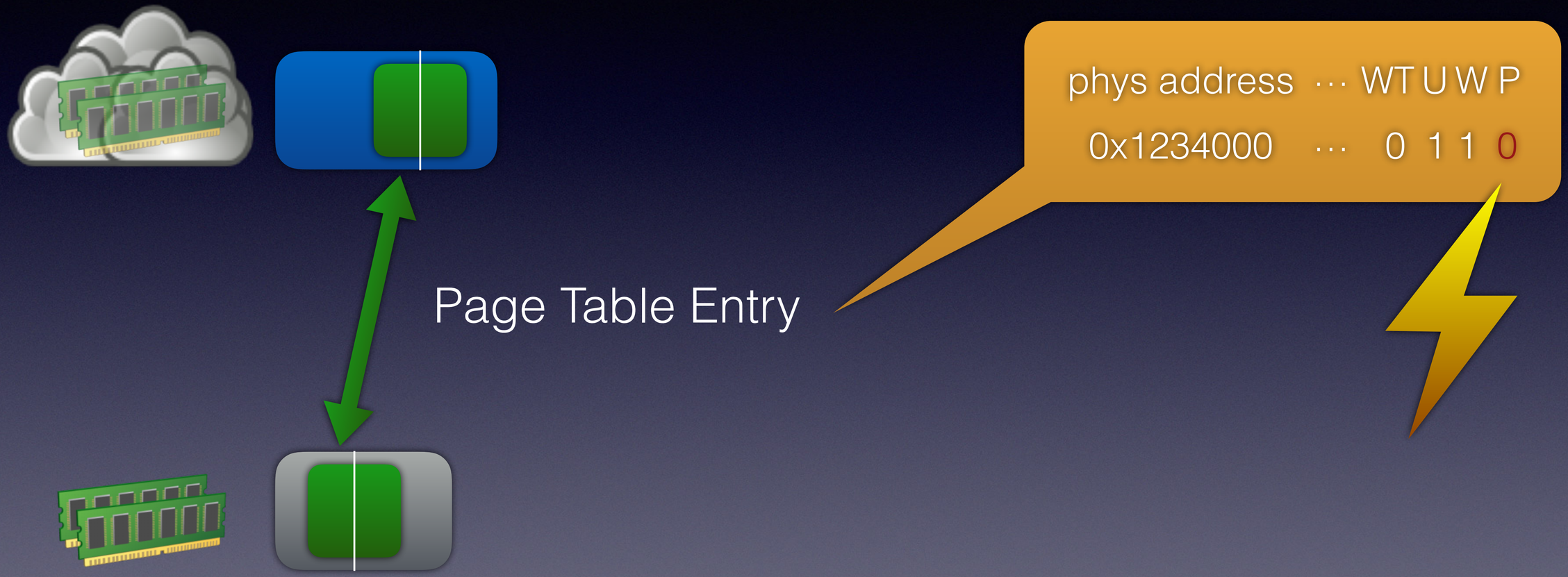
Paging



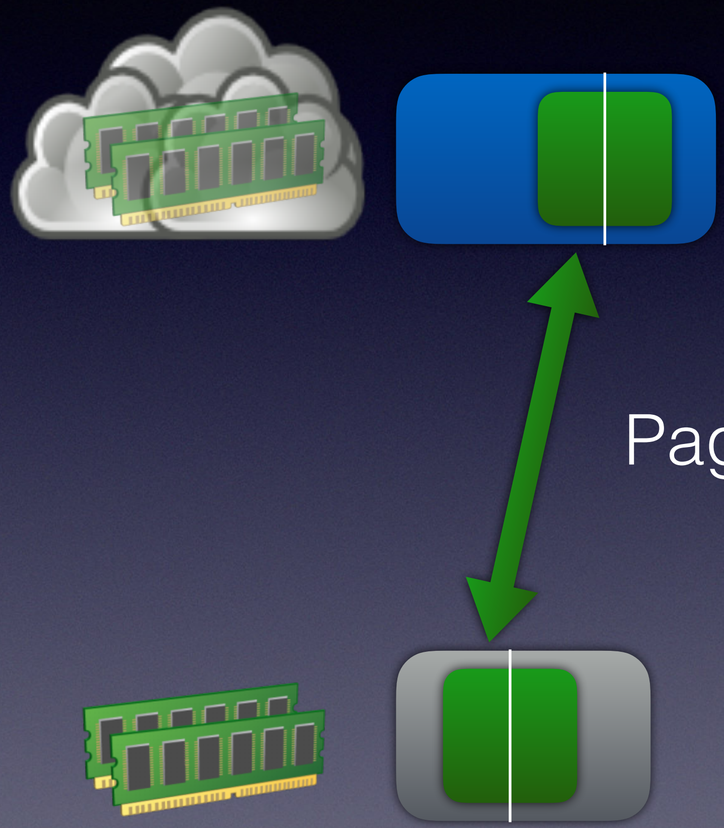
Page Table Entry

phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	1

Paging



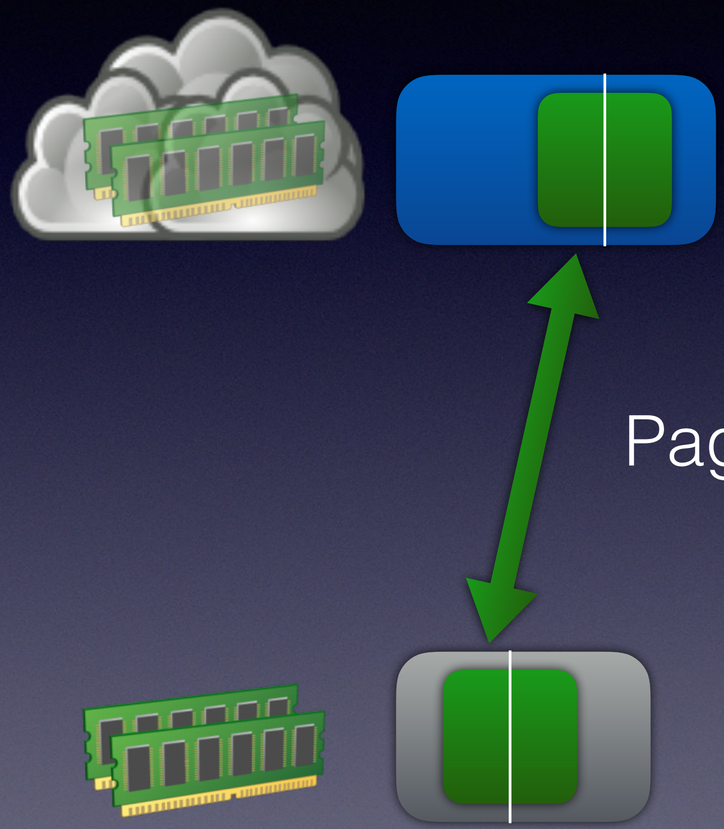
Paging



Page Table Entry

L1d

Paging



Page Table Entry

Paging

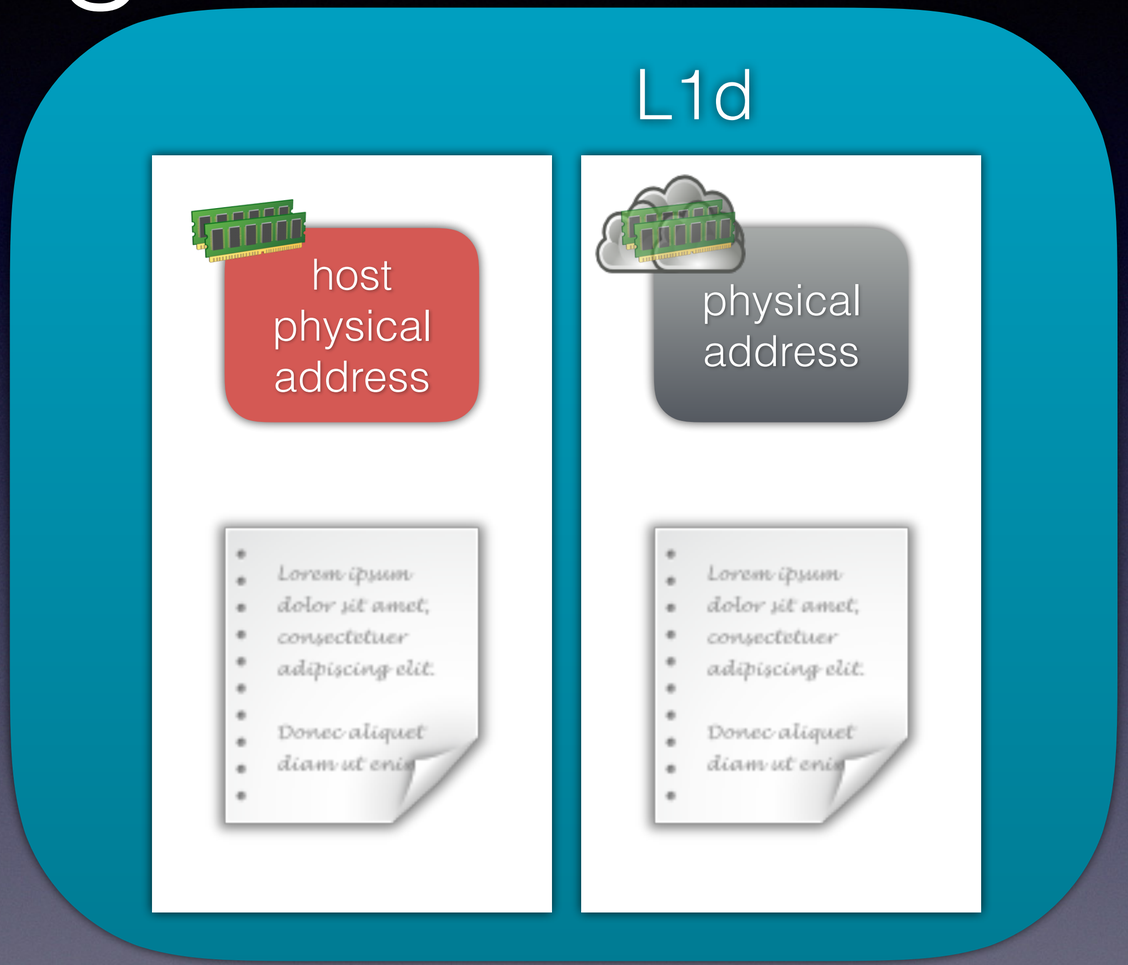
Page Table Entry

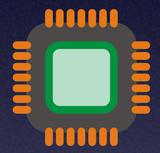
Extended
Page Table Entry

Paging

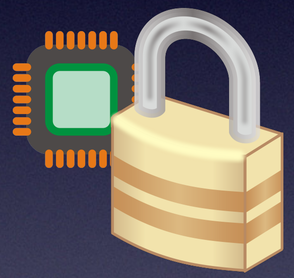
Page Table Entry

Extended
Page Table Entry

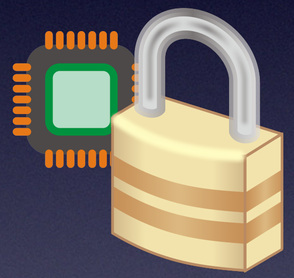


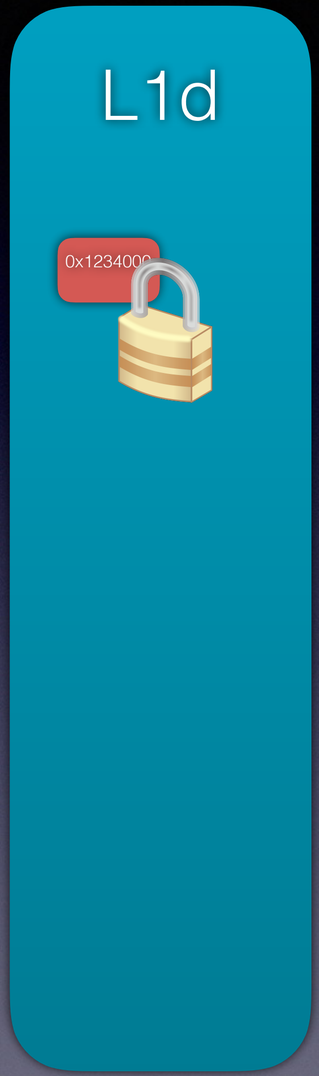


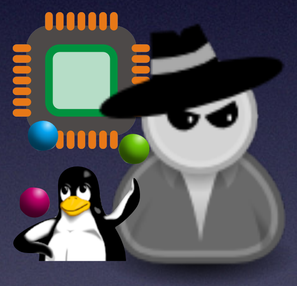
L1d



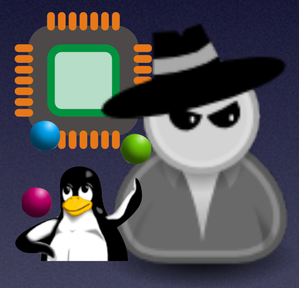
L1d



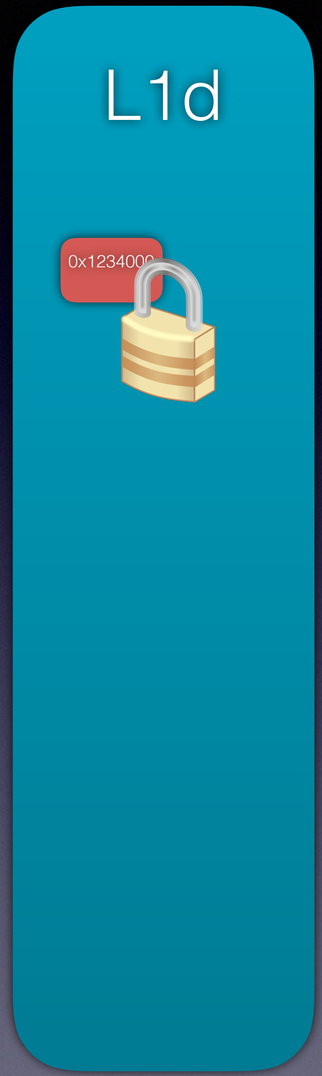

L1d
0x1234000

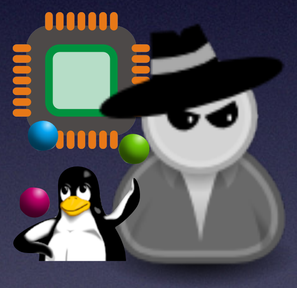


L1d

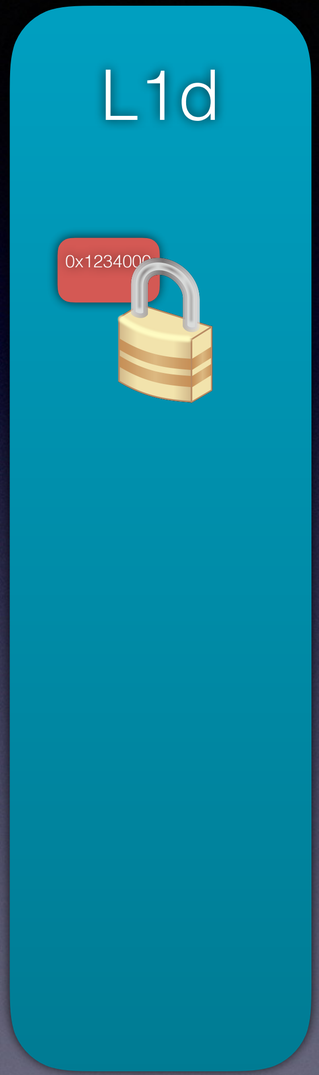


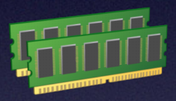
```
foo = *(char *)0x1000;
bar = oracle[foo * 4096];
```





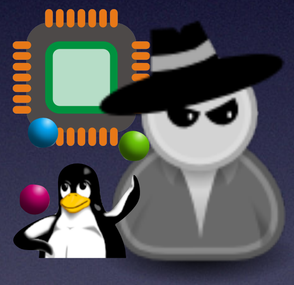
```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```





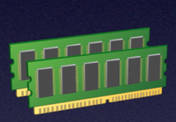
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	1

Page Table Entry



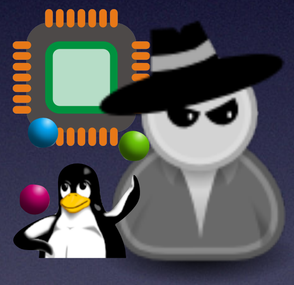
```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d



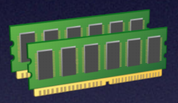
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	1

Page Table Entry



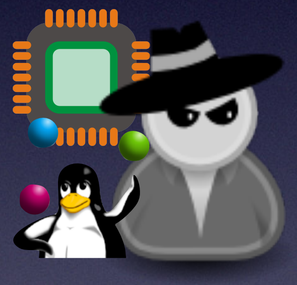
```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d



phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	1

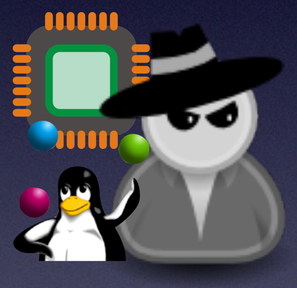
Page Table Entry



```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

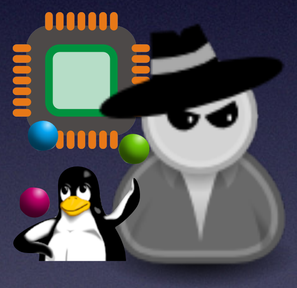
L1d

phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	1

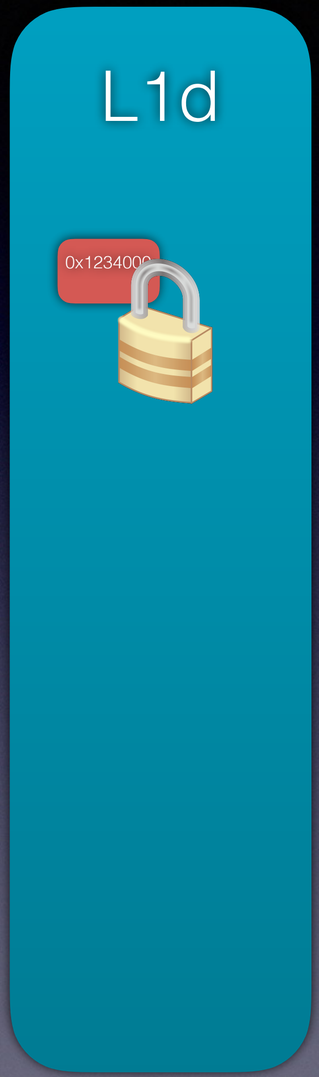


```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

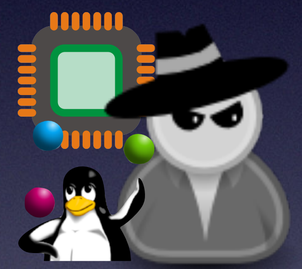
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0



```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```



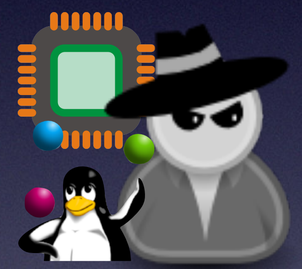
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0



```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d

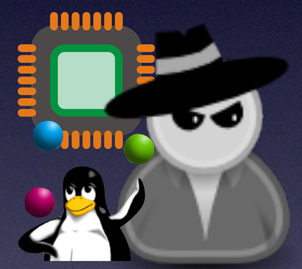
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0

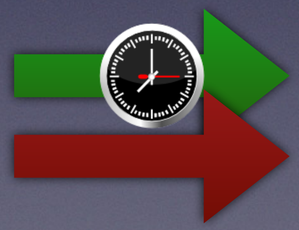


```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d

phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0

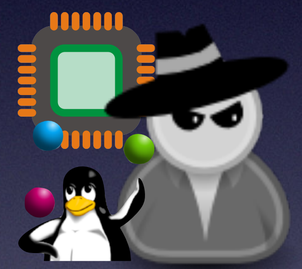




```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d

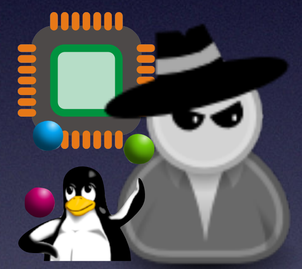
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0



```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

L1d

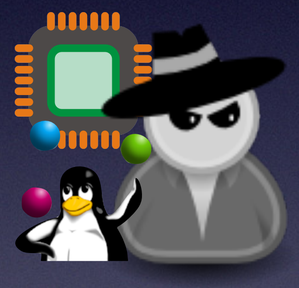
phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0



```
movsbq 0x1000,%rax
shl    $0xc,%rax
movsbq oracle(%rax),%rax
```

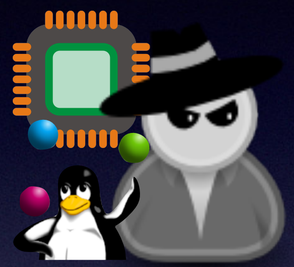
L1d

phys address	⋯	WT	U	W	P
0x1234000	⋯	0	1	1	0

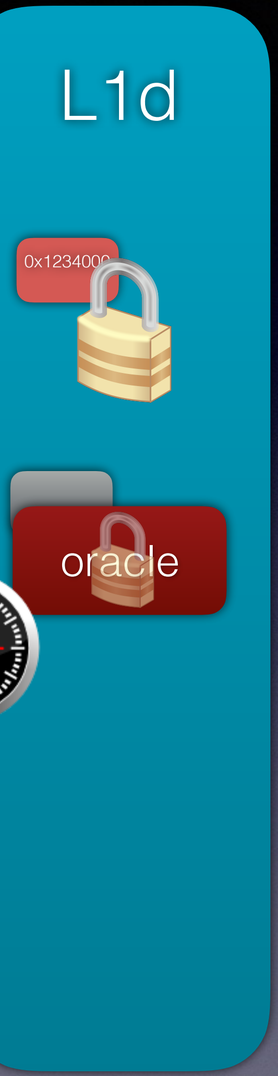


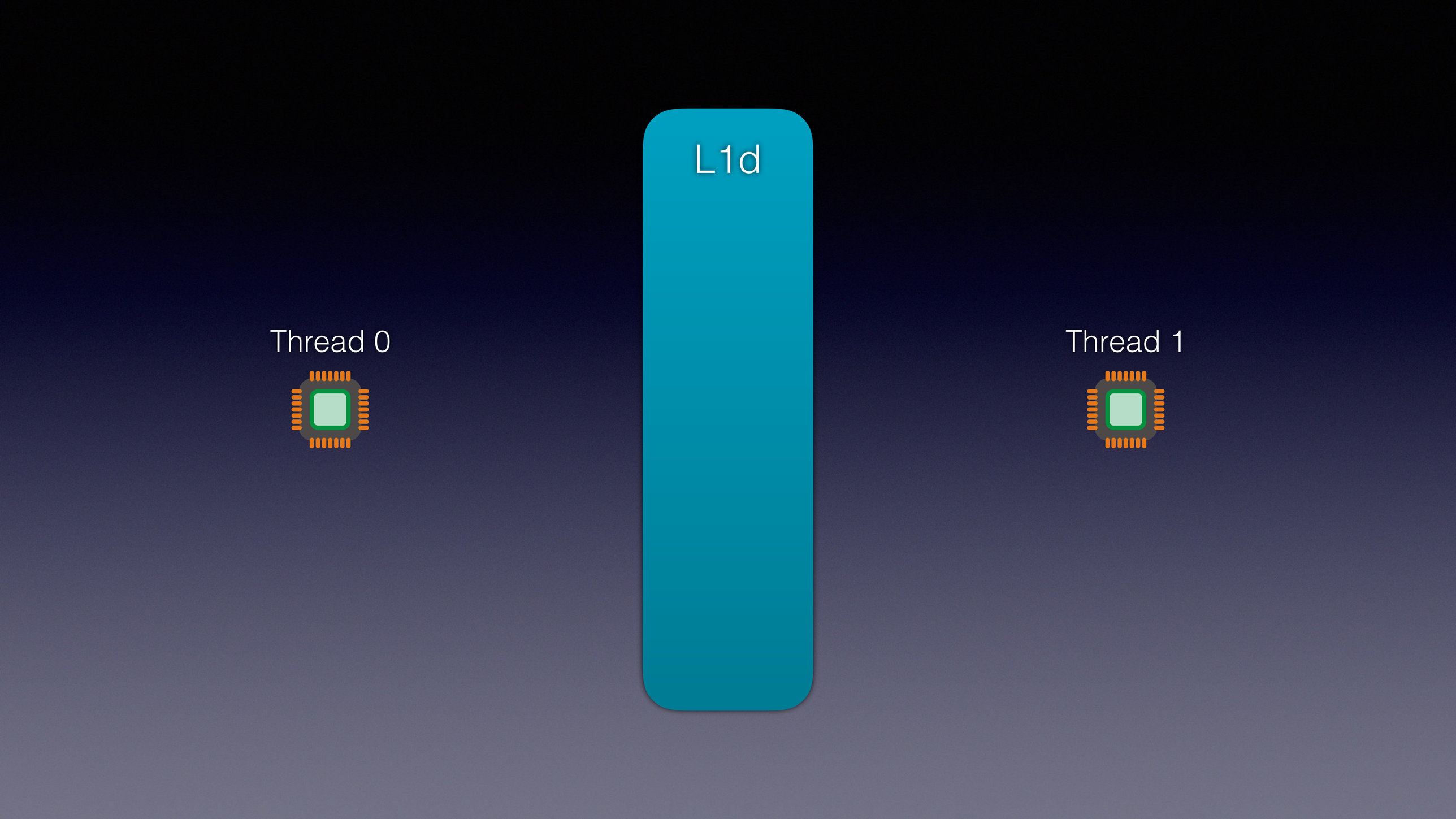
```
movsbq 0x1000,%rax
```

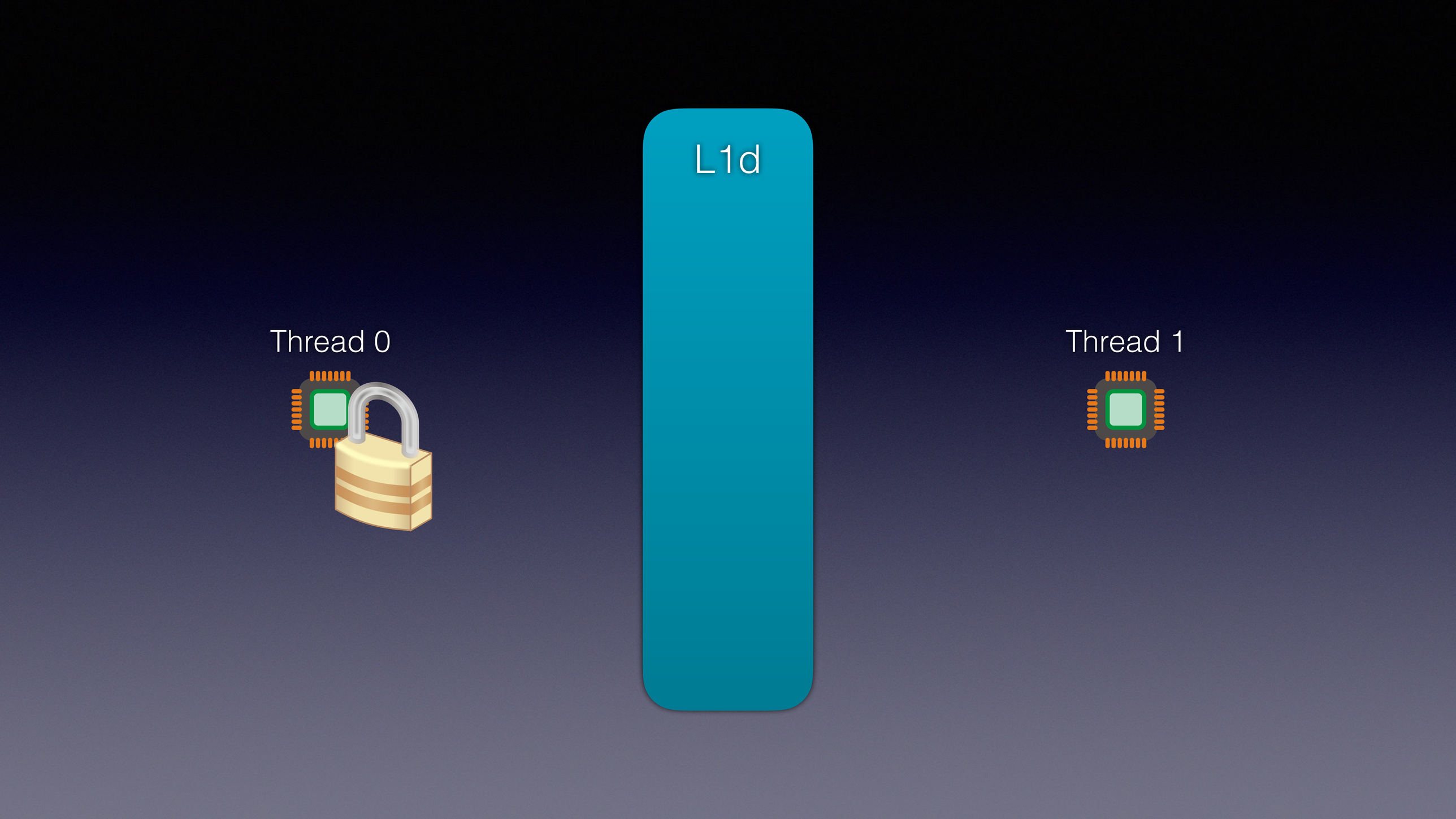
L1d




oracle

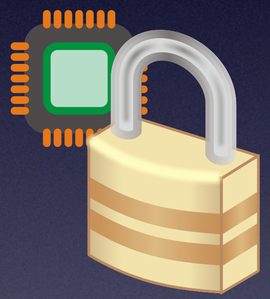

L1d
0x1234000
oracle


L1d
Thread 0
Thread 1

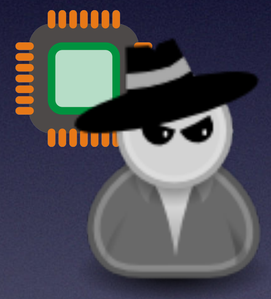

L1d
Thread 0
Thread 1

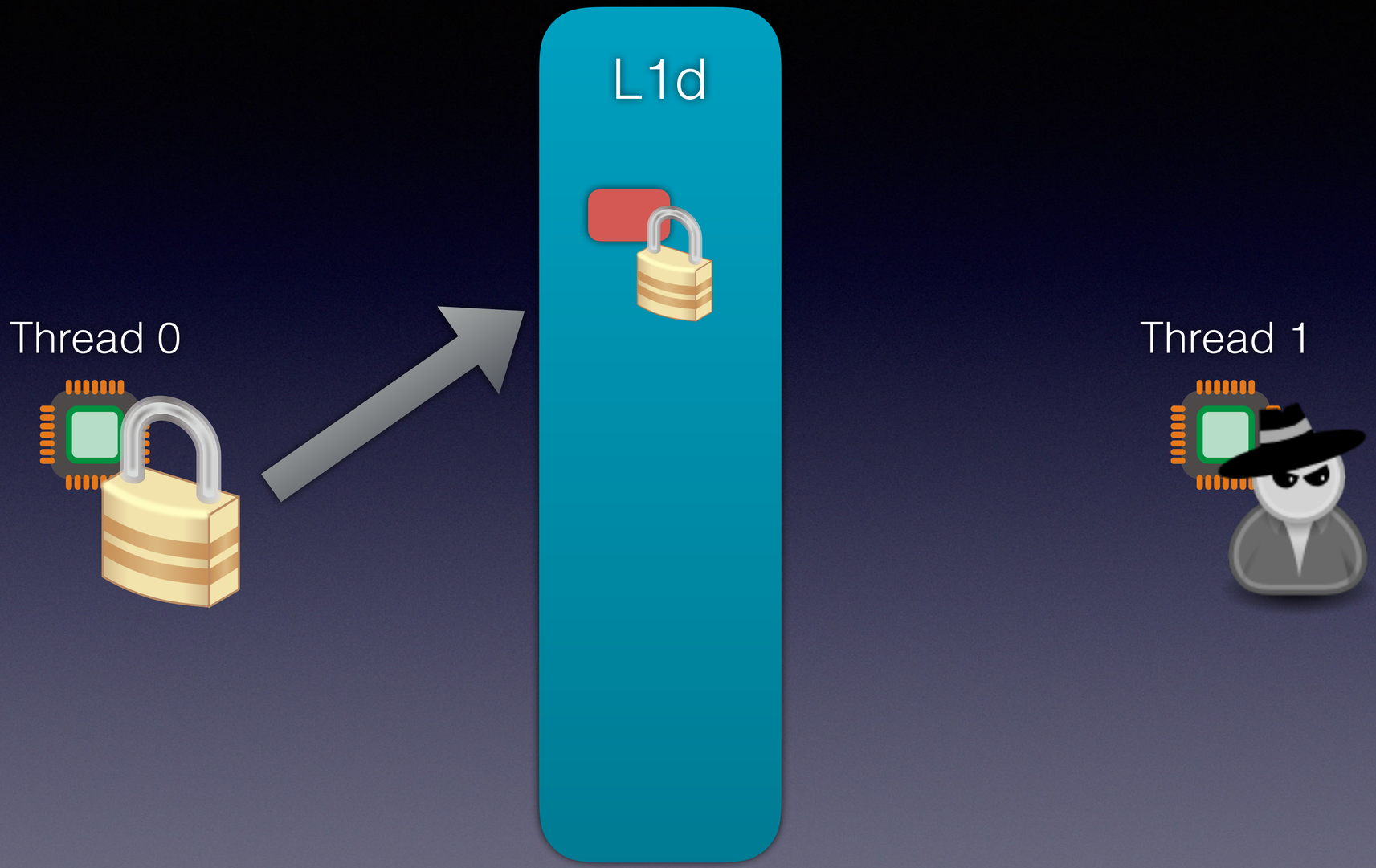
L1d

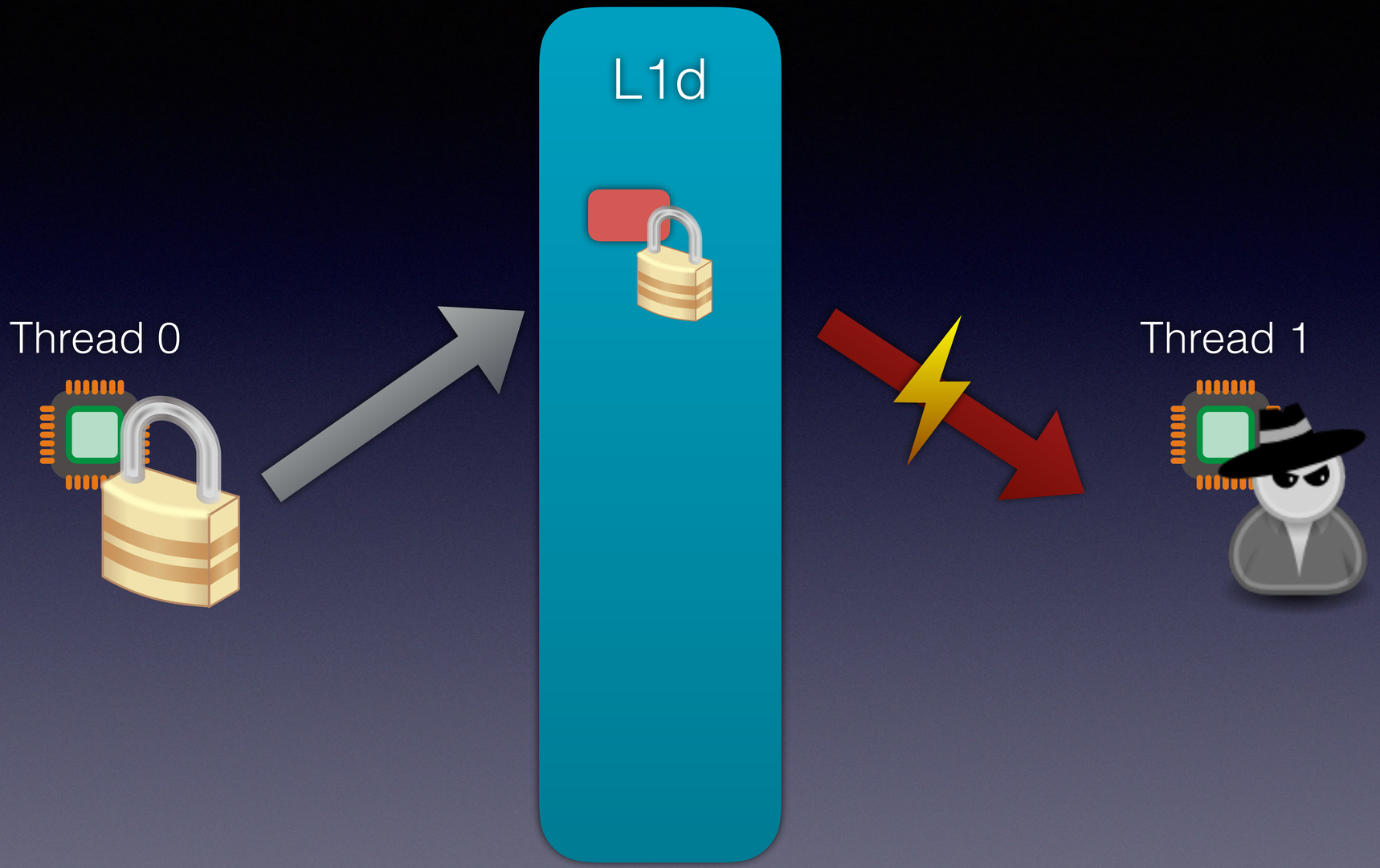
Thread 0



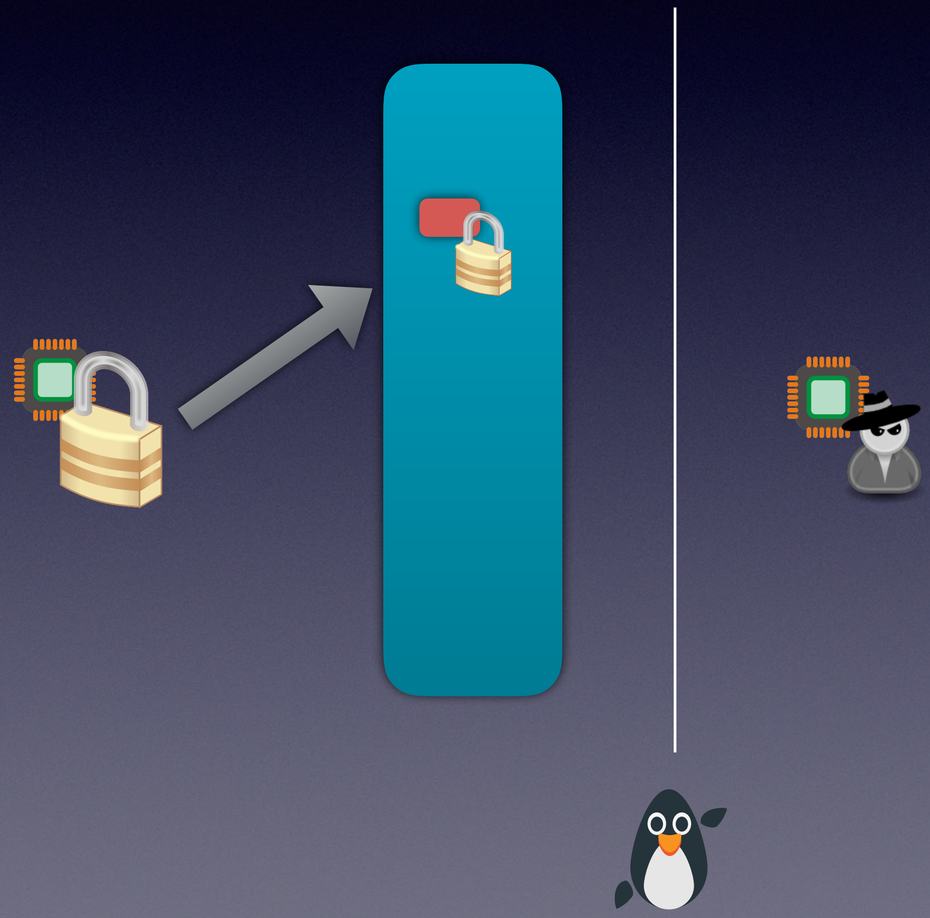
Thread 1



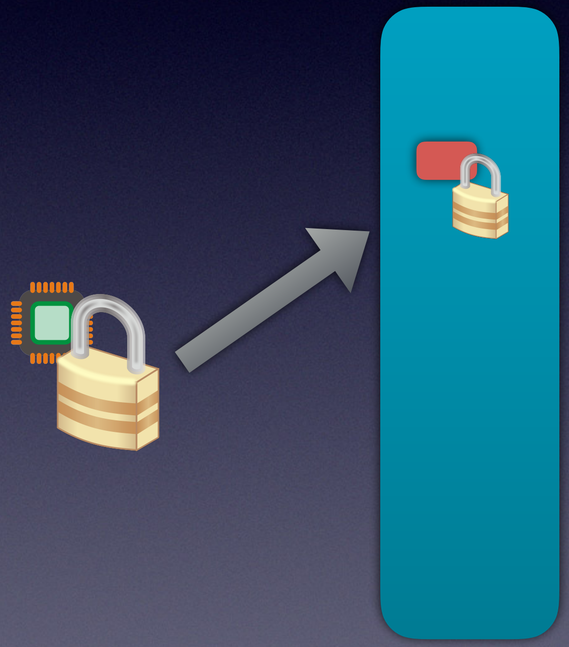

L1d
Thread 0
Thread 1

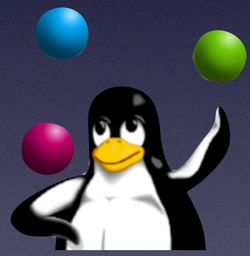

L1d
Thread 0
Thread 1

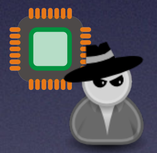
Mitigation



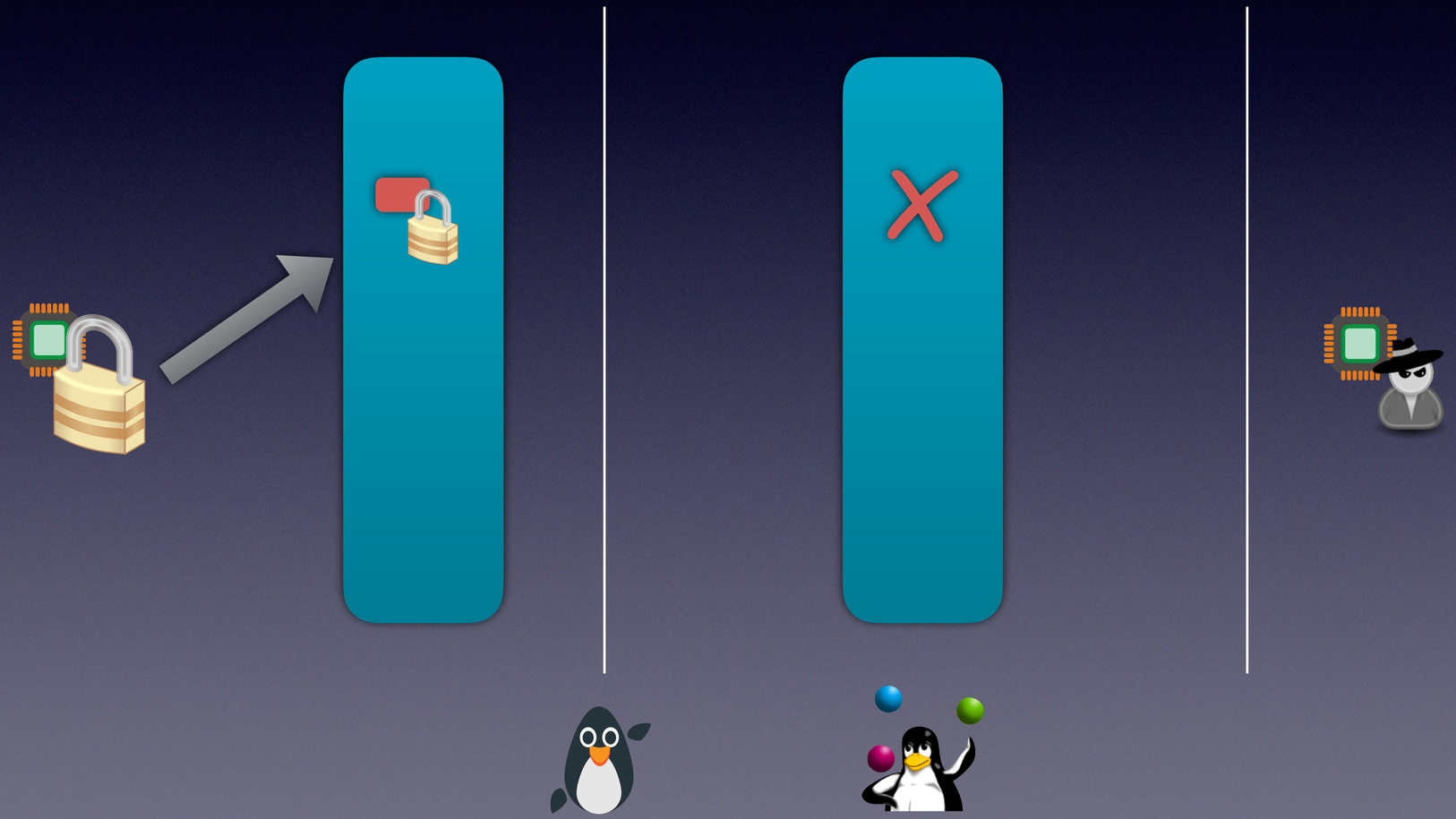
Mitigation



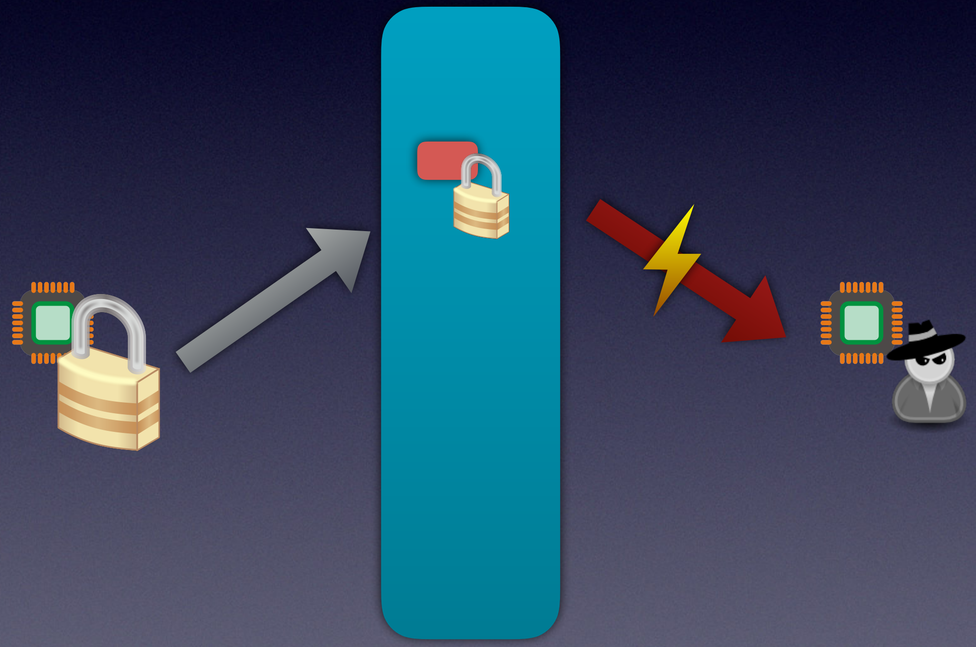




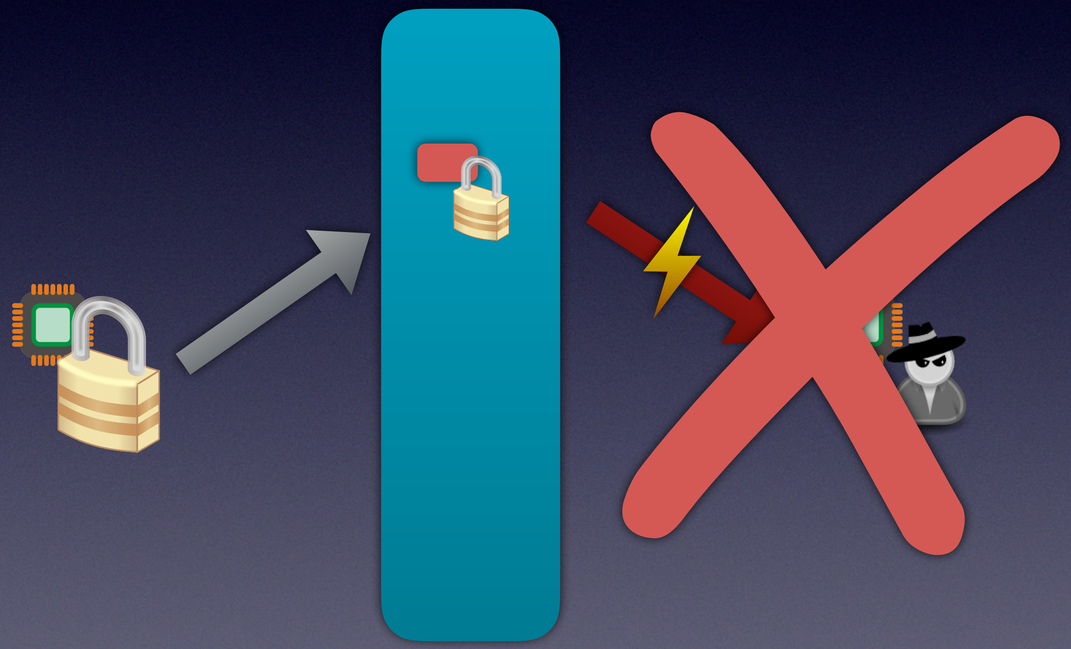
Mitigation



Mitigation

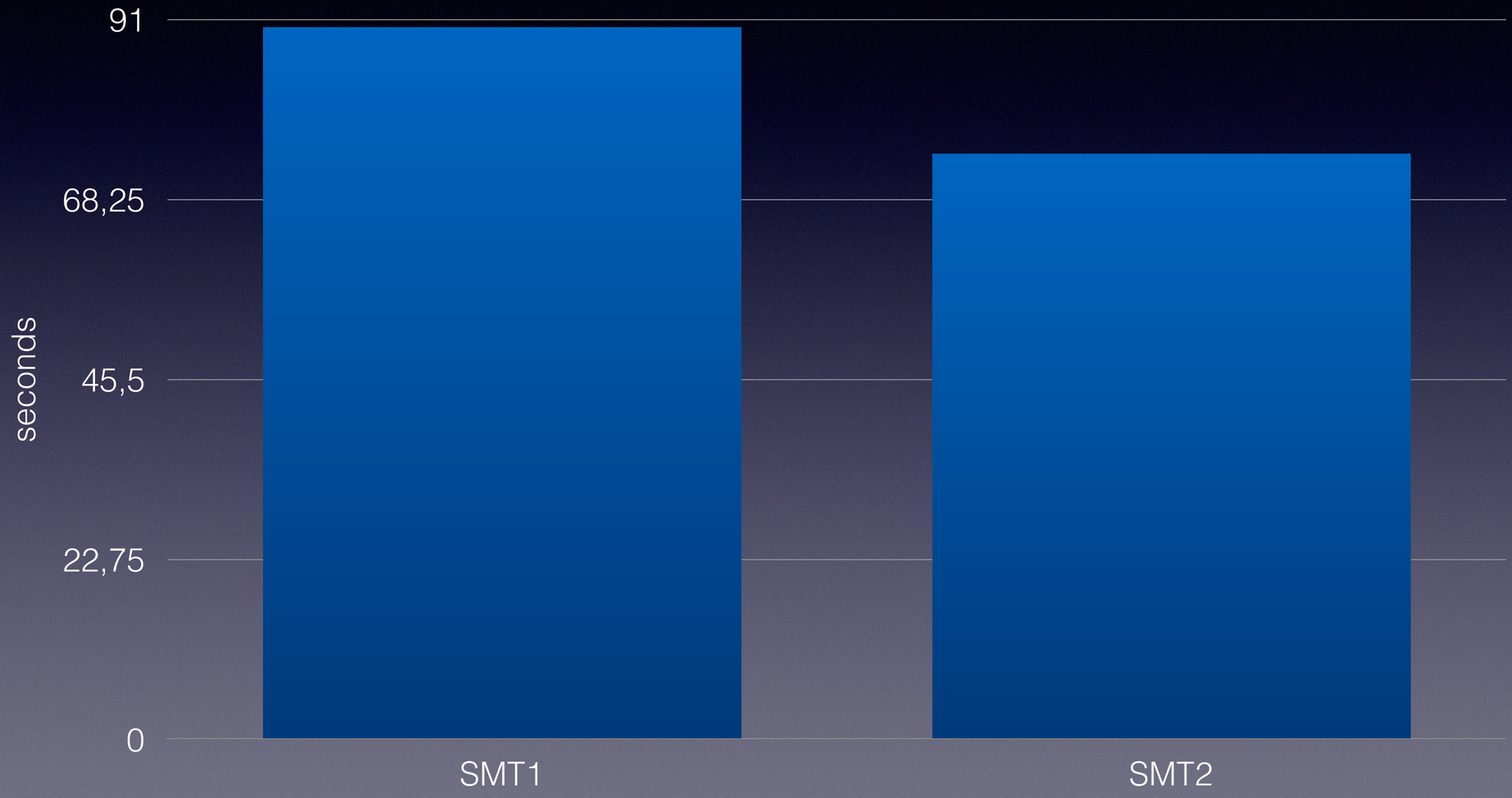


Mitigation

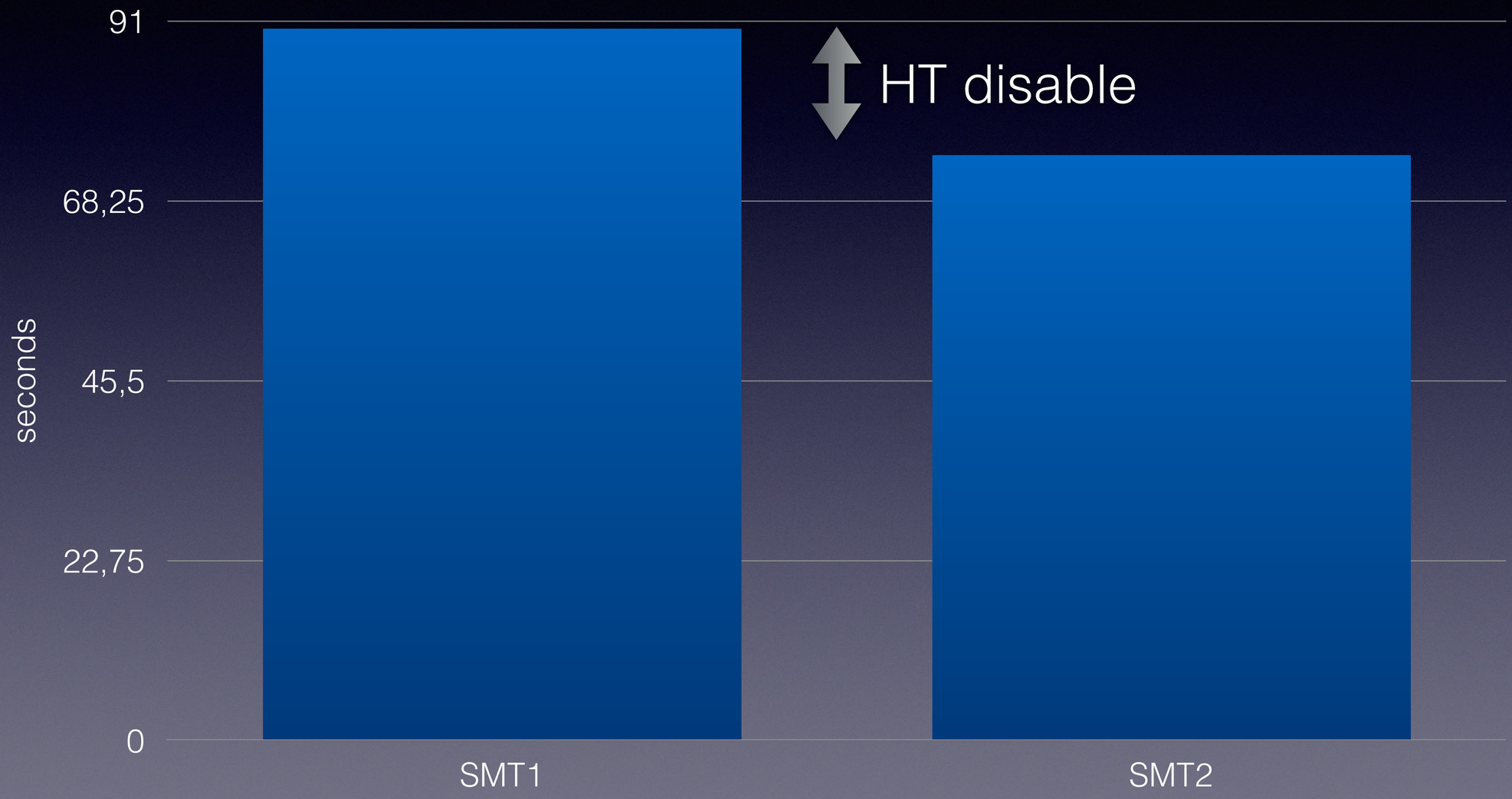


Kernel build

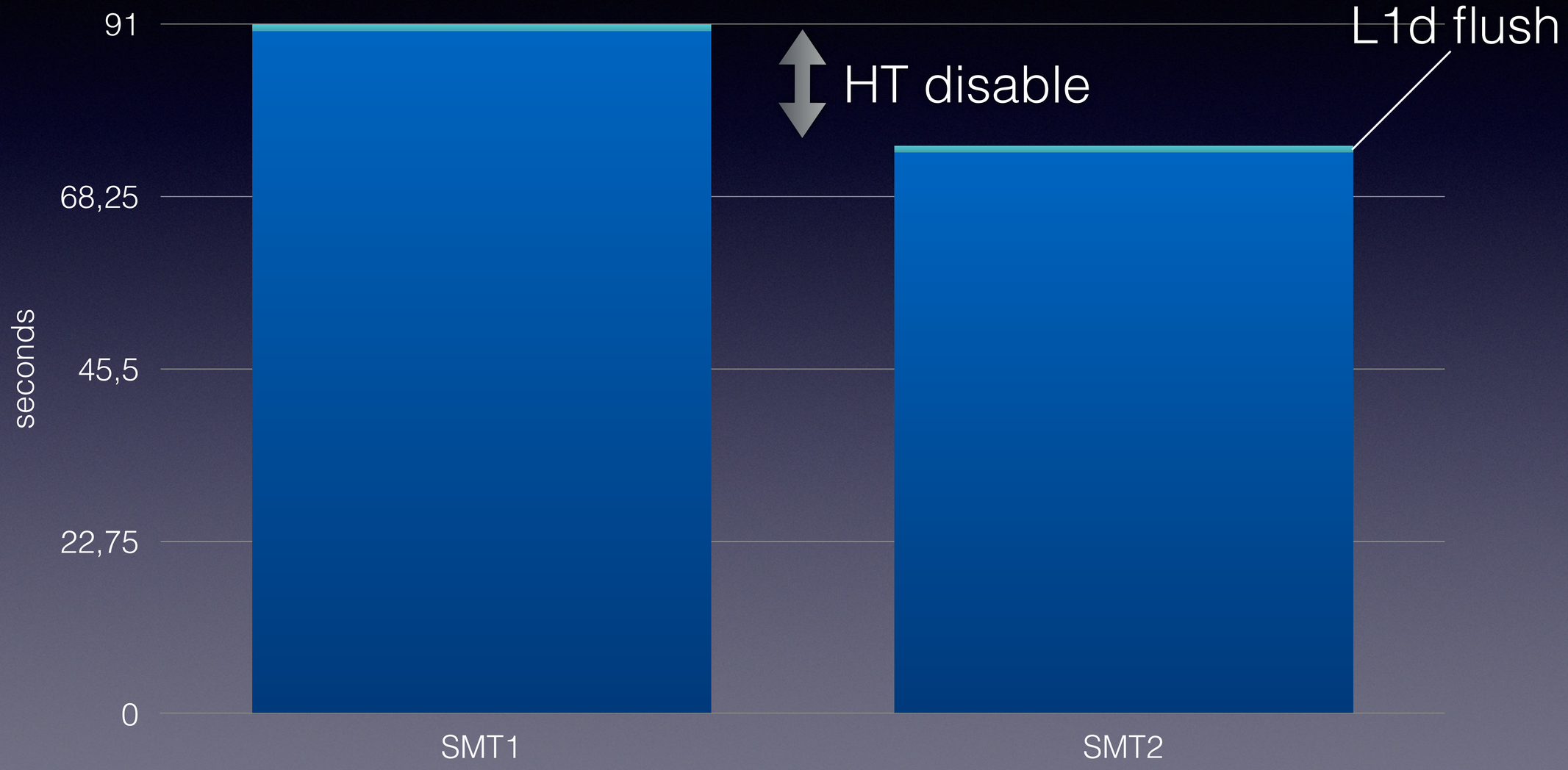
Kernel build



Kernel build

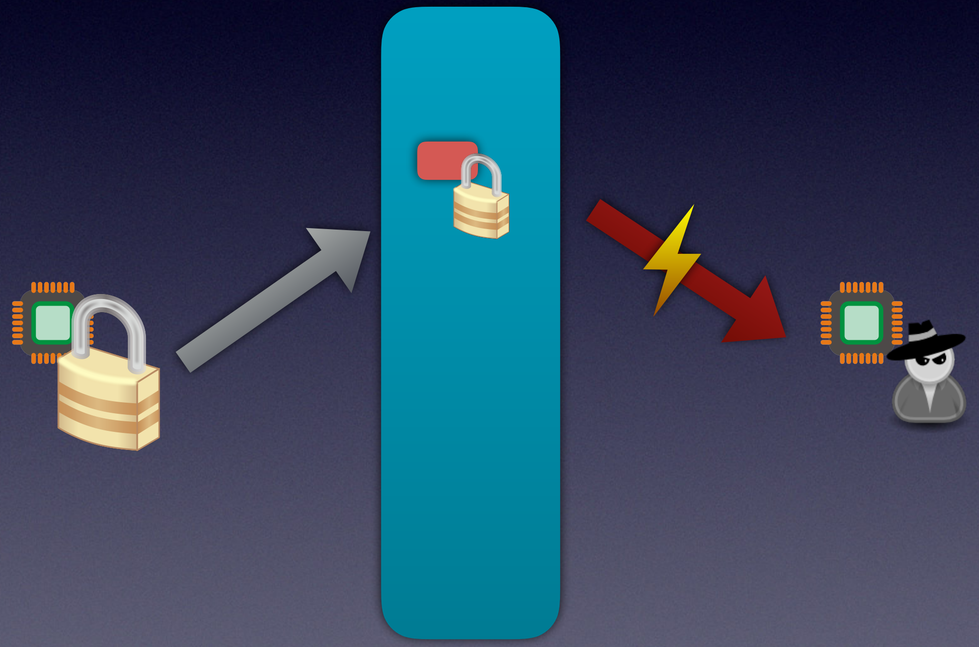


Kernel build

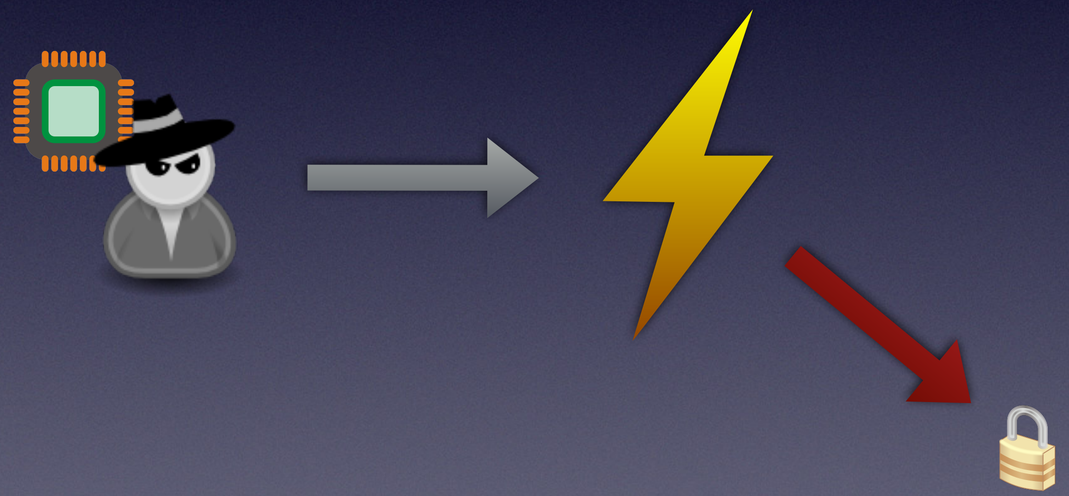


Alternative Mitigation

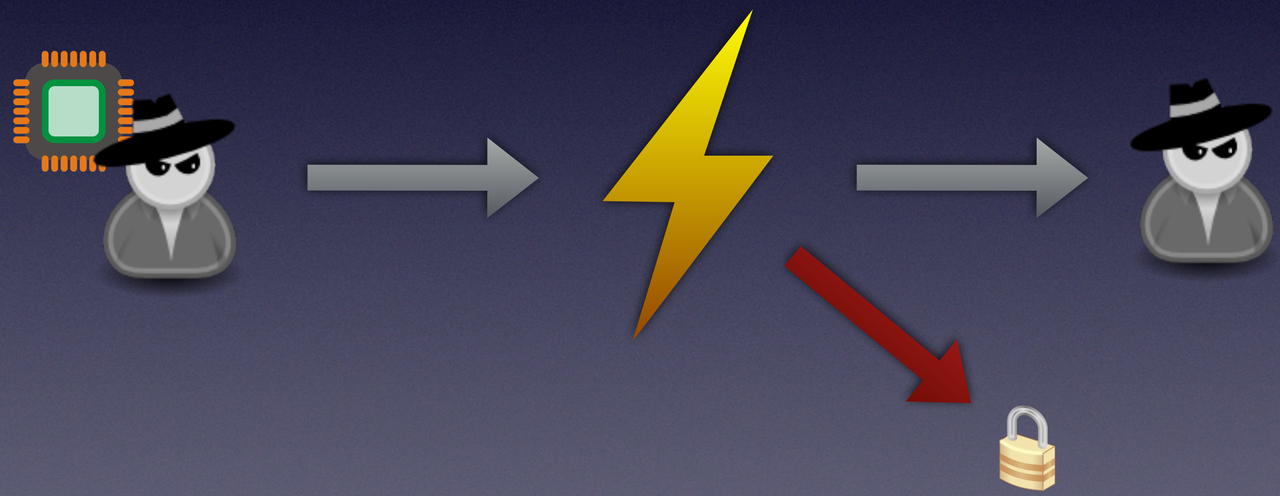
Alternative Mitigation



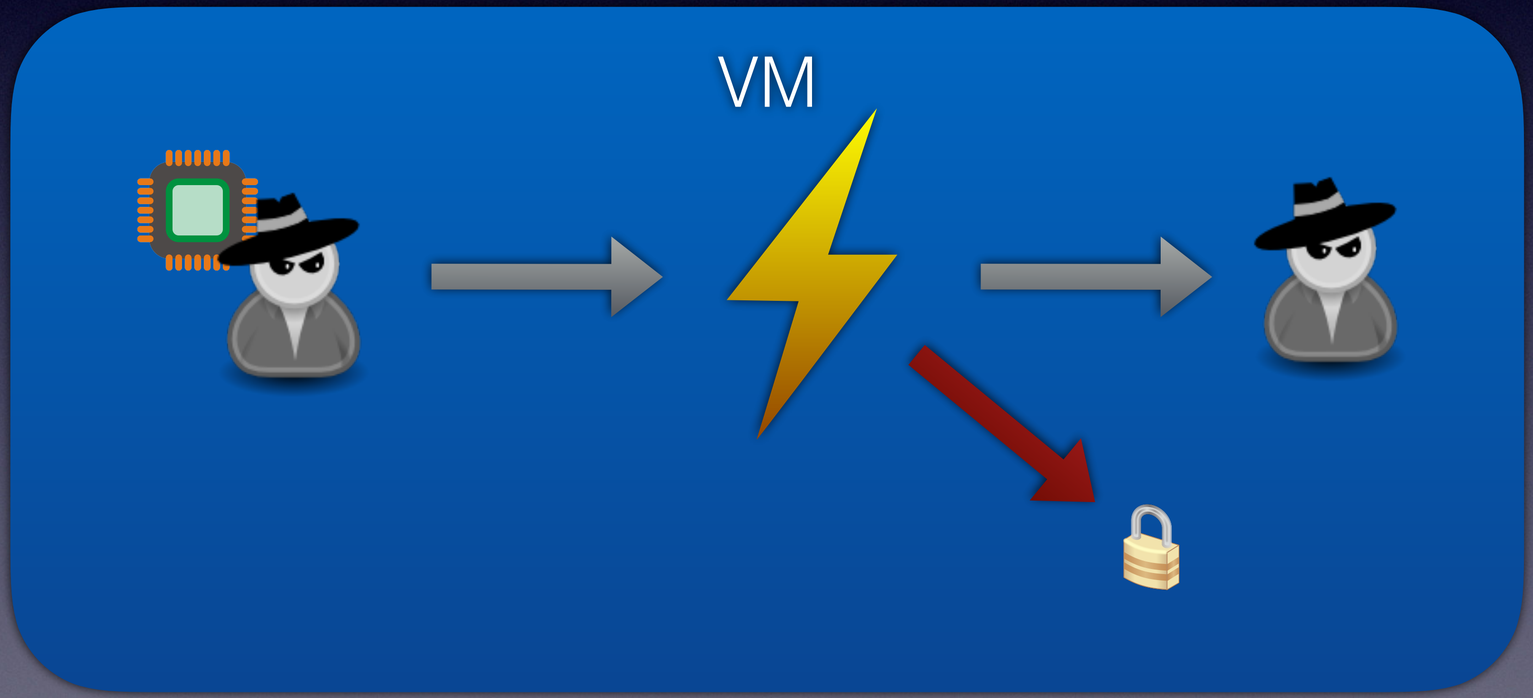
Alternative Mitigation



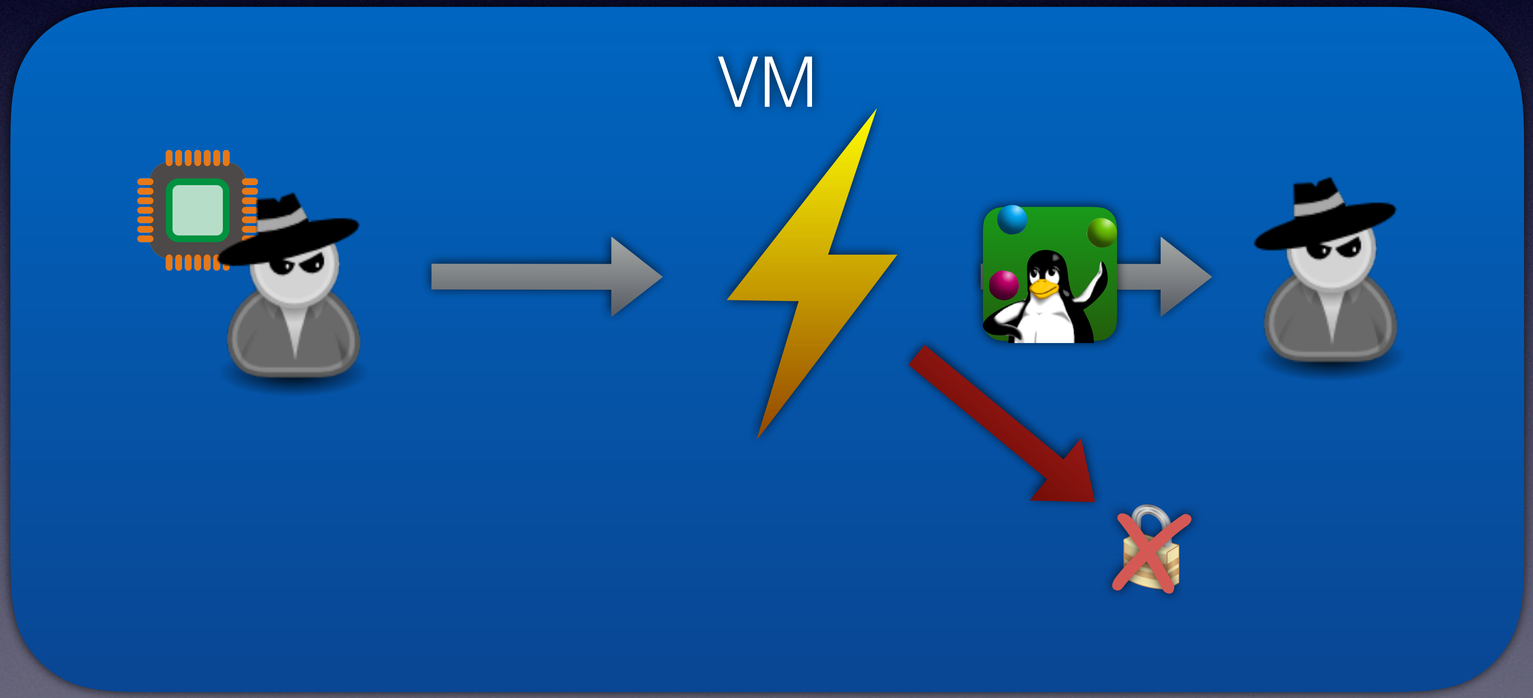
Alternative Mitigation



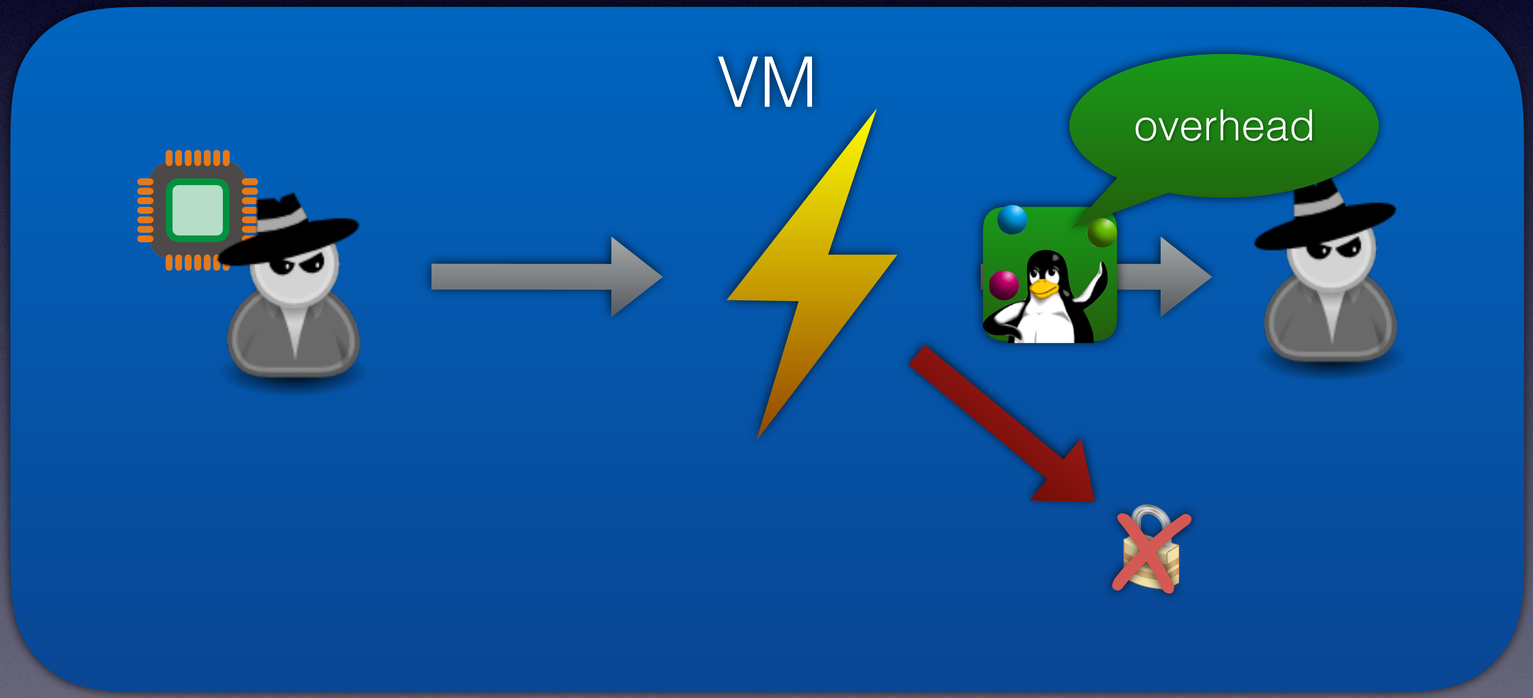
Alternative Mitigation



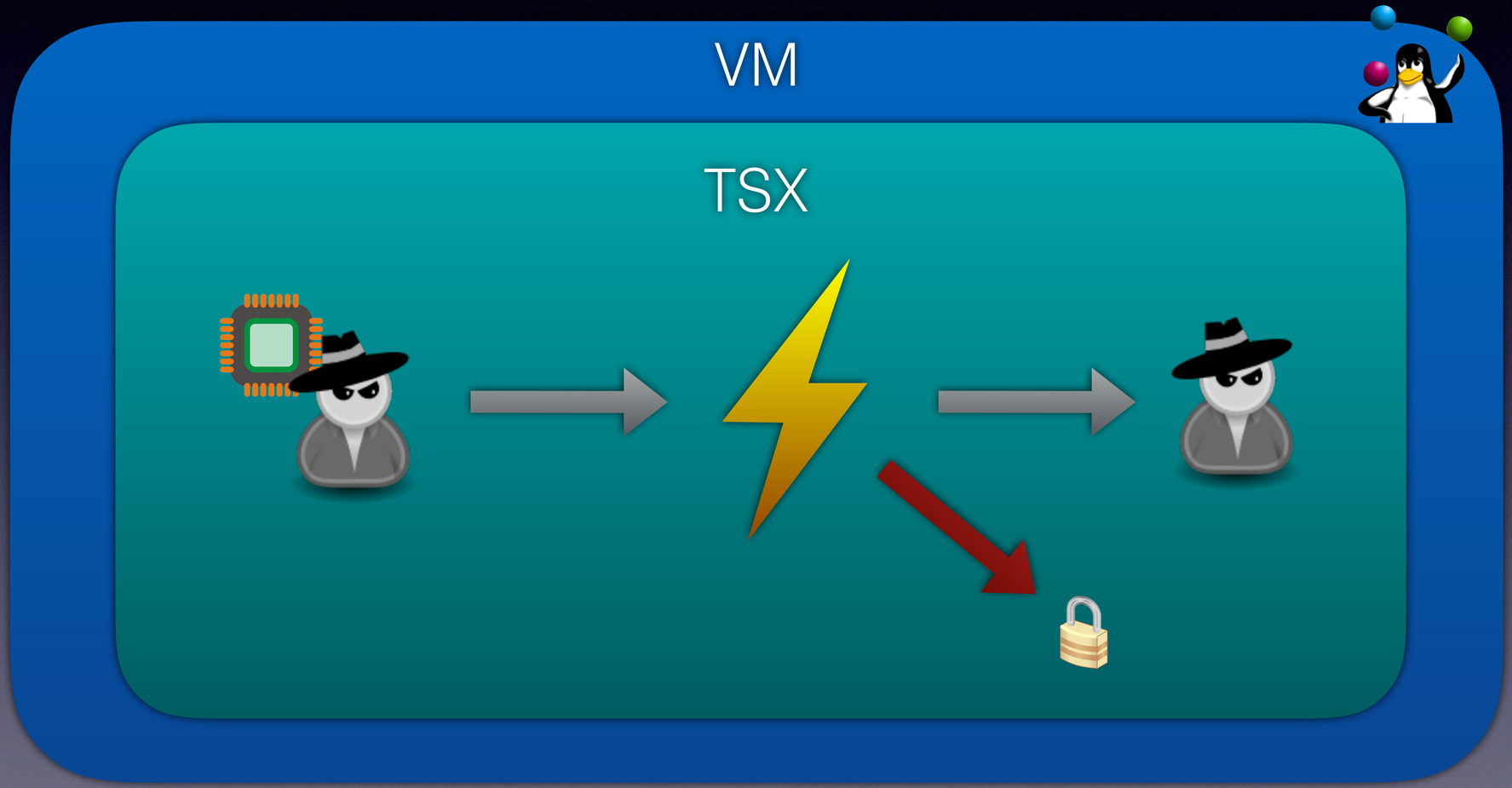
Alternative Mitigation

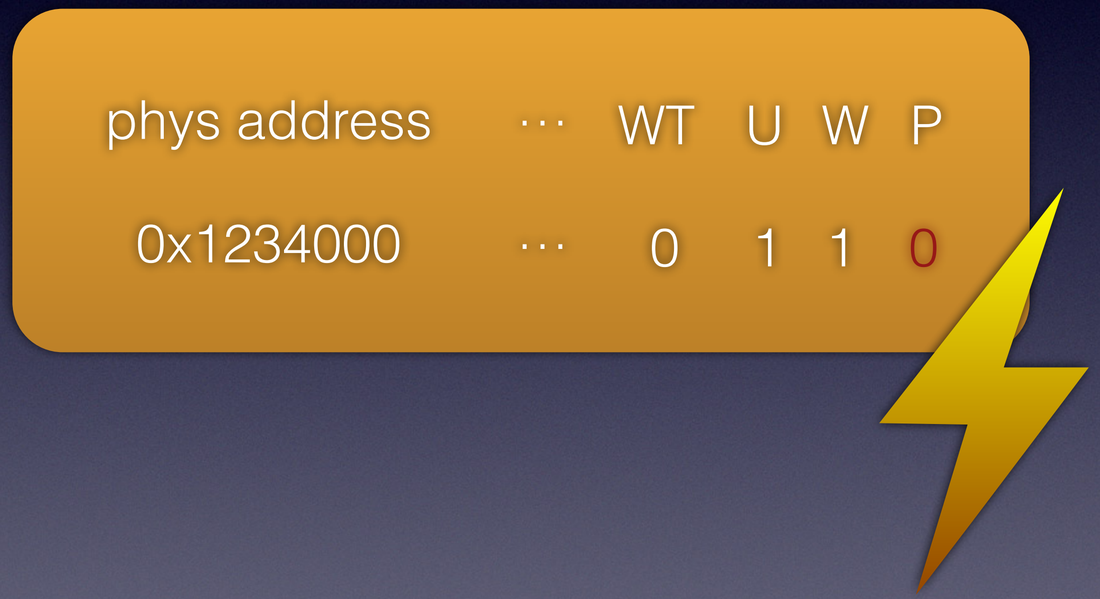


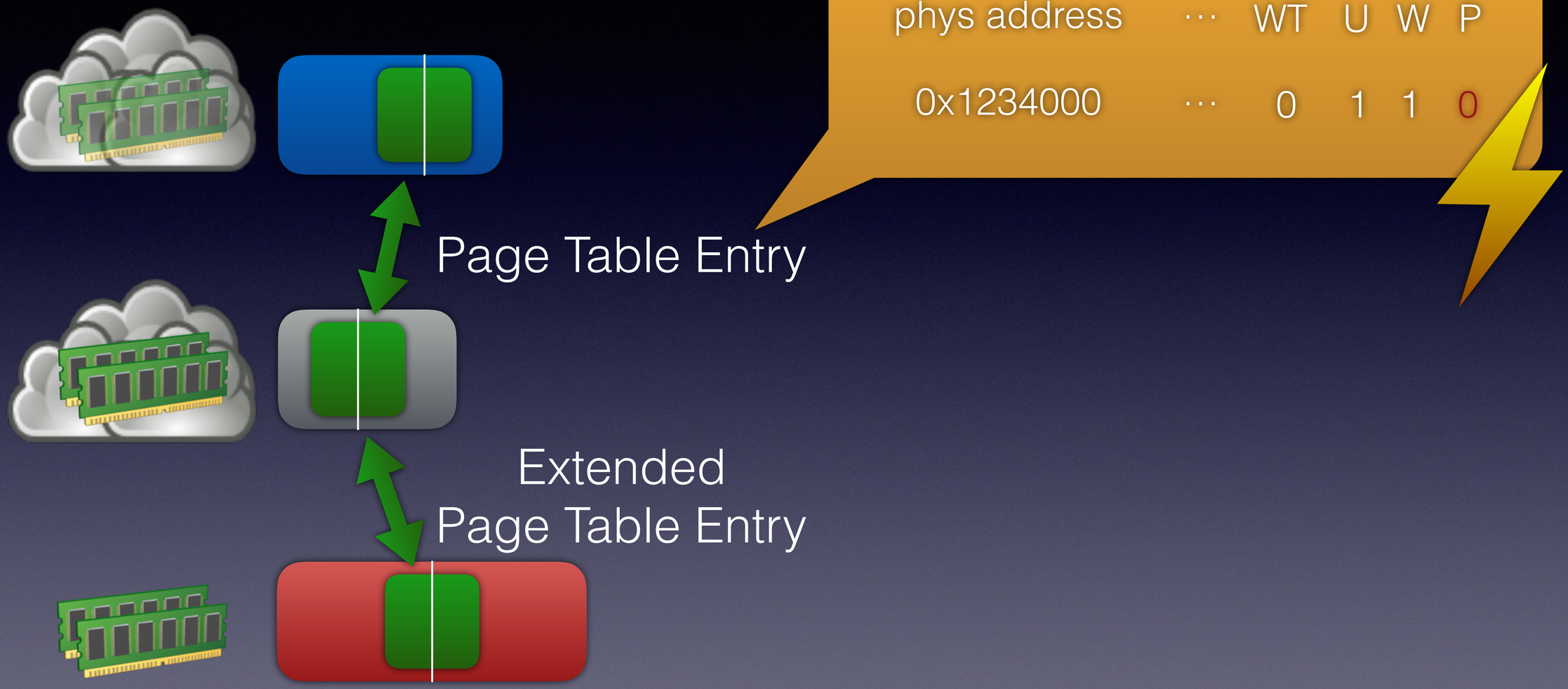
Alternative Mitigation

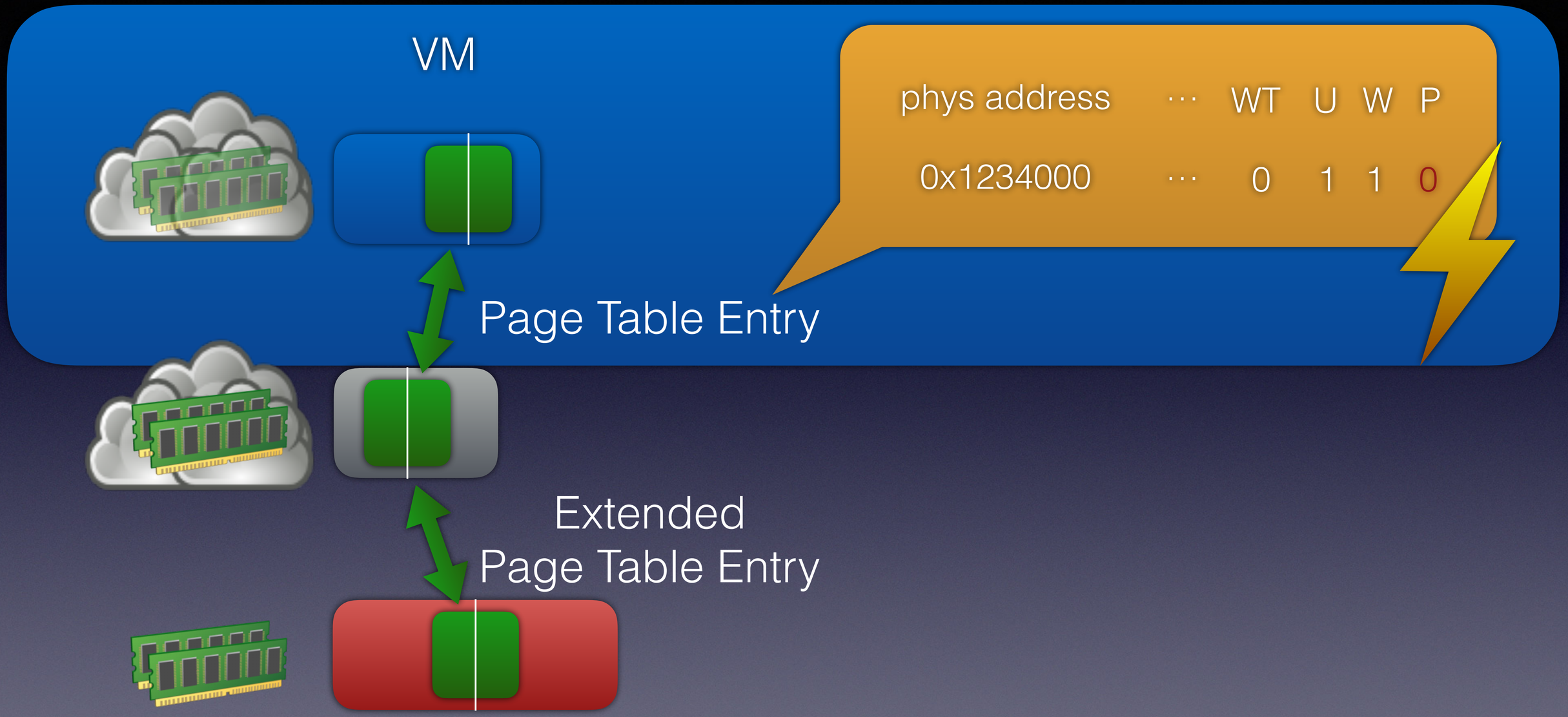


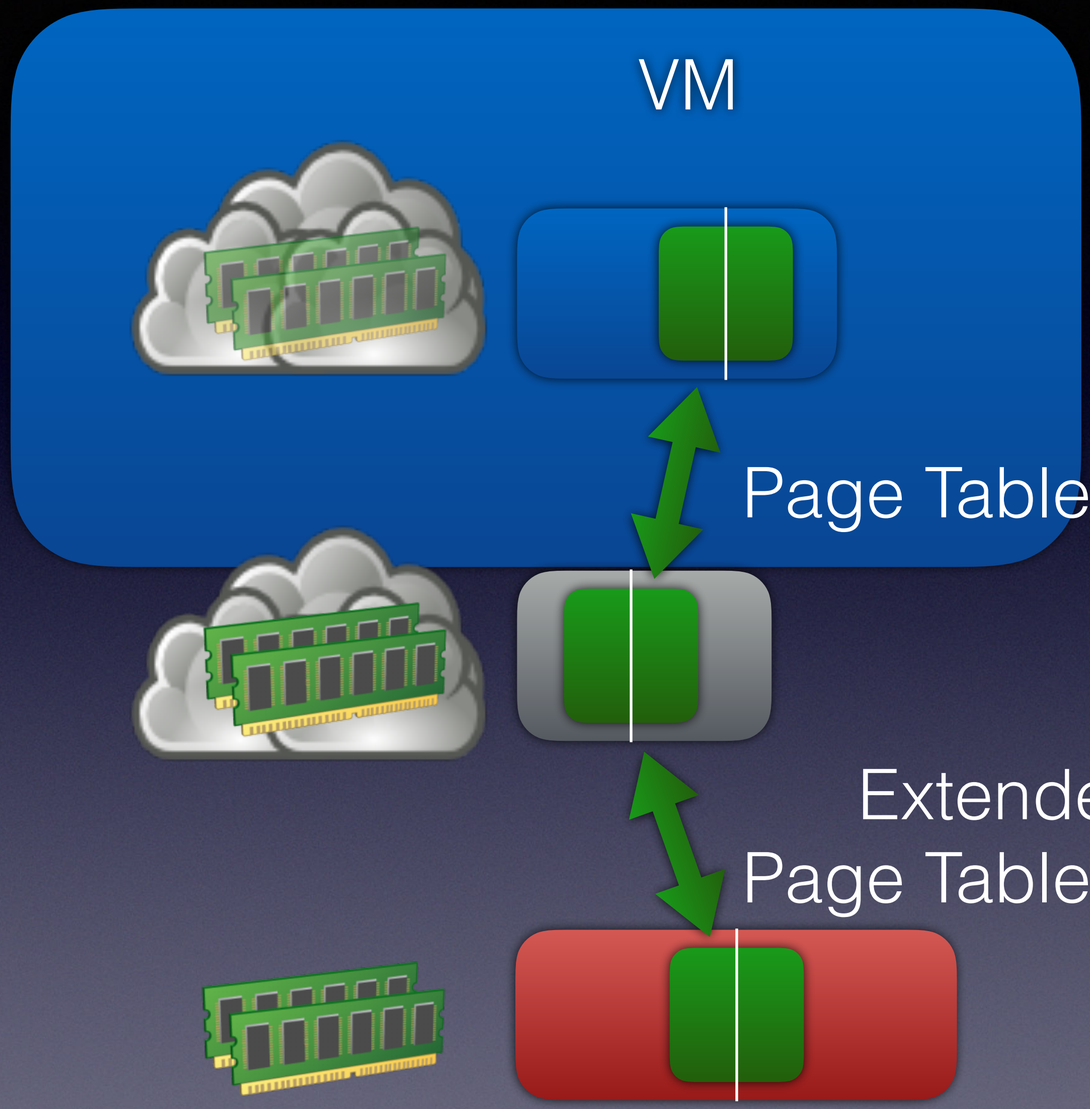
Alternative Mitigation

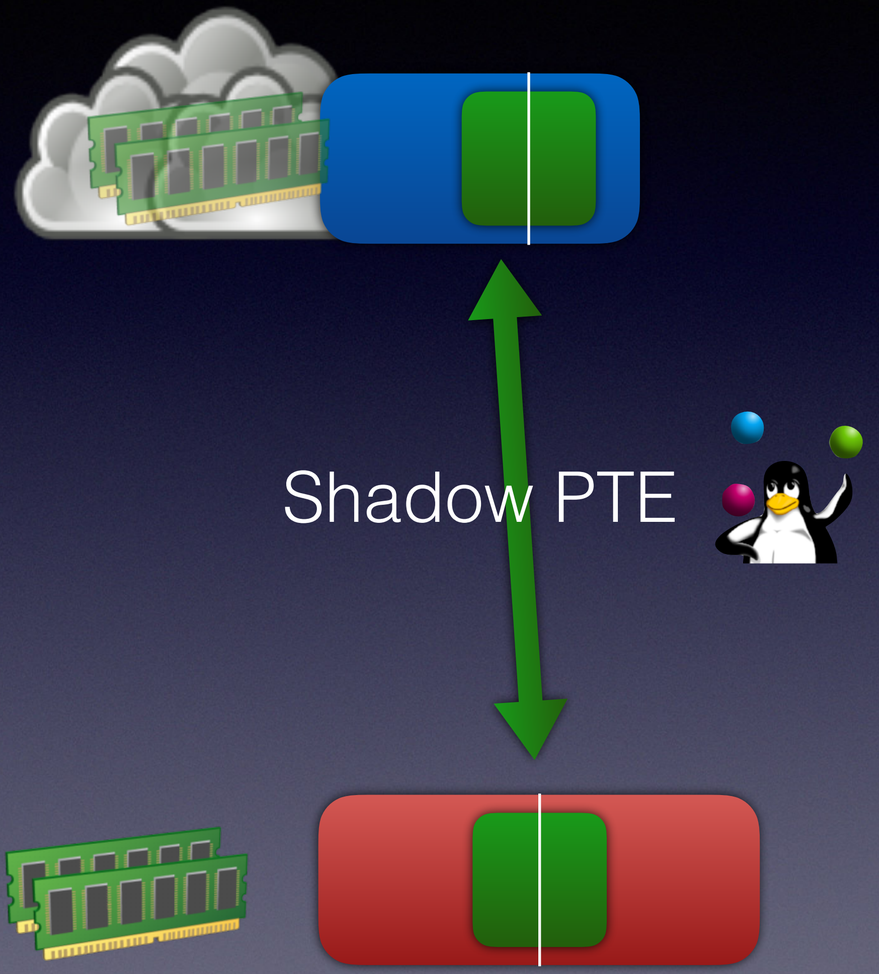


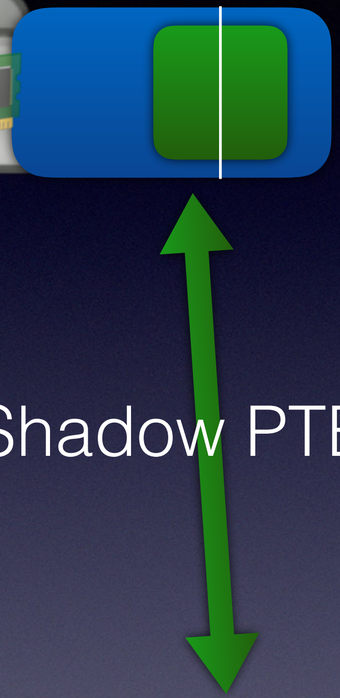

phys address ⋯ WT U W P
0x1234000 ⋯ 0 1 1 0

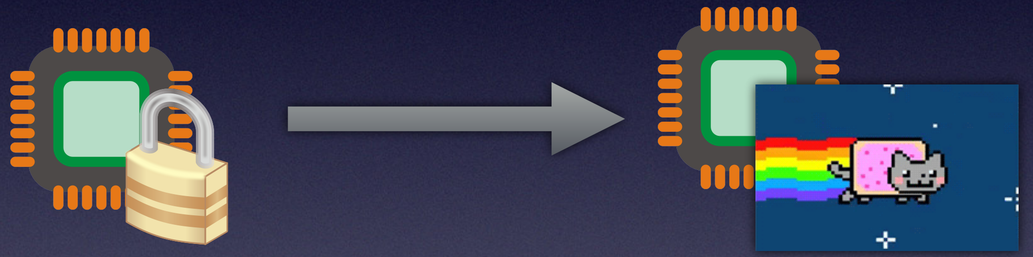

phys address ··· WT U W P
0x1234000 ··· 0 1 1 0
Page Table Entry
Extended
Page Table Entry

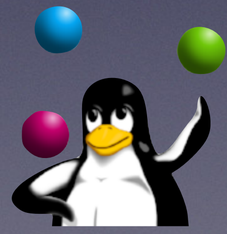

VM
phys address ⋯ WT U W P
0x1234000 ⋯ 0 1 1 0
Page Table Entry
Extended
Page Table Entry

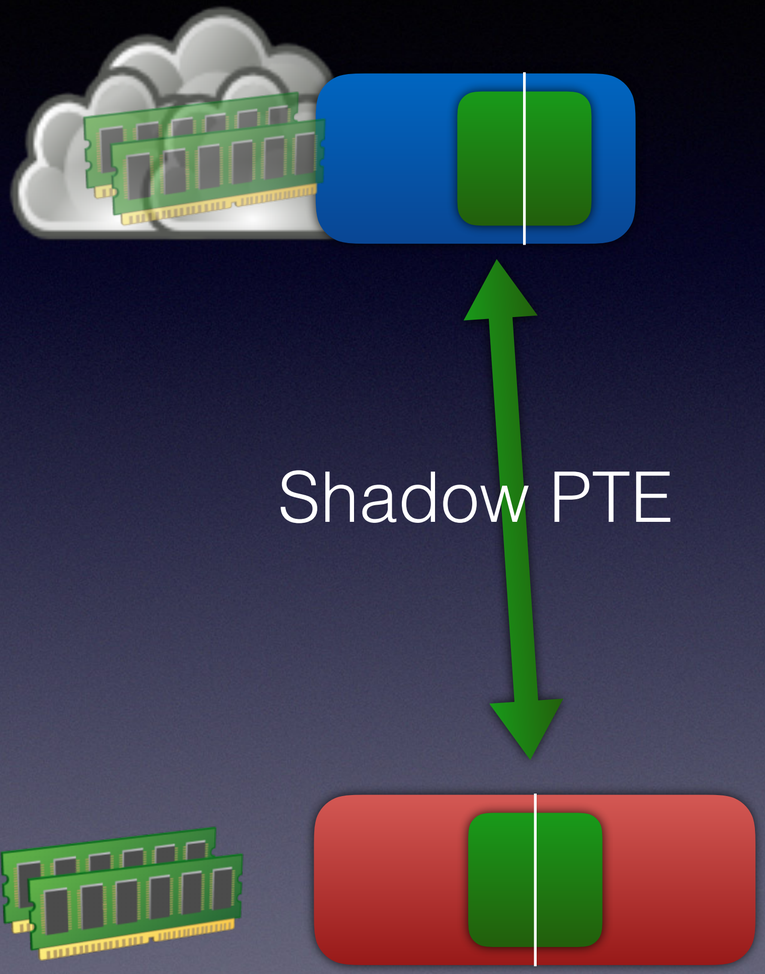

VM
Page Table Entry
Extended
Page Table Entry

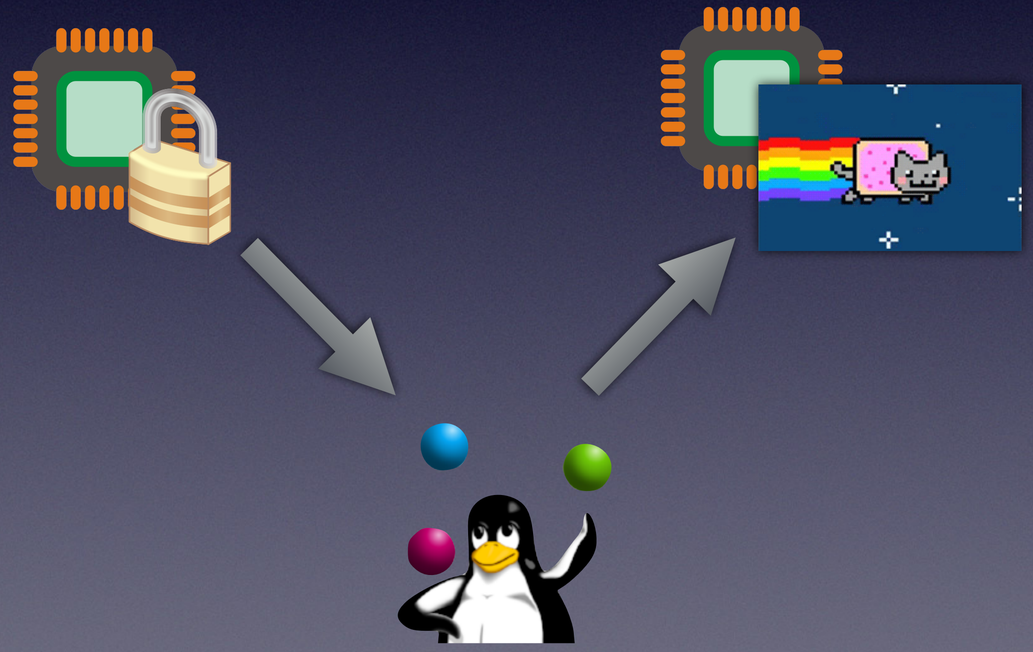

Shadow PTE

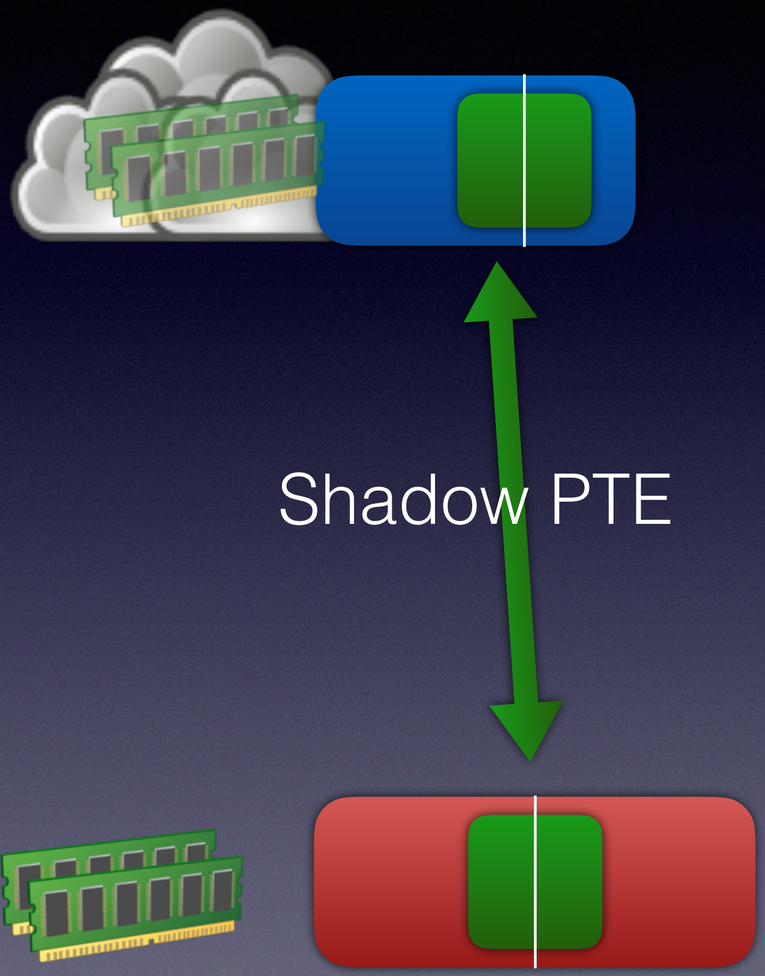

Shadow PTE

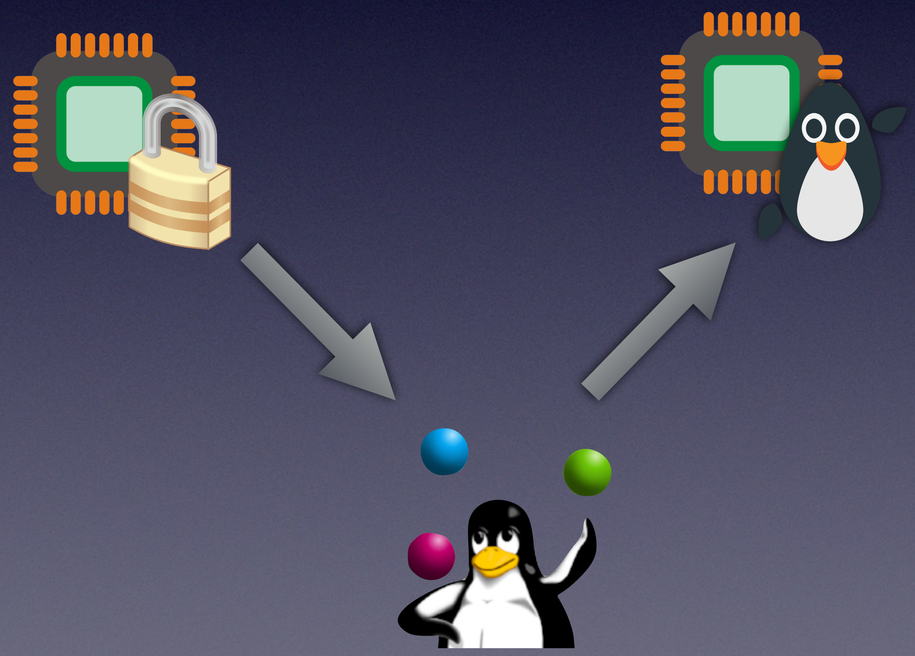




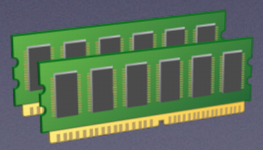

Shadow PTE

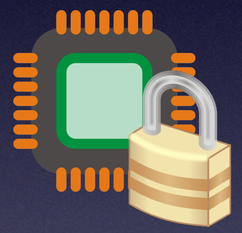



Shadow PTE

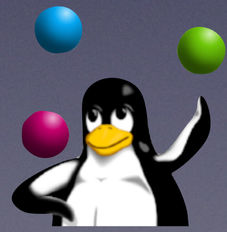


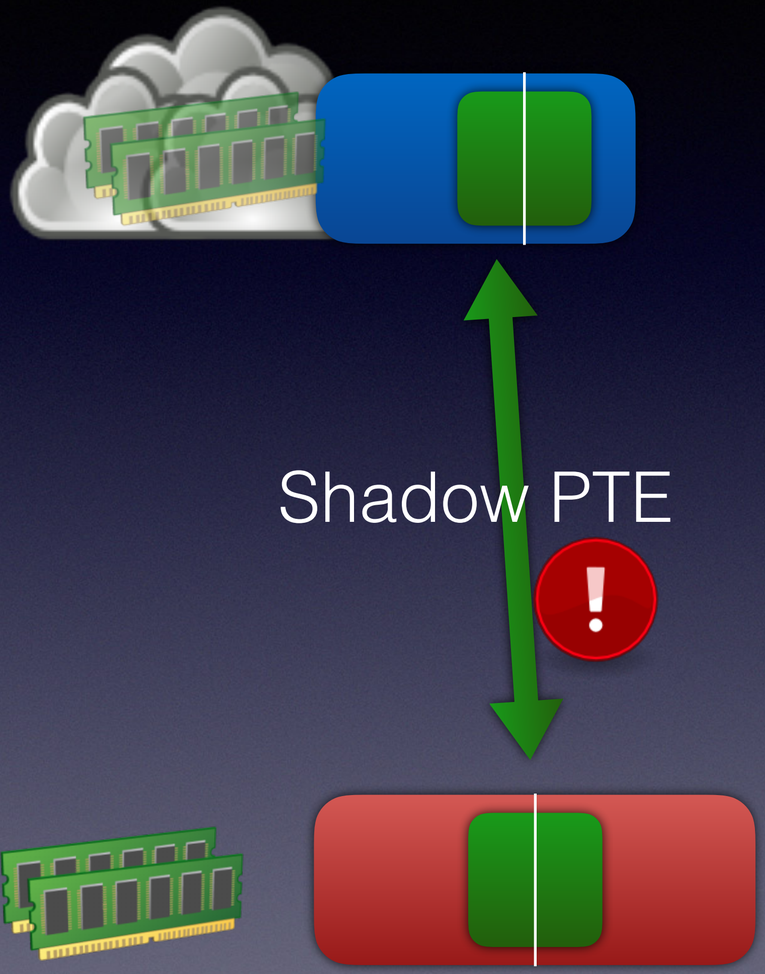
Shadow PTE

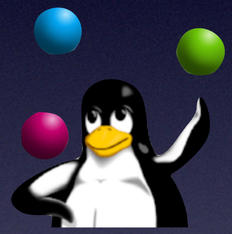


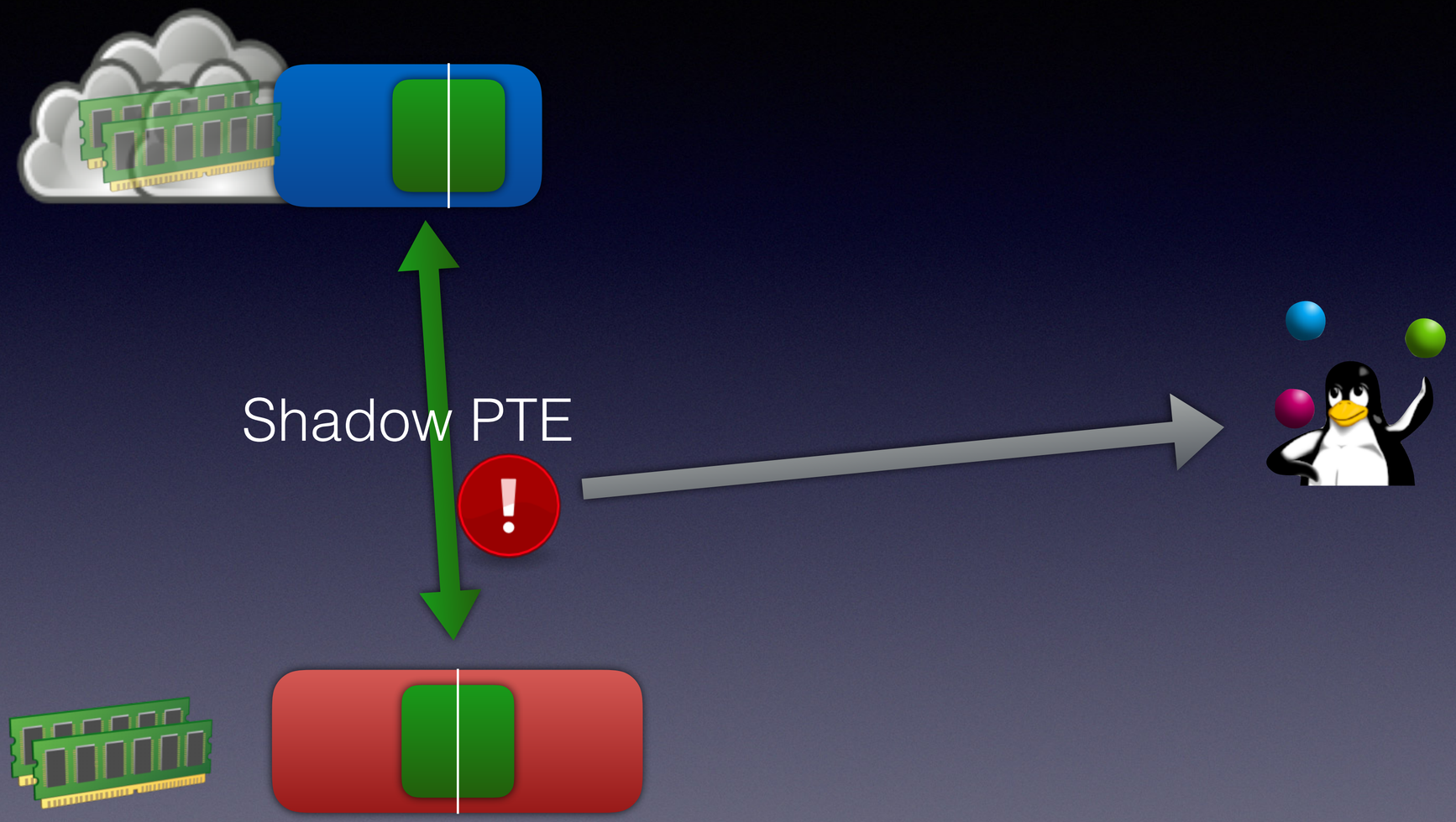


CR3
whitelist

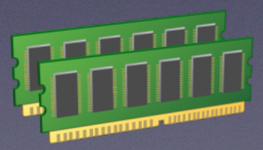



Shadow PTE

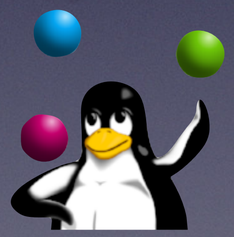



Shadow PTE
!

Shadow PTE



PML



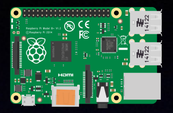
Alternative Mitigation

- Core Scheduling
- Hide host secrets

Demo

Thank You

External Sources

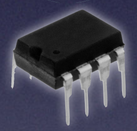


https://commons.wikimedia.org/wiki/File:Raspberry_Pi_B%2B_illustration.svg

https://commons.wikimedia.org/wiki/File:Sd-card-1377140.svg

http://eu.mophie.com/shop/media/catalog/product/cache/3/small_image/270x330/9df78eab33525d08d6e5fb8d27136e95/u/s/usb-micro3-40-blk_usb-tip-detail_front-back_540px.jpg

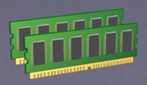
https://commons.wikimedia.org/wiki/File:Circle-icons-submarine.svg



https://commons.wikimedia.org/wiki/File:150-8-DIP.jpg

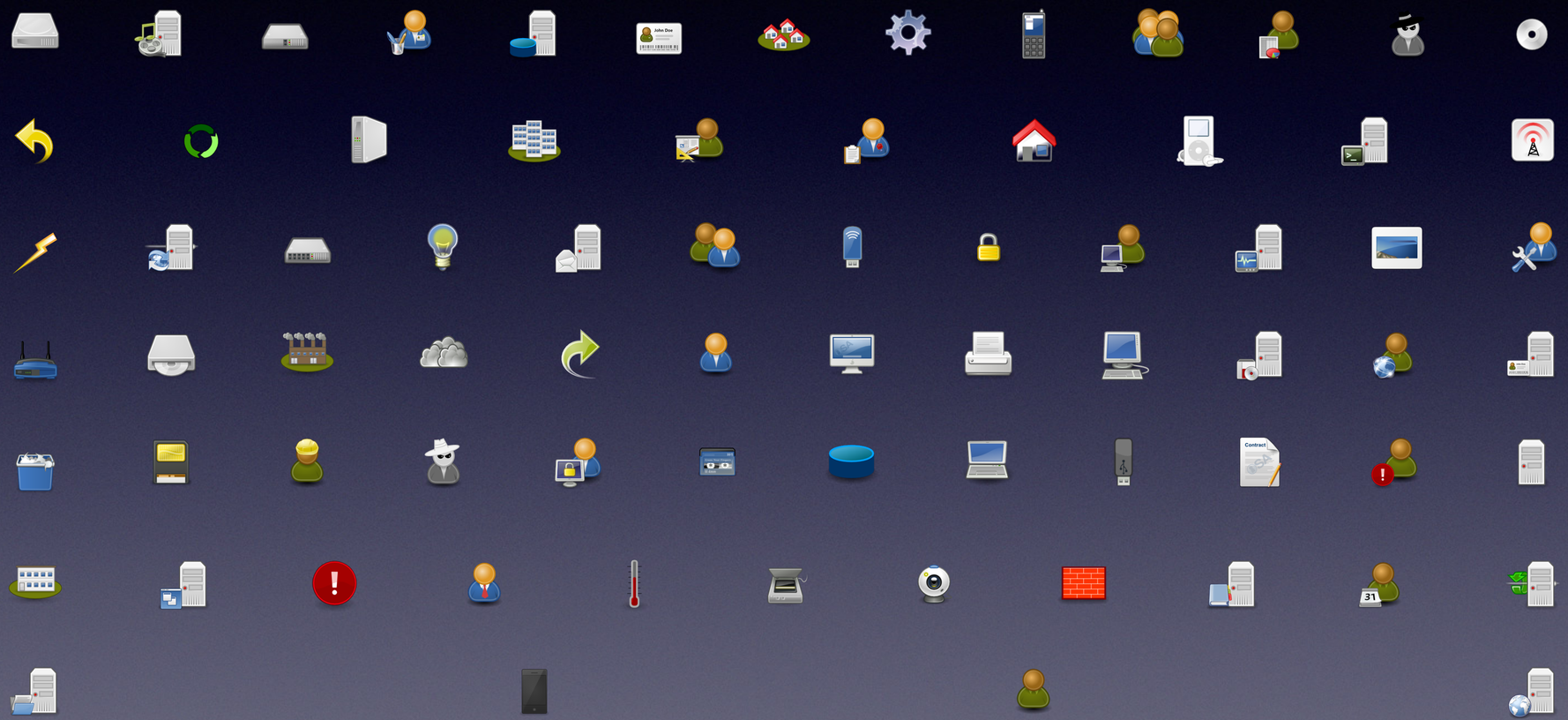
https://commons.wikimedia.org/wiki/File:Hdd_icon.svg

https://commons.wikimedia.org/wiki/File:ARM_CPU_icon.svg



http://findicons.com/icon/177982/memory#

OSA Icons



Icons received from http://www.opensecurityarchitecture.org/cms/library/icon-library

emojione Icons



Icons received from http://www.opensecurityarchitecture.org/cms/library/icon-library

Other Icons

http://findicons.com/icon/202613/folder_library

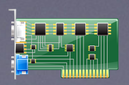
http://findicons.com/icon/download/234261/clock/128/png

http://findicons.com/icon/439269/button_power

https://fosdem.org/2017/schedule/event/grub_new_maintainers/attachments/slides/1768/export/events/attachments/grub_new_maintainers/slides/1768/slides.pdf

https://de.wikipedia.org/wiki/BeagleBoard#/media/File:Beagle_Board_big.jpg

https://thenounproject.com/term/folder-tree/27307/



https://commons.wikimedia.org/wiki/File:Crystal_Project_Hardware.png

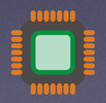
Other Icons

http://tumboy.tumblr.com/post/10052361836

http://findicons.com/icon/132807/b_leg_embossed

http://findicons.com/icon/237892/text_plain

http://findicons.com/icon/226957/package_games_board



https://pixabay.com/en/cpu-processor-intel-amd-chip-152656/

https://commons.wikimedia.org/wiki/File:Crypto_stub.svg