Following the Linux Kernel Defence Map

Alexander Popov

Positive Technologies

About Me

- Alexander Popov
- Linux kernel developer since 2013
- Security researcher at POSITIVE TECHNOLOGIES
- Focused on
  - Linux kernel vulnerability discovery
  - Exploitation techniques
  - Defensive technologies

Agenda

1. Linux Kernel Defence Map
2. kconfig-hardened-check tool

source: http://moscowmarathon.org/

Linux Kernel Security

Linux kernel security is a complex area, there are:

- Vulnerability classes
- Exploitation techniques
- Bug detection mechanisms
- Defence technologies
  - Mainline
  - Out-of-tree
  - Commercial
  - Provided by hardware

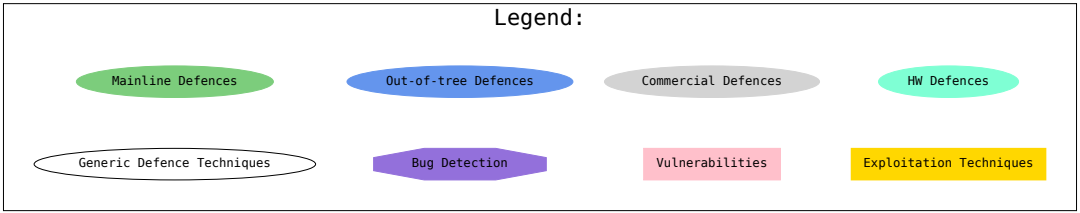
They all have complex relationships

We need a map for easier navigation

Drawn by Daniel Reeve, made by weta

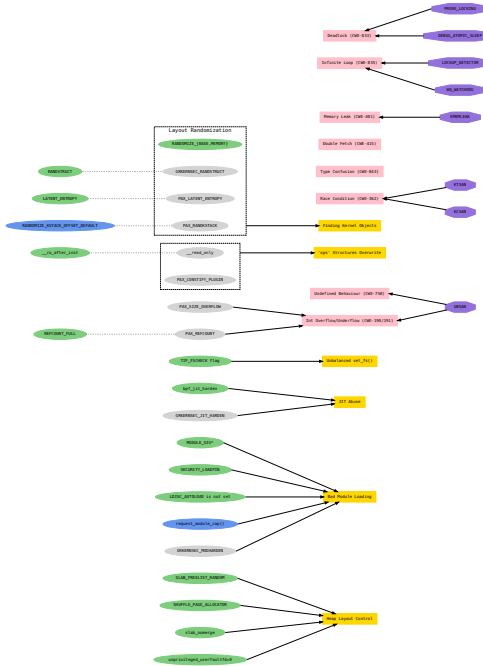
Linux Kernel Defence Map

- So I created a Linux Kernel Defence Map
  https://github.com/a13xp0p0v/linux-kernel-defence-map
- Started to work on it in 2018, still improving and updating
- Key concepts:



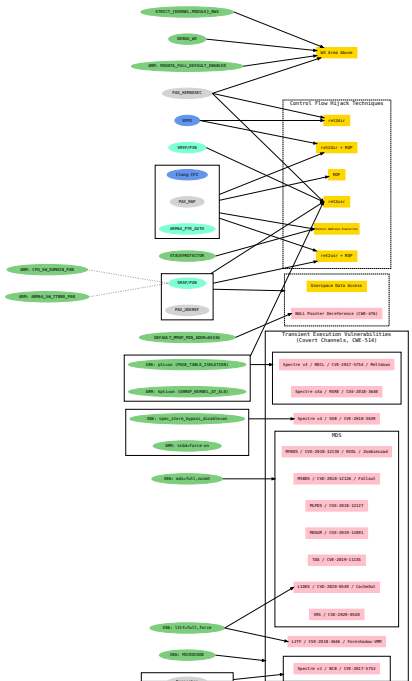
- Each connection between nodes represents a relationship
- (!) This map doesn't cover cutting attack surface

Linux Kernel Defence Map
whole picture (1/4)



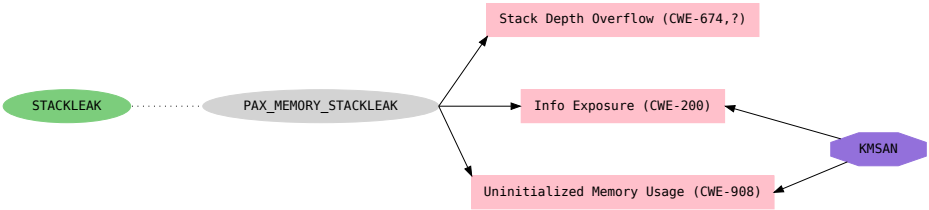
Linux Kernel Defence Map
whole picture (2/4)

Linux Kernel Defence Map
whole picture (3/4)

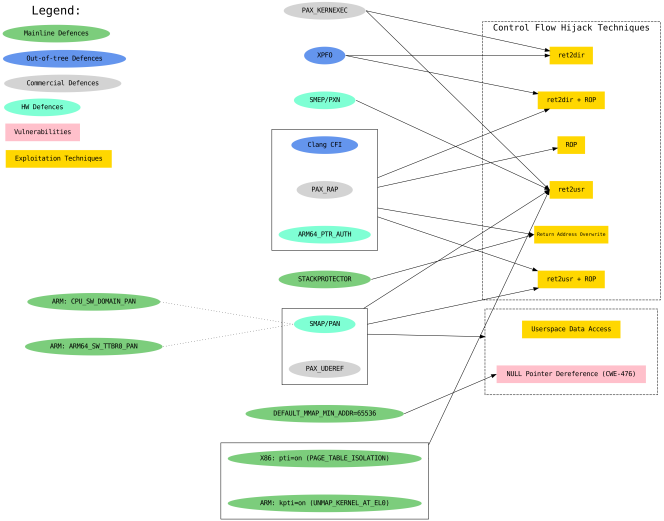


Linux Kernel Defence Map
whole picture (4/4)

Examples from the Map: STACKLEAK



Examples from the Map: Control Flow Hijack



Map Implementation

- Requirements:
  - I need to update the Map (at least every kernel release)
  - I **want** to develop it in text and use VCS
  - I **don't want** to place nodes and edges manually
- So I chose the DOT language provided by the Graphviz package
- Command to generate SVG file:

```
dot -Tsvg map.dot -o map.svg
```

Map Code Example

```
// Defences relations
edge [style=dotted, arrowhead=none, dir=none, headport=_, tailport=_];
"STACKLEAK":e -> "PAX_MEMORY_STACKLEAK":w;
// Bug Detection Mechanisms vs Vulnerabilities
edge [style=solid, arrowhead=normal, dir=back, headport=_, tailport=_];
"Uninitialized Memory Usage (CWE-908)":e -> "KMSAN";
"Info Exposure (CWE-200)":e -> "KMSAN";
// Defences vs Vulnerabilities and Exploitation Techniques
edge [style=solid, arrowhead=normal, dir=forward, headport=_, tailport=_];
"PAX_MEMORY_STACKLEAK":e -> "Stack Depth Overflow (CWE-674,?)":sw;
"PAX_MEMORY_STACKLEAK":e -> "Uninitialized Memory Usage (CWE-908)":nw;
"PAX_MEMORY_STACKLEAK":e -> "Info Exposure (CWE-200)":w;
```

Linux Kernel Defence Map: Knowledge Sources

- grsecurity features
- Linux kernel security documentation
- Kernel Self Protection Project recommended settings
- Linux kernel mitigation checklist by Shawn C
- Trends, challenge, and shifts in software vulnerability mitigation by MSRC
- And more at https://github.com/a13xp0p0v/linux-kernel-defence-map/blob/master/README.md

Nice Map!

kconfig-hardened-check

- There are **plenty** of Linux kernel hardening config options
- A lot of them are **not** enabled by the major distros
- **Nobody** likes checking configs manually
- **So let the computers do their job!**
- I created **kconfig-hardened-check** for checking security-related options in the Linux kernel Kconfig option list

  https://github.com/a13xp0p0v/kconfig-hardened-check
- Started to work on it in 2018, still improving and updating

kconfig-hardened-check: Output Example (1/4)

```
[a13x@hackbase kconfig-hardened-check]$ ./bin/kconfig-hardened-check -c kconfig_hardened_check/config_files/distros/ubuntu-focal.config
[+] Config file to check: kconfig_hardened_check/config_files/distros/ubuntu-focal.config
[+] Detected architecture: X86_64
[+] Detected kernel version: 5.4
=============================================================================================================
                option name                 | desired val | decision |       reason       |   check result
=============================================================================================================
CONFIG_BUG                                  |      y      |defconfig |  self_protection   |   OK
CONFIG_SLUB_DEBUG                           |      y      |defconfig |  self_protection   |   OK
CONFIG_GCC_PLUGINS                          |      y      |defconfig |  self_protection   |   FAIL: not found
CONFIG_STACKPROTECTOR_STRONG                |      y      |defconfig |  self_protection   |   OK
CONFIG_STRICT_KERNEL_RWX                    |      y      |defconfig |  self_protection   |   OK
CONFIG_STRICT_MODULE_RWX                    |      y      |defconfig |  self_protection   |   OK
CONFIG_REFCOUNT_FULL                        |      y      |defconfig |  self_protection   |   FAIL: "is not set"
CONFIG_IOMMU_SUPPORT                        |      y      |defconfig |  self_protection   |   OK
CONFIG_MICROCODE                            |      y      |defconfig |  self_protection   |   OK
CONFIG_RETPOLINE                            |      y      |defconfig |  self_protection   |   OK
CONFIG_X86_SMAP                             |      y      |defconfig |  self_protection   |   OK
CONFIG_SYN_COOKIES                          |      y      |defconfig |  self_protection   |   OK
CONFIG_X86_UMIP                             |      y      |defconfig |  self_protection   |   OK: CONFIG_X86_INTEL_UMIP "y"
CONFIG_PAGE_TABLE_ISOLATION                 |      y      |defconfig |  self_protection   |   OK
CONFIG_RANDOMIZE_MEMORY                     |      y      |defconfig |  self_protection   |   OK
CONFIG_INTEL_IOMMU                          |      y      |defconfig |  self_protection   |   OK
CONFIG_AMD_IOMMU                            |      y      |defconfig |  self_protection   |   OK
CONFIG_VMAP_STACK                           |      y      |defconfig |  self_protection   |   OK
CONFIG_RANDOMIZE_BASE                       |      y      |defconfig |  self_protection   |   OK
CONFIG_THREAD_INFO_IN_TASK                  |      y      |defconfig |  self_protection   |   OK
CONFIG_BUG_ON_DATA_CORRUPTION               |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_DEBUG_WX                             |      y      |   kspp   |  self_protection   |   OK
CONFIG_SCHED_STACK_END_CHECK                |      y      |   kspp   |  self_protection   |   OK
CONFIG_SLAB_FREELIST_HARDENED               |      y      |   kspp   |  self_protection   |   OK
CONFIG_SLAB_FREELIST_RANDOM                 |      y      |   kspp   |  self_protection   |   OK
CONFIG_SHUFFLE_PAGE_ALLOCATOR               |      y      |   kspp   |  self_protection   |   OK
CONFIG_FORTIFY_SOURCE                       |      y      |   kspp   |  self_protection   |   OK
CONFIG_DEBUG_LIST                           |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_DEBUG_SG                             |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_DEBUG_CREDENTIALS                    |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_DEBUG_NOTIFIERS                      |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_INIT_ON_ALLOC_DEFAULT_ON             |      y      |   kspp   |  self_protection   |   OK
CONFIG_GCC_PLUGIN_LATENT_ENTROPY            |      y      |   kspp   |  self_protection   |   FAIL: not found
```

kconfig-hardened-check: Output Example (2/4)

```
CONFIG_GCC_PLUGIN_LATENT_ENTROPY          |      y      |   kspp   |  self_protection   |   FAIL: not found
CONFIG_GCC_PLUGIN_RANDSTRUCT              |      y      |   kspp   |  self_protection   |   FAIL: not found
CONFIG_HARDENED_USERCOPY                  |      y      |   kspp   |  self_protection   |   OK
CONFIG_HARDENED_USERCOPY_FALLBACK         | is not set  |   kspp   |  self_protection   |   FAIL: "y"
CONFIG_MODULE_SIG                         |      y      |   kspp   |  self_protection   |   OK
CONFIG_MODULE_SIG_ALL                     |      y      |   kspp   |  self_protection   |   OK
CONFIG_MODULE_SIG_SHA512                  |      y      |   kspp   |  self_protection   |   OK
CONFIG_MODULE_SIG_FORCE                   |      y      |   kspp   |  self_protection   |   FAIL: "is not set"
CONFIG_INIT_STACK_ALL                     |      y      |   kspp   |  self_protection   |   FAIL: not found
CONFIG_INIT_ON_FREE_DEFAULT_ON            |      y      |   kspp   |  self_protection   |   OK: CONFIG_PAGE_POISONING "y"
CONFIG_GCC_PLUGIN_STACKLEAK               |      y      |   kspp   |  self_protection   |   FAIL: not found
CONFIG_DEFAULT_MMAP_MIN_ADDR              |    65536    |   kspp   |  self_protection   |   OK
CONFIG_SECURITY_DMESG_RESTRICT            |      y      |  clipos  |  self_protection   |   FAIL: "is not set"
CONFIG_DEBUG_VIRTUAL                      |      y      |  clipos  |  self_protection   |   FAIL: "is not set"
CONFIG_STATIC_USERMODEHELPER              |      y      |  clipos  |  self_protection   |   FAIL: "is not set"
CONFIG_EFI_DISABLE_PCI_DMA                |      y      |  clipos  |  self_protection   |   FAIL: not found
CONFIG_SLAB_MERGE_DEFAULT                 | is not set  |  clipos  |  self_protection   |   FAIL: "y"
CONFIG_RANDOM_TRUST_BOOTLOADER            | is not set  |  clipos  |  self_protection   |   FAIL: "y"
CONFIG_RANDOM_TRUST_CPU                   | is not set  |  clipos  |  self_protection   |   FAIL: "y"
CONFIG_GCC_PLUGIN_RANDSTRUCT_PERFORMANCE  | is not set  |  clipos  |  self_protection   |   FAIL: CONFIG_GCC_PLUGIN_RANDSTRUCT not "y"
CONFIG_STACKLEAK_METRICS                  | is not set  |  clipos  |  self_protection   |   FAIL: CONFIG_GCC_PLUGIN_STACKLEAK not "y"
CONFIG_STACKLEAK_RUNTIME_DISABLE          | is not set  |  clipos  |  self_protection   |   FAIL: CONFIG_GCC_PLUGIN_STACKLEAK not "y"
CONFIG_INTEL_IOMMU_SVM                    |      y      |  clipos  |  self_protection   |   OK
CONFIG_INTEL_IOMMU_DEFAULT_ON             |      y      |  clipos  |  self_protection   |   FAIL: "is not set"
CONFIG_SLUB_DEBUG_ON                      |      y      |    my    |  self_protection   |   FAIL: "is not set"
CONFIG_RESET_ATTACK_MITIGATION            |      y      |    my    |  self_protection   |   OK
CONFIG_AMD_IOMMU_V2                       |      y      |    my    |  self_protection   |   FAIL: "m"
CONFIG_SECURITY                           |      y      |defconfig |  security_policy   |   OK
CONFIG_SECURITY_YAMA                      |      y      |   kspp   |  security_policy   |   OK
CONFIG_SECURITY_WRITABLE_HOOKS            | is not set  |    my    |  security_policy   |   OK: not found
CONFIG_SECURITY_LOCKDOWN_LSM              |      y      |  clipos  |  security_policy   |   OK
CONFIG_SECURITY_LOCKDOWN_LSM_EARLY        |      y      |  clipos  |  security_policy   |   OK
CONFIG_LOCK_DOWN_KERNEL_FORCE_CONFIDENTIALITY|      y      |  clipos  |  security_policy   |   FAIL: "is not set"
CONFIG_SECURITY_SAFESETID                 |      y      |    my    |  security_policy   |   OK
CONFIG_SECURITY_LOADPIN                   |      y      |    my    |  security_policy   |   FAIL: "is not set"
CONFIG_SECURITY_LOADPIN_ENFORCE           |      y      |    my    |  security_policy   |   FAIL: CONFIG_SECURITY_LOADPIN not "y"
CONFIG_SECCOMP                            |      y      |defconfig | cut_attack_surface |   OK
CONFIG_SECCOMP_FILTER                     |      y      |defconfig | cut_attack_surface |   OK
CONFIG_STRICT_DEVMEM                      |      y      |defconfig | cut_attack_surface |   OK
CONFIG_ACPI_CUSTOM_METHOD                 | is not set  |   kspp   | cut_attack_surface |   OK
```

kconfig-hardened-check: Output Example (3/4)

```
CONFIG_ACPI_CUSTOM_METHOD                | is not set  |   kspp   | cut_attack_surface |   OK
CONFIG_COMPAT_BRK                        | is not set  |   kspp   | cut_attack_surface |   OK
CONFIG_DEVKMEM                           | is not set  |   kspp   | cut_attack_surface |   OK
CONFIG_COMPAT_VDSO                       | is not set  |   kspp   | cut_attack_surface |   OK
CONFIG_BINFMT_MISC                       | is not set  |   kspp   | cut_attack_surface |   FAIL: "m"
CONFIG_INET_DIAG                         | is not set  |   kspp   | cut_attack_surface |   FAIL: "m"
CONFIG_KEXEC                             | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_PROC_KCORE                        | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_LEGACY_PTYS                       | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_HIBERNATION                       | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_IA32_EMULATION                    | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_X86_X32                           | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_MODIFY_LDT_SYSCALL                | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_OABI_COMPAT                       | is not set  |   kspp   | cut_attack_surface |   OK: not found
CONFIG_MODULES                           | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_DEVMEM                            | is not set  |   kspp   | cut_attack_surface |   FAIL: "y"
CONFIG_IO_STRICT_DEVMEM                  |      y      |   kspp   | cut_attack_surface |   FAIL: "is not set"
CONFIG_LEGACY_VSYSCALL_NONE              |      y      |   kspp   | cut_attack_surface |   FAIL: "is not set"
CONFIG_ZSMALLOC_STAT                     | is not set  |grsecurity| cut_attack_surface |   OK
CONFIG_PAGE_OWNER                        | is not set  |grsecurity| cut_attack_surface |   OK
CONFIG_DEBUG_KMEMLEAK                    | is not set  |grsecurity| cut_attack_surface |   OK
CONFIG_BINFMT_AOUT                       | is not set  |grsecurity| cut_attack_surface |   OK: not found
CONFIG_KPROBES                           | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_UPROBES                           | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_GENERIC_TRACER                    | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_PROC_VMCORE                       | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_PROC_PAGE_MONITOR                 | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_USELIB                            | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_CHECKPOINT_RESTORE                | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_USERFAULTFD                       | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_HWPOISON_INJECT                   | is not set  |grsecurity| cut_attack_surface |   FAIL: "m"
CONFIG_MEM_SOFT_DIRTY                    | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_DEVPORT                           | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_DEBUG_FS                          | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_NOTIFIER_ERROR_INJECTION          | is not set  |grsecurity| cut_attack_surface |   FAIL: "m"
CONFIG_X86_PTDUMP                        | is not set  |grsecurity| cut_attack_surface |   OK
CONFIG_DRM_LEGACY                        | is not set  |maintainer| cut_attack_surface |   OK
CONFIG_FB                                | is not set  |maintainer| cut_attack_surface |   FAIL: "y"
CONFIG_VT                                | is not set  |maintainer| cut_attack_surface |   FAIL: "y"
CONFIG_AIO                               | is not set  |grapheneos| cut_attack_surface |   FAIL: "y"
```

kconfig-hardened-check: Output Example (4/4)

```
CONFIG_MEM_SOFT_DIRTY                    | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_DEVPORT                           | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_DEBUG_FS                          | is not set  |grsecurity| cut_attack_surface |   FAIL: "y"
CONFIG_NOTIFIER_ERROR_INJECTION          | is not set  |grsecurity| cut_attack_surface |   FAIL: "m"
CONFIG_X86_PTDUMP                        | is not set  |grsecurity| cut_attack_surface |   OK
CONFIG_DRM_LEGACY                        | is not set  |maintainer| cut_attack_surface |   OK
CONFIG_FB                                | is not set  |maintainer| cut_attack_surface |   FAIL: "y"
CONFIG_VT                                | is not set  |maintainer| cut_attack_surface |   FAIL: "y"
CONFIG_AIO                               | is not set  |grapheneos| cut_attack_surface |   FAIL: "y"
CONFIG_STAGING                           | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_KSM                               | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_KALLSYMS                          | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_X86_VSYSCALL_EMULATION            | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_MAGIC_SYSRQ                       | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_KEXEC_FILE                        | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_USER_NS                           | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_X86_MSR                           | is not set  |  clipos  | cut_attack_surface |   FAIL: "m"
CONFIG_X86_CPUID                         | is not set  |  clipos  | cut_attack_surface |   FAIL: "m"
CONFIG_IO_URING                          | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_X86_IOPL_IOPERM                   | is not set  |  clipos  | cut_attack_surface |   OK: not found
CONFIG_LDISC_AUTOLOAD                    | is not set  |  clipos  | cut_attack_surface |   FAIL: "y"
CONFIG_X86_INTEL_TSX_MODE_OFF            |      y      |  clipos  | cut_attack_surface |   OK
CONFIG_ACPI_TABLE_UPGRADE                | is not set  | lockdown | cut_attack_surface |   FAIL: "y"
CONFIG_EFI_TEST                          | is not set  | lockdown | cut_attack_surface |   FAIL: "m"
CONFIG_BPF_SYSCALL                       | is not set  | lockdown | cut_attack_surface |   FAIL: "y"
CONFIG_MMIOTRACE_TEST                    | is not set  | lockdown | cut_attack_surface |   OK
CONFIG_MMIOTRACE                         | is not set  |    my    | cut_attack_surface |   FAIL: "y"
CONFIG_LIVEPATCH                         | is not set  |    my    | cut_attack_surface |   FAIL: "y"
CONFIG_IP_DCCP                           | is not set  |    my    | cut_attack_surface |   FAIL: "m"
CONFIG_IP_SCTP                           | is not set  |    my    | cut_attack_surface |   FAIL: "m"
CONFIG_FTRACE                            | is not set  |    my    | cut_attack_surface |   FAIL: "y"
CONFIG_BPF_JIT                           | is not set  |    my    | cut_attack_surface |   FAIL: "y"
CONFIG_VIDEO_VIVID                       | is not set  |    my    | cut_attack_surface |   FAIL: "m"
CONFIG_INPUT_EVBUG                       | is not set  |    my    | cut_attack_surface |   FAIL: "m"
CONFIG_INTEGRITY                         |      y      |defconfig |userspace_hardening |   OK
CONFIG_ARCH_MMAP_RND_BITS                |     32      |  clipos  |userspace_hardening |   FAIL: "28"

[+] Config check is finished: 'OK' - 57 / 'FAIL' - 81
[a13x@hackbase kconfig-hardened-check]$
```

kconfig-hardened-check: Knowledge Sources

- KSPP recommended settings
- CLIP OS kernel configuration
- Last public grsecurity patch (options which they disable)
- SECURITY_LOCKDOWN_LSM patchset
- Direct feedback from Linux kernel maintainers

kconfig-hardened-check: About the Project

- GPL-3.0 License
- In Python (please don't cry if my code looks like C code, I'm just a kernel developer)
- CI: automatic functional tests, code coverage 93%
- Distribution via pip/setuptools
- Nice contributors (kudos!)
- Is used by several Linux distributions (I'm glad!)

Conclusion (The Main Slide)

1. The **Linux Kernel Defence Map** helps to:
   - Get Linux Kernel security **overview**
   - Develop a **threat model** for your GNU/Linux system
   - Learn about kernel defences that **can help** against these threats
2. **kconfig-hardened-check** tool helps to control **security-related** options in your kernel config
3. Please **don't** change these options **without** knowing your **threat model**

Thanks!
Enjoy the Conference!

My contacts:
alex.popov@linux.com
https://a13xp0p0v.github.io/

Positive Technologies:
https://www.ptsecurity.com/