4 An Easter Egg in PCI Express

by Jacob Torrey

Dear Pastor Laphroaig,

Please consider the following submission to your church newsletter. I hope you think it worthy of your holy parishioners and readers.

Our friends at Intel are always providing Easter eggs for us to enjoy, and having stumbled across a new one for x86, the most neighborly option was naturally to share with all interested parties. This PoC is a weird quirk in which a newer x86 feature-set breaks invariants/security guarantees from older version. Specifically, the newer PCI Express configuration space access mechanism breaks virtual memory. Virtual memory is orchestrated by the CR3 register (storing the physical *address* of the page tables) and the page tables themselves. An issue with kernel shell-code and live memory forensics is that unless the *virtual address* of the page tables is known, it is impossible to map them (or any other physical address for that matter) into virtual memory, resulting in a chicken-andegg problem. Luckily, most operating systems keep the page tables at a known virtual address (0xC0000000 on many Windows systems), but this Easter egg allows access to the page tables on any OS.

In kernel space, CR3 can be read, providing the physical address of the OS page tables; however, due to Intel's virtual memory protections, there is no way to create a recursive virtual mapping to that physical address. All that is needed to do so, is a way to write an arbitrary 32-bits (which will become a PDE mapping in the page tables) to a known physical location.



This is the crux of the issue, and the security of virtual memory depends on it. Luckily, with the advent of PCI Express, there is now the "Enhanced Configuration Access Mechanism" (ECAM), which shadows PCI configuration space registers into physical memory at an address kept in the PCIEXPBAR register (D0:FO offset: 0x60). This is typically enabled on all the systems the author has come across, but your mileage may vary. With this ECAM, changes made to the configuration space via the legacy port I/O mechanism (0xCF8/0xCFC) will be reflected in physical memory. Now all that is needed is a register in configuration space that is at least 32-bits wide and can be changed to an arbitrary value without impacting the system. Again, Intel is looking out for our church, and through their grace, they provide a "Scratchpad Data" register (D0:F0 offset: 0xDC) that has no semantic meaning, just a location for software to store data. Now we have the function ModifyPM() for physical memory. (This is for Windows 32-bit without PAE, running as driver code.)

12// Utilize the scratch pad register as our mini-PDE 14 mov ebx, cr3 and ebx, 0xFFC00000// This is going to hold our new PDE (The bits in CR3 with the least significant stuff removed) 16or ebx, 0x83 //P | RW | PS18mov dx, 0x0cf8 // Offset 0x37 (0xDC / 4) 20mov eax, 0x80000DC ${\rm out}\ dx\,,\ eax$ 22mov dx, 0x0CFC 24mov eax, ebx out dx, eax // Write our PDE 26// Determine where in physical memory we can find the PDE mov dx, $0 \times 0 \times 168$ 28mov eax, 0x80000060 30 out dx, eax 32mov dx, 0x0CFCin eax. dx 34mov MMIORange, eax // Save our value and BAM! 36 popad } 38**if** (VDEBUG) 40DbgPrint("MMIO Base Address: %x", MMIORange); 42return MMIORange; }

Once the scratchpad register is primed and ready, and the physical address of the ECAM is known, the next step is to treat the register as a PDE mapping in the OS page tables to add a recursive mapping at a known location.

```
/**
1
       Sets up a recursive mapping to the OS page directory
       I commented it very thoroughly because it's quite complex.
3
5
      Basically it:
      -> Saves the current (real) CR3 value
7
      -> Creates a new PDE to map in the (real) PDT
      -> Creates a virtual address using the (fake) PDE we inserted in ModifyPM
9
      -> Switches to the (fake) CR3 and utilizes the constructed virtual
          address to insert the new recursive mapping into the (real) PDT
11
      -> Switches the CR3 back and continues on smugly
13 ULONG recurMap()
   ł
      ULONG MMIORange = 0;
15
      ULONG PDEBase = 0;
17
      ULONG PDE offset = 0;
       // Sets up the (fake) PDE and
19
      MMIORange = ModifyPM();
      MMIORange &= 0 \times F0000000;
21
23
       if (VDEBUG)
           DbgPrint("Mapping PDT to itself");
25
       \_asm {
```

27cli 29pushad // Save the current CR3, seems like overkill, but it makes sense 31mov ebx, cr3 // A copy to use to construct our virtual address mov ecx, cr3 // Save a copy so we don't mess up things up too much 33 mov edx, MMIORange // Our new CR3 val 3537 // Setup our virtual address // Gets us our offset into stuff
// Reference the PDE offset of (0x37 << 22)</pre> and ebx , 0x003FFFFFor ebx, 0x0DC0000039 // EBX should now have our virtual address :) 41 // Tests to see if the PDE is free for use 43test_pde: 45add ebx, 0x4 // Offset to unused PDE 47// Keep the offset var up to date (but uint32 aligned, not uint8) mov eax, PDEoffset 49add eax, 0x1 mov PDEoffset, eax 5153mov cr3, edx // Inject our new CR3 55// case it could cause later problems. 5759mov cr3, ecx // Restore everything nicely //************ END CRITICAL SECTION cmp eax, 0 // Can we use this entry? 61je inject_pde // Try the next one jmp test_pde // Found an empty one, w00t! 63 // Injects our recursive PDE into the PDT 65 inject_pde: // Setup our recursive PDE (again) 67 mov eax, cr3 // A copy to modify for our new recursive PDE and eax, 0xFFC00000 // Only the most significant bits stay for 4M pages 69or eax, 0x93 // P / \widetilde{RW} / \widetilde{PS} / PCD 71// EAX now holds the same PDE to put into the 'real' PDT mov cr3, edx // Inject our new CR3 7375// case it could cause later problems 77 mov cr3, ecx // Restore everything nicely 79//*********** END CRITICAL SECTION 81 // Determine the virtual address of the base of the PDT 83 // (remembering the differences in alignment) mov eax, cr3 // A copy to modify for our new recursive PDE 85 and eax, 0x003FFFFF // Only the most significant bits stay for 4M pages 87 mov ebx, PDEoffset shl ebx, 22 // Offset into the PDT 89 ${\rm or}\ {\rm eax}\,,\ {\rm ebx}$ mov PDEoffset, eax 91

0.2	popad
95	$_{ m sti}$
95	}
07	
97	DbgPrint("Mapping complete should be mapped in at 0x%x!", PDEoffset):
99	
	return PDEoffset;
101	}

The above, on a 32-bit non-PAE system, will return the virtual address that maps in the page directory and allows you to map in arbitrary physical memory as a known location. It should be noted that kernel privileges are needed (to access CR3) and to operate on a kernel page marked as Global so as to persist through the CR3 changes. The author hopes you enjoyed this weird machine and remember to treat your input data as formally as code, for only you can prevent vulnerabilities!

Sincerely, @JacobTorrey

