CSC 405
Computer Security

Linux Security

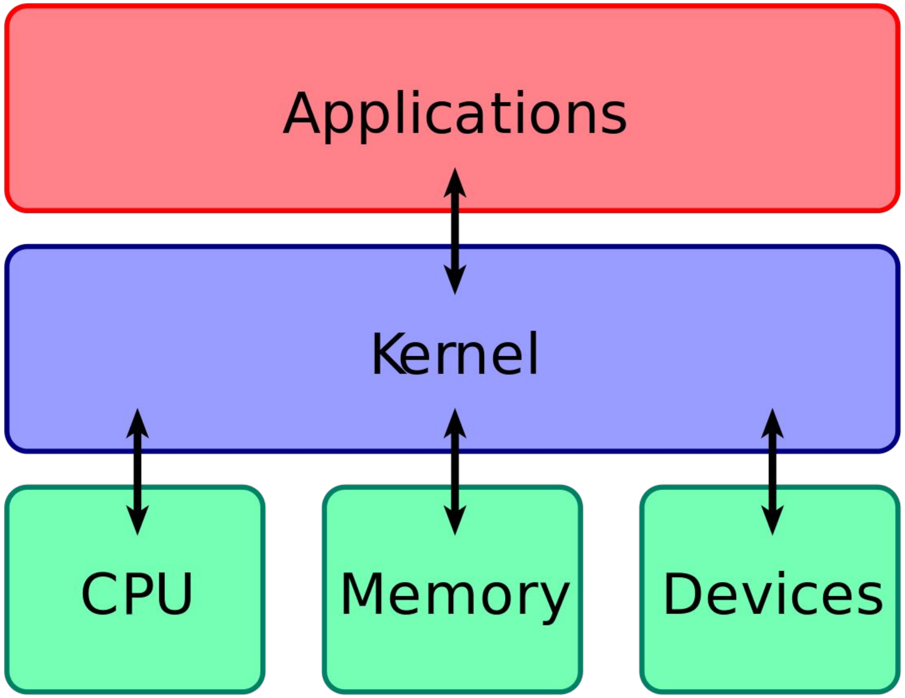
Alexandros Kapravelos
akaprav@ncsu.edu

Unix / Linux

- Started in 1969 at AT&T / Bell Labs
- Split into a number of popular branches
  - BSD, System V (commercial, AT&T), Solaris, HP-UX, AIX
- Inspired a number of Unix-like systems
  - Linux, Minix
- Standardization attempts
  - POSIX, Single Unix Specification (SUS), Filesystem Hierarchy Standard (FHS), Linux Standard Base (LSB), ELF

OS Security

- Kernel vulnerability
  - usually leads to complete system compromise
  - attacks performed via system calls



Kernel vulnerabilities

<table>
<tr><th>#</th><th>CVE ID</th><th>CWE ID</th><th># of Exploits</th><th>Vulnerability Type(s)</th><th>Publish Date</th><th>Update Date</th><th>Score</th><th>Gained Access Level</th><th>Access</th><th>Complexity</th><th>Authentication</th><th>Conf.</th><th>Integ.</th><th>Avail.</th></tr>
<tr><td>1</td><td>CVE-2017-12762</td><td>119</td><td></td><td>Overflow</td><td>2017-08-09</td><td>2017-08-25</td><td>10.0</td><td>None</td><td>Remote</td><td>Low</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">In /drivers/isdn/i4l/isdn_net.c: A user-controlled buffer is copied into a local buffer of constant size using strcpy without a length check which can cause a buffer overflow. This affects the Linux kernel 4.9-stable tree, 4.12-stable tree, 3.18-stable tree, and 4.4-stable tree.</td></tr>
<tr><td>2</td><td>CVE-2017-11176</td><td>416</td><td></td><td>DoS</td><td>2017-07-11</td><td>2017-08-07</td><td>10.0</td><td>None</td><td>Remote</td><td>Low</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">The mq_notify function in the Linux kernel through 4.11.9 does not set the sock pointer to NULL upon entry into the retry logic. During a user-space close of a Netlink socket, it allows attackers to cause a denial of service (use-after-free) or possibly have unspecified other impact.</td></tr>
<tr><td>3</td><td>CVE-2017-8890</td><td>415</td><td></td><td>DoS</td><td>2017-05-10</td><td>2017-05-24</td><td>10.0</td><td>None</td><td>Remote</td><td>Low</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">The inet_csk_clone_lock function in net/ipv4/inet_connection_sock.c in the Linux kernel through 4.10.15 allows attackers to cause a denial of service (double free) or possibly have unspecified other impact by leveraging use of the accept system call.</td></tr>
<tr><td>4</td><td>CVE-2017-7895</td><td>189</td><td></td><td></td><td>2017-04-28</td><td>2017-05-11</td><td>10.0</td><td>None</td><td>Remote</td><td>Low</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">The NFSv2 and NFSv3 server implementations in the Linux kernel through 4.10.13 lack certain checks for the end of a buffer, which allows remote attackers to trigger pointer-arithmetic errors or possibly have unspecified other impact via crafted requests, related to fs/nfsd/nfs3xdr.c and fs/nfsd/nfsxdr.c.</td></tr>
<tr><td>5</td><td>CVE-2017-0648</td><td>264</td><td></td><td>Exec Code</td><td>2017-06-14</td><td>2017-07-07</td><td>9.3</td><td>None</td><td>Remote</td><td>Medium</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">An elevation of privilege vulnerability in the kernel FIQ debugger could enable a local malicious application to execute arbitrary code within the context of the kernel. This issue is rated as High due to the possibility of a local permanent device compromise, which may require reflashing the operating system to repair the device. Product: Android. Versions: Kernel-3.10. Android ID: A-36101220.</td></tr>
<tr><td>6</td><td>CVE-2017-0605</td><td>264</td><td></td><td>Exec Code</td><td>2017-05-12</td><td>2017-05-19</td><td>9.3</td><td>None</td><td>Remote</td><td>Medium</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">An elevation of privilege vulnerability in the kernel trace subsystem could enable a local malicious application to execute arbitrary code within the context of the kernel. This issue is rated as Critical due to the possibility of a local permanent device compromise, which may require reflashing the operating system to repair the device. Product: Android. Versions: Kernel-3.10, Kernel-3.18. Android ID: A-35399704. References: QC-CR#1048480.</td></tr>
<tr><td>7</td><td>CVE-2017-0564</td><td>264</td><td></td><td>Exec Code</td><td>2017-04-07</td><td>2017-07-10</td><td>9.3</td><td>None</td><td>Remote</td><td>Medium</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">An elevation of privilege vulnerability in the kernel ION subsystem could enable a local malicious application to execute arbitrary code within the context of the kernel. This issue is rated as Critical due to the possibility of a local permanent device compromise, which may require reflashing the operating system to repair the device. Product: Android. Versions: Kernel-3.10, Kernel-3.18. Android ID: A-34276203.</td></tr>
<tr><td>8</td><td>CVE-2017-0563</td><td>264</td><td></td><td>Exec Code</td><td>2017-04-07</td><td>2017-07-10</td><td>9.3</td><td>None</td><td>Remote</td><td>Medium</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">An elevation of privilege vulnerability in the HTC touchscreen driver could enable a local malicious application to execute arbitrary code within the context of the kernel. This issue is rated as Critical due to the possibility of a local permanent device compromise, which may require reflashing the operating system to repair the device. Product: Android. Versions: Kernel-3.10. Android ID: A-32089409.</td></tr>
<tr><td>9</td><td>CVE-2017-0561</td><td>264</td><td></td><td>Exec Code</td><td>2017-04-07</td><td>2017-08-15</td><td>10.0</td><td>None</td><td>Remote</td><td>Low</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">A remote code execution vulnerability in the Broadcom Wi-Fi firmware could enable a remote attacker to execute arbitrary code within the context of the Wi-Fi SoC. This issue is rated as Critical due to the possibility of remote code execution in the context of the Wi-Fi SoC. Product: Android. Versions: Kernel-3.10, Kernel-3.18. Android ID: A-34199105. References: B-RB#110814.</td></tr>
<tr><td>10</td><td>CVE-2017-0528</td><td>264</td><td></td><td>Exec Code Bypass</td><td>2017-03-07</td><td>2017-07-17</td><td>9.3</td><td>None</td><td>Remote</td><td>Medium</td><td>Not required</td><td>Complete</td><td>Complete</td><td>Complete</td></tr>
<tr><td colspan="15">An elevation of privilege vulnerability in the kernel security subsystem could enable a local malicious application to to execute code in the context of a privileged process. This issue is rated as High because it is a general bypass for a kernel level defense in depth or exploit mitigation technology. Product: Android. Versions: Kernel-3.18. Android ID: A-33351919.</td></tr>
</table>

OS Security

- More recent Linux vulnerabilities
  - Linux message interface (August 2005, CAN-2005-2490)
  - race condition - `proc` and `prctl` (July 2006, CVE-2006-3626)
  - local privilege escalation - (September 2007, CVE 2007-4573)
  - security bypass and DoS (May 2008, CVE-2008-2148, CVE-2008-2137)
  - local privilege escalation (August 2009, CVE-2009-2692)
  - local privilege escalation (September 2010, CVE-2010-3081)
  - code execution (June 2016, CVE-2016-4440)

- Device driver code is particularly vulnerable
  - (most) drivers run in kernel mode, either kernel modules or compiled-in
  - often not well audited
  - very large code based compared to core services

- Examples
  - `aironet`, `asus_acpi`, `decnet`, `mpu401`, `msnd`, and `pss` (2004) found by sparse (tool developed by Linus Torvalds)
  - remote root (MadWifi - 2006, Broadcom - 2006)

Kernel exploitation research is active

Unleashing Use-Before-Initialization Vulnerabilities in the Linux Kernel Using Targeted Stack Spraying

- reliably exploiting uninitialized uses on the kernel stack has been considered infeasible
- code executed prior to triggering the vulnerability must leave an attacker-controlled pattern on the stack
- a fully automated targeted stackspraying approach for the Linux kernel that reliably facilitates the exploitation of uninitialized uses
- published in NDSS 2017

source: https://www.cc.gatech.edu/~klu38/publications/ubi-ndss17.pdf

Unix

- Code running in user mode is always linked to a certain identity
  - security checks and access control decisions are based on user identity
- Unix is user-centric
  - no roles
- User
  - identified by username (UID), group name (GID)
  - typically authenticated by password (stored encrypted)
- User `root`
  - superuser, system administrator
  - special privileges (access resources, modify OS)
  - cannot decrypt user passwords

Process Management

- Process
  - implements user-activity
  - entity that executes a given piece of code
  - has its own execution stack, memory pages, and file descriptors table
  - separated from other processes using the virtual memory abstraction
- Thread
  - separate stack and program counter
  - share memory pages and file descriptor table

Process Management

- Process Attributes
  - process ID (PID)
    - uniquely identified process
  - (real) user ID (UID)
    - ID of owner of process
  - effective user ID (EUID)
    - ID used for permission checks (e.g., to access resources)
  - saved user ID (SUID)
    - to temporarily drop and restore privileges
  - lots of management information
    - scheduling
    - memory management, resource management

Process Management

- Switching between IDs
  - uid-setting system calls

    int setuid(uid_t uid)

    int seteuid(uid_t uid)

    int setresuid(uid_t ruid, uid_t euid, uid_t suid)

- Can be tricky
  - POSIX 1003.1:

    If the process has appropriate privileges, the setuid(newuid) function sets the real user ID, effective user ID, and the [saved user ID] to newuid.
  - what are appropriate privileges?

    Solaris: EUID = 0; FreeBSD: newuid = EUID;

    Linux: SETUID capability

Process Management

Bug in sendmail 8.10.1:

- call to setuid(getuid()) to clear privileges (effective UID is root)
- on Linux, attacker could clear SETUID capability
- call clears EUID, but SUID remains root

Further reading

Setuid Demystified
Hao Chen, David Wagner, and Drew Dean
11th USENIX Security Symposium, 2002

User Authentication

- How does a process get a user ID?
- Authentication
- Passwords
  - user passwords are used as keys for `crypt()` function
  - uses SHA-512
  - 8-byte "salt"
    - chosen from date, not secret
    - prevent same passwords to map onto same string
    - make dictionary attacks more difficult
- Password cracking
  - dictionary attacks, rainbow tables
  - `Crack, JohnTheRipper`

User Authentication

- Shadow passwords
  - password file is needed by many applications to map user ID to user names
  - encrypted passwords are not
- `/etc/shadow`
  - holds encrypted passwords
  - account information
    - last change date
    - expiration (warning, disabled)
    - minimum change frequency
  - readable only by superuser and privileged programs
  - SHA-512 hashed passwords (default on Ubuntu) to slow down guessing

User Authentication

- Shadow passwords
  - a number of other encryption / hashing algorithms were proposed
  - blowfish, SHA-1, …
- Other authentication means possible
  - Linux PAM (pluggable authentication modules)
  - Kerberos
  - Active directory (Windows)

Group Model

- Users belong to one or more groups
  - primary group (stored in `/etc/password`)
  - additional groups (stored in `/etc/group`)
  - possibility to set group password
  - and become group member with `newgrp`

- `/etc/group`

```
groupname : password : group id : additional users
root:x:0:root
bin:x:1:root,bin,daemon
users:x:100:akaprav
```

- Special group `wheel/sudo`
  - protect `root` account by limiting user accounts that can perform `su`

File System

- File tree
  - primary repository of information
  - hierarchical set of directories
  - directories contain file system objects (FSO)
  - root is denoted "/"

- File system object
  - files, directories, symbolic links, sockets, device files
  - referenced by *inode* (index node)

File System

- Access Control
  - permission bits
  - `chmod, chown, chgrp, umask`
  - file listing:

```
     -        rwx     rwx      rwx
(file type) (user)  (group)  (other)
```

Type	r	w	x	s	t
File	read access	write access	execute	suid / sgid inherit id	sticky bit
Directory	list files	insert and remove files	stat / execute files, chdir	new files have dir-gid	files/dirs only delete-able by owner

Sticky bit

- What's a very common directory with sticky bit?

```
$  ls -ld /tmp
 drwxrwxrwt 26 root root 69632 Sep  7 15:24 /tmp
 $ ls -l test
 -rw-rw-r-- 1 kapravel kapravel 0 Sep  7 15:29 test
 $  chmod +t test; ls -l test
 -rw-rw-r-T 1 kapravel kapravel 0 Sep  7 15:29 test
```

SUID Programs

- Each process has real and effective user / group ID
  - usually identical
  - real IDs
    - determined by current user
    - authentication (login, su)
  - effective IDs
    - determine the "rights" of a process
    - system calls (e.g., setuid())
  - suid / sgid bits
    - to start process with effective ID different from real ID
    - attractive target for attacker
- Never use SUID shell scripts (multiplying problems)

File System

- Shared resource
  - susceptible to race condition problems
- Time-of-Check, Time-of-Use (**TOCTOU**)
  - common race condition problem
  - problem:
    - Time-Of-Check (t_1): validity of assumption A on entity E is checked
    - Time-Of-Use (t_2): assuming A is still valid, E is used
    - Time-Of-Attack (t_3): assumption A is invalidated

$$t_1 < t_3 < t_2$$

TOCTOU

- Steps to access a resource
  1. obtain reference to resource
  2. query resource to obtain characteristics
  3. analyze query results
  4. if resource is fit, access it

- Often occurs in Unix file system accesses
  - check permissions for a certain file name (e.g., using **`access(2)`**)
  - open the file, using the file name (e.g., using **`fopen(3)`**)
  - four levels of indirection (symbolic link - hard link - inode - file descriptor)

- Windows uses file handles and includes checks in API open call

Overview

```
/* access returns 0 on success */
if(!access(file, W_OK)) {
   f = fopen(file, "wb+");
   write_to_file(f);
} else {
   fprintf(stderr, "Permission denied \
                   when trying to open %s.\n", file);
}
```

- Attack

```
$ touch dummy; ln -s dummy pointer
$ rm pointer; ln -s /etc/passwd pointer
```

Examples

- TOCTOU Examples
  - Setuid Scripts
    1. exec() system call invokes seteuid() call prior to executing program
    2. program is a script, so command interpreter is loaded first
    3. program interpreted (with root privileges) is invoked on script name
    4. attacker can replace script content between step 2 and 3

Examples

- TOCTOU Examples
  - Directory operations
    - rm can remove directory trees, traverses directories depth-first
    - issues chdir("..") to go one level up after removing a directory branch
    - by relocating subdirectory to another directory, arbitrary files can be deleted
  - Temporary files
    - commonly opened in /tmp or /var/tmp
    - often guessable file names

Temporary Files

"Secure" procedure for creating temporary files

1. pick a prefix for your filename
2. generate at least 64 bits of high-quality randomness
3. base64 encode the random bits
4. concatenate the prefix with the encoded random data
5. set umask appropriately (0066 is usually good)
6. use fopen(3) to create the file, opening it in the proper mode
7. delete the file immediately using unlink(2)
8. perform reads, writes, and seeks on the file as necessary
9. finally, close the file

Temporary Files

- Library functions to create temporary files can be insecure
  - mktemp(3) is not secure, use mkstemp(3) instead
  - old versions of mkstemp(3) did not set umask correctly

- Temp Cleaners
  - programs that clean "old" temporary files from temp directories
  - first lstat(2) file, then use unlink(2) to remove files
  - vulnerable to race condition when attacker replaces file between lstat(2) and unlink(2)
  - arbitrary files can be removed
  - delay program long enough until temp cleaner removes active file

Prevention

- "Handbook of Information Security Management" suggests
  1. increase number of checks
  2. move checks closer to point of use
  3. immutable bindings

- Only number 3 is secure!
- Immutable bindings
  - operate on file descriptors
  - do not check access by yourself (i.e., no use of access(2)) drop privileges instead and let the file system do the job
- Use the O_CREAT | O_EXCL flags to create a new file with open(2)
  and be prepared to have the open call fail

Prevention

Series of papers on the access system call

Fixing races for fun and profit: how to use access(2)
D. Dean and A. Hu
Usenix Security Symposium, 2004

Fixing races for fun and profit: howto abuse atime
N. Borisov, R. Johnson, N. Sastry, and D. Wagner
Usenix Security Symposium, 2005

Portably Solving File TOCTTOU Races with Hardness Amplification
D. Tsafrir, T. Hertz, D. Wagner, and D.Da Silva
Usenix Conference on File and Storage Technologies (FAST), 2008

Locking

- Ensures exclusive access to a certain resource
- Used to circumvent accidental race conditions
  - advisory locking (processes need to cooperate)
  - not mandatory, therefore not secure
- Often, files are used for locking
  - portable (files can be created nearly everywhere)
  - “stuck” locks can be easily removed
- Simple method
  - create file using the O_EXCL flag

Shell

- Shell
  - one of the core Unix application
  - both a command language and programming language
  - provides an interface to the Unix operating system

  - rich features such as control-flow primitives, parameter passing, variables, and string substitution
  - communication between shell and spawned programs via redirection and pipes

  - different flavors
    - bash and sh, tcsh and csh, ksh, zsh

Shell Attacks

- Environment Variables
  - $HOME and $PATH can modify behavior of programs that operate with relative path names
  - $IFS – internal field separator
    - used to parse tokens
    - usually set to [\t\n] but can be changed to “/“
    - “/bin/ls“ is parsed as “bin ls“ calling bin locally
    - IFS now only used to split expanded variables
  - preserve attack (/usr/lib/preserve is SUID)
    - called “/bin/mail“ when vi crashes to preserve file
    - change IFS, create bin as link to /bin/sh, kill vi

Shell Attacks

- Control and escape characters
  - can be injected into command string
  - modify or extend shell behavior
  - user input used for shell commands has to be rigorously sanitized
  - easy to make mistakes
  - classic examples are `;’ and `&’
- Applications that are invoked via shell can be targets as well
  - increased vulnerability surface
- Restricted shell
  - invoked with -r or rbash
  - more controlled environment

Shell Attacks

- `system(char *cmd)`
  - function called by programs to execute other commands
  - invokes shell
  - executes string argument by calling /bin/sh –c string
  - makes binary program vulnerable to shell attacks
  - especially when user input is utilized

- popen(char *cmd, char *type)
  - forks a process, opens a pipe and invokes shell for cmd

File Descriptor Attacks

- SUID program opens file
- forks external process
  - sometimes under user control
- on-execute flag
  - if close-on-exec flag is not set, then new process inherits file descriptor
  - malicious attacker might exploit such weakness
- Linux Perl 5.6.0
  - getpwuid() leaves /etc/shadow opened (June 2002)
  - problem for Apache with mod_perl
  - web browsers and flash

Resource Limits

- File system limits
  - quotas
  - restrict number of storage blocks and number of inodes
  - hard limit
    - can never be exceeded (operation fails)
  - soft limit
    - can be exceeded temporarily
  - can be defined per mount-point
  - defend against resource exhaustion (denial of service)

- Process resource limits
  - number of child processes, open file descriptors

Signals

- Signal
  - simple form of interrupt
  - asynchronous notification
  - can happen anywhere for process in user space
  - used to deliver segmentation faults, reload commands, …
  - kill command

- Signal handling
  - process can install signal handlers
  - when no handler is present, default behavior is used
    - ignore or kill process
  - possible to catch all signals except SIGKILL (-9)

Signals

- Security issues
  - code has to be re-entrant
    - atomic modifications
    - no global data structures
  - race conditions
  - unsafe library calls, system calls
  - examples
    - wu-ftpd 2001, sendmail 2001 + 2006, stunnel 2003, ssh 2006
- Secure signals
  - write handler as simple as possible
  - block signals in handler

Shared Libraries

- Library
  - collection of object files
  - included into (linked) program as needed
  - code reuse

- Shared library
  - multiple processes share a single library copy
  - save disk space (program size is reduced)
  - save memory space (only a single copy in memory)
  - used by virtually all Unix applications (at least libc.so)
  - check binaries with ldd

Shared Libraries

- Static shared library
  - address binding at link-time
  - not very flexible when library changes
  - code is fast
- Dynamic shared library
  - address binding at load-time
  - uses procedure linkage table (PLT) and global offset table (GOT)
  - code is slower (indirection)
  - loading is slow (binding has to be done at run-time)
  - classic .so or .dll libraries
- PLT and GOT entries are very popular attack targets
  - buffer overflows

Shared Libraries

- Management
  - stored in special directories (listed in `/etc/ld.so.conf`)
  - manage cache with `ldconfig`

- Preload
  - override (substitute) with other version
  - use `/etc/ld.so.preload`
  - can also use environment variables for override
  - possible security hazard
  - now disabled for SUID programs (old Solaris vulnerability)

Advanced Security Features

- Address space protection
  - address space layout randomization (ASLR)
  - non-executable stack (based on NX bit or PAX patches)
- Mandatory access control extensions
  - SELinux
  - role-based access control extensions
  - capability support
- Miscellaneous improvements
  - hardened chroot jails
  - better auditing
- https://wiki.ubuntu.com/Security/Features

RIP Solaris

solaris™