

## Operating Systems Project Report

<b>Project Number (01 / 02 / 03):</b>	01
<b>Name:</b>	蕭望緯
<b>Student ID:</b>	0811521
<b>YouTube link (Format youtube.com/watch?v=[key]):</b>	<a href="https://youtu.be/Y0dAnP7n9mw">https://youtu.be/Y0dAnP7n9mw</a>
<b>Date (YYYY-MM-DD):</b>	2021-10-20
<b>Names of the files uploaded to E3:</b>	OS_Project01_0811521.pdf
<b>Physical Machine Total RAM (Example: 8.0 GB):</b>	16GB
<b>Physical Machine CPU (Example: Intel i7-2600K):</b>	11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz 2.42 GHz

Checklist	
Yes/No	Item
Yes	The report name follows the format "OS_ProjectXX_StudentID.pdf".
Yes	The report was uploaded to E3 before the deadline.
Yes	The YouTube video is public, and anyone with the link can watch it.
Yes	The audio of the video has a good volume.
Yes	The pictures in your report and video have a good quality.
Yes	All the questions and exercises were answered inside the report.
Yes	I understand that late submission is late submission, regardless of the time uploaded.
Yes	I understand that any cheating in my report / video / code will not be tolerated.

## Individual Questions:

1. What is a Kernel? What are the differences between *mainline*, *stable* and *longterm*? What is a Kernel panic?

Ans:

Kernel: lies in the center part of an operating system, responsible for process and memory management, file systems, device control and networking.

*mainline*, *stable* and *longterm* are three sections that release some kernel versions:

mainline: provide versions that contain new features

stable: versions in the mainline section will be sent to the stable section, and bugfix versions will be released frequently

longterm: bugfix versions are released less frequently

Kernel panic: occurs when any fatal error is detected, and the OS will stop to protect the system

2. What are the differences between *building*, *debugging* and *profiling*?

Ans:

building: install a new kernel

debugging: for the development of kernel

profiling: monitor the performance of processes and resources allocation of the kernel

3. What are GCC, GDB, and KGDB, and what they are used for?

Ans:

GCC: compiler for various programming languages, including C, C++, Fortran, Ada, Go, and etc.

GDB: debugger that aids the development of software programs

KGDB: debugger specific to development of kernel, runs on the Target machine to be debugged

4. What are the `/usr/`, `/boot/`, `/home/`, `/boot/grub` folders for?

Ans:

`/usr`: contains important files such as data files, libraries, binaries, etc, which support the system.

`/boot`: contains files related to the booting of the system

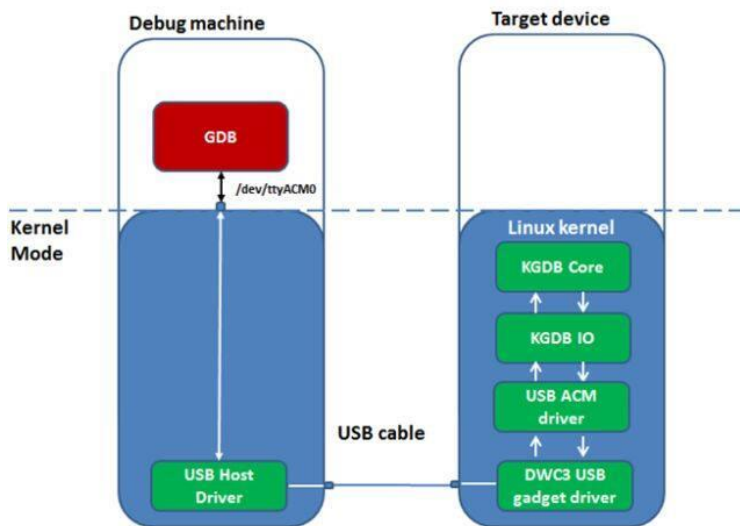
`/home`: personal workspace for users where the users have full access to writing and creating files

`/boot/grub`: contains files related to GRUB

5. What are the general steps to debug a Linux Kernel? (Add a figure)

Ans:

- A debugger machine and a target machine
- Connection between them
- gdb on the debugger machine and kgdb on the target machine
- establish debugging session



source: [1]

6. For this project, why do we need two virtual machines?

Ans:

Using kgdb requires two virtual machines. Debugger machine is used as a debugging tool and runs gdb, while Target machine is debugged, runs kgdb, and can be frozen by the debugger. Also, this project simulates remote debugging in real world situation.

7. In Section 3.2, what are the differences between **make**, **make modules\_install** and **make install**?

Ans:

**make:** compile the kernel

**make modules\_install:** install compiled kernel modules

**make install:** install the kernel so that we can execute it

8. In Section 3.3, what are the commands **kgdbwait** and **kgdboc=ttyS1,115200** for?

Ans:

**kgdbwait:** Upon booting the kernel, the system will wait for the connection to the KGDB

**kgdboc:** configure the device through which gdb and kgdb communicate

9. What is **grub**? What is **grub.cfg**?

Ans:

grub: a boot loader that manages the booting process and allows users to launch one of the installed operating system

grub.cfg: GRUB configuration file that contains menu setting and individual setting for each installed kernel. It is generated by several grub.d files.

10. List at least 10 commands you can use with GDB.

Ans:

- connect to remote machine: **target remote /dev/ttyS1**
- unfreeze the target machine: **continue**
- set breakpoint at the beginning of a function: **break function**
- list information of all current breakpoints: **info breakpoints**
- delete all breakpoints: **delete**
- delete breakpoints set in the function: **clear function**
- print the parameter of the function hitting a breakpoint: **print parameter**
- list source code nearby the index: **list index**
- execute the next line of program (execute the entire function call): **next**
- execute the next line of program (include each line of a function call): **step**
- show the recent locations (several frames, the entire stack) in the program: **backtrace**
- show current stack frame: **frame**
- list information of current frame: **info frame**

11. What is a kernel function? What is a system call?

Ans:

kernel function: jobs supported by the kernel

system call: request services from the kernel

12. What is KASLR? What is it for?

Ans:

“Kernel address space layout randomization”

It enables address space randomization to protect the memory from attack and illegal access.

13. What are GDB’s non-stop and all-stop modes?

Ans:

non-stop: in multi-threaded system, GDB freezes certain threads while other threads keep running

all-stop modes: GDB freezes the entire system (all threads)

14. Explain what the command **echo g > /proc/sysrq-trigger** does.

Ans:

stop the kernel execution and give control to kgdb

15. What are these functions: **clone**, **mmap**, **write** and **open**?

Ans:

**clone**: create a child process

**mmap**: creates a new mapping of files to memory in the virtual address space of the calling process

**write**: writes bytes from the buffer to the file pointed by file descriptor.

**open**: opens the file specified by a pathname.

16. Why is there no **fork** system call? What is the difference between **fork** and **clone**?

Ans:

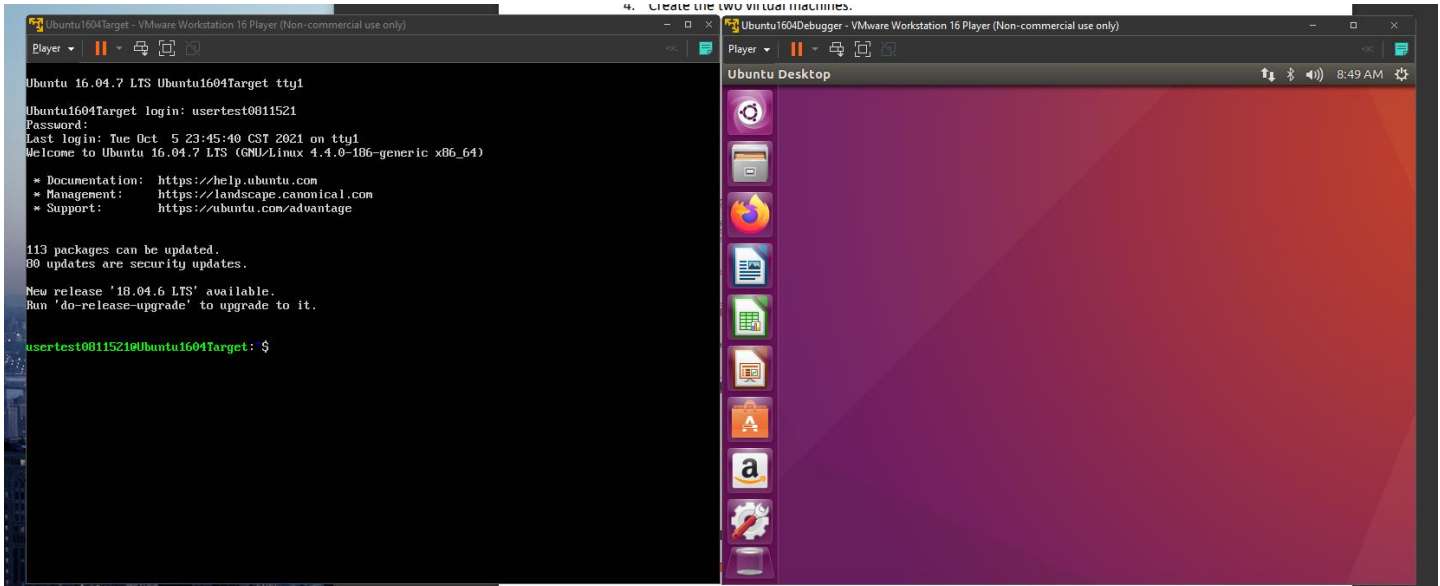
**fork** will call **clone** system call and do similar things like **clone**.

**fork**: the created child process doesn't share resources (e.g., memory space, files) with the parent

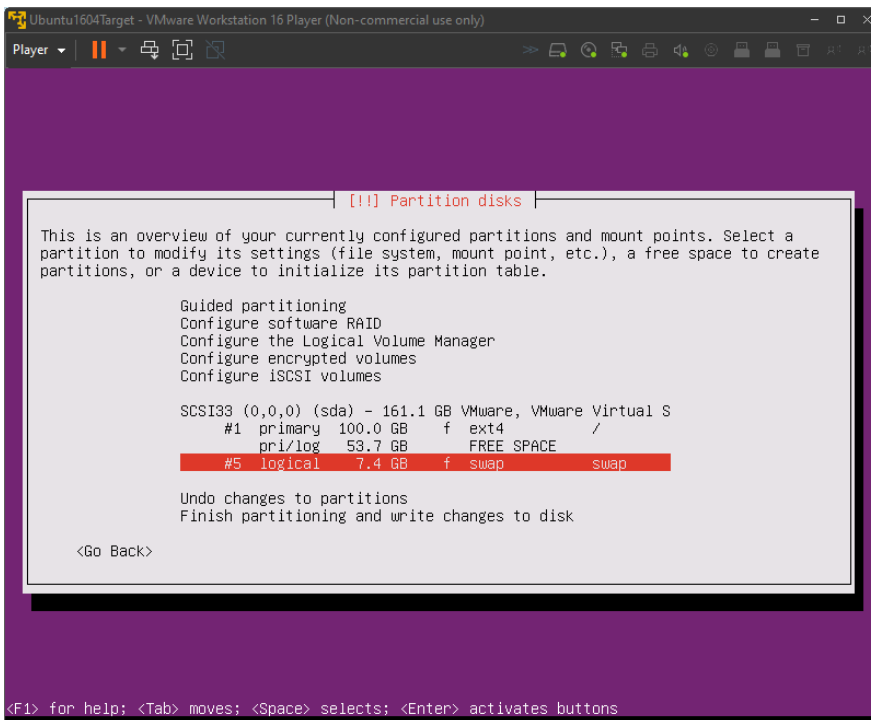
**clone**: the created child process can share resources with the parent

## Screenshot #1

Target(left) & Debugger(right) Machine



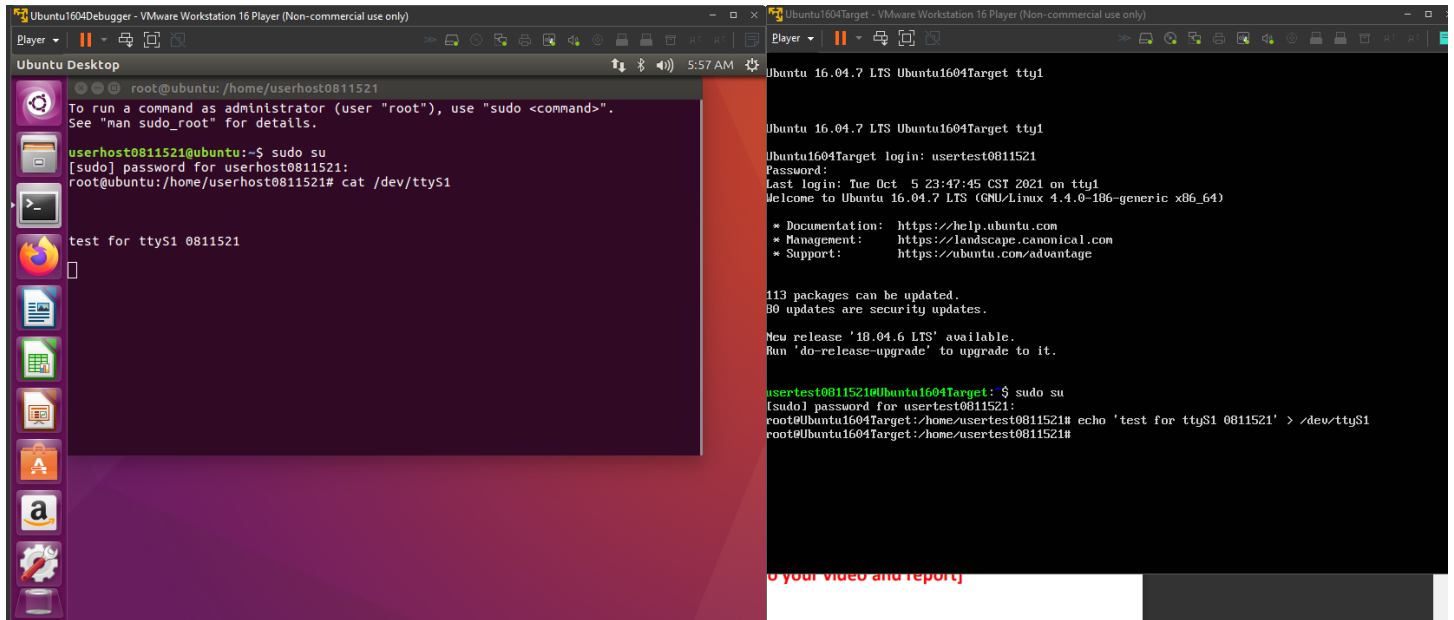
Target machine disks configuration



## Screenshot #2

Serial port communication testing after some configuration

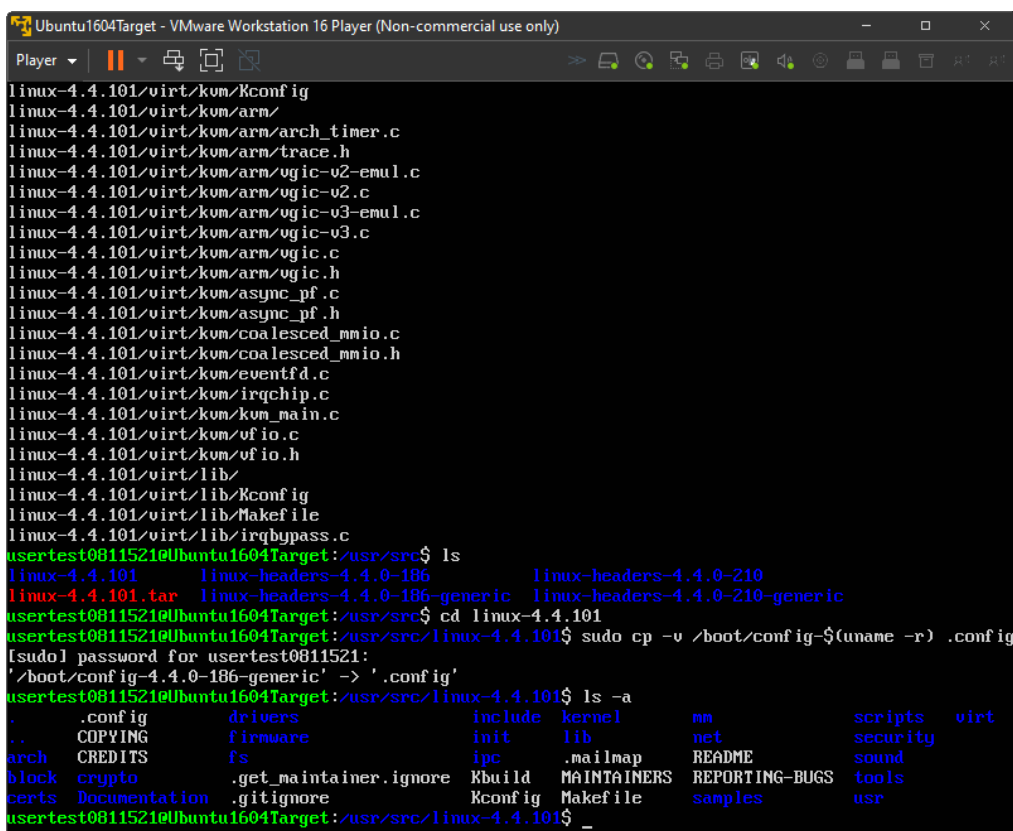
Message passed from Target to Debugger successfully through serial port communication



## Screenshot #3

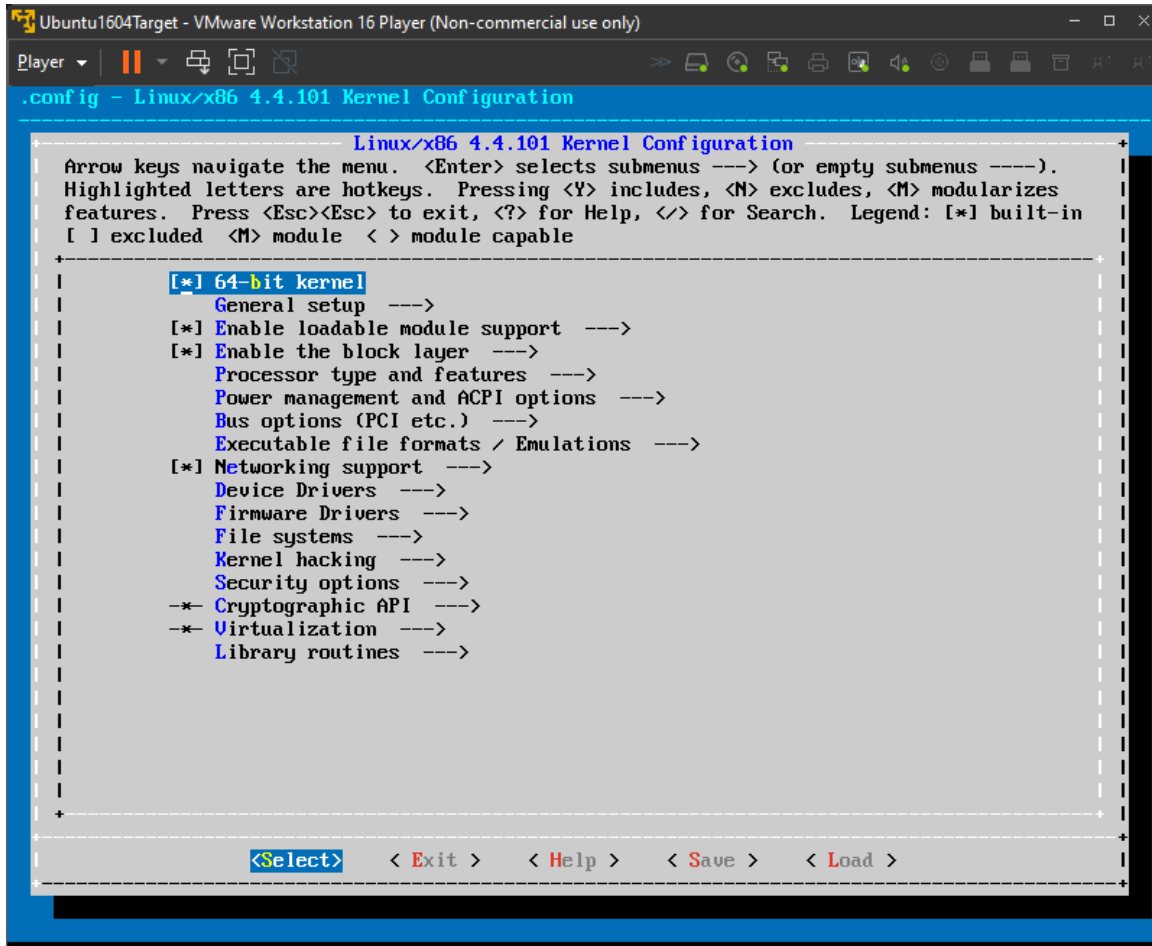
Use the predefined configuration from one of the other versions of kernel

Copy the .config file to the directory of kernel used for this project



## Screenshot #4

Open Kernel configuration interface and examine the setup



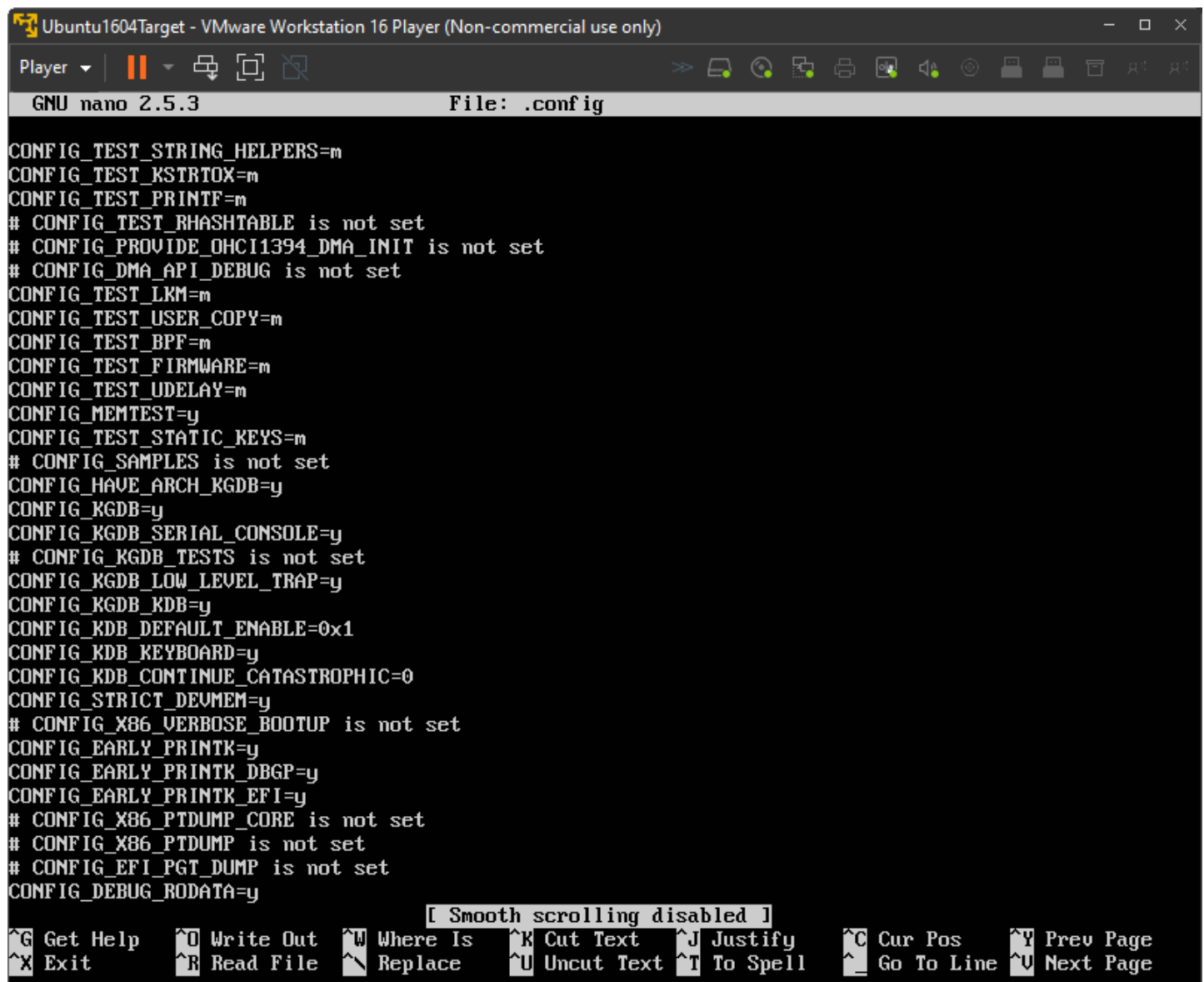


## Screenshot #5

Open the file: `/usr/src/linux-4.4.101/.config`

Double-check these four lines so that we can use kgdb for this kernel in later sections:

```
CONFIG_FRAME_POINTER=y
CONFIG_KGDB=y
CONFIG_KGDB_SERIAL_CONSOLE=y
CONFIG_KGDB_KDB=y
CONFIG_KDB_KEYBOARD=y
```



```
Ubuntu1604Target - VMware Workstation 16 Player (Non-commercial use only)
Player ▾ | [Icons]
GNU nano 2.5.3 File: .config

CONFIG_TEST_STRING_HELPERS=m
CONFIG_TEST_KSTRTOX=m
CONFIG_TEST_PRINTF=m
# CONFIG_TEST_RHASHTABLE is not set
# CONFIG_PROVIDE_OHCI1394_DMA_INIT is not set
# CONFIG_DMA_API_DEBUG is not set
CONFIG_TEST_LKM=m
CONFIG_TEST_USER_COPY=m
CONFIG_TEST_BPF=m
CONFIG_TEST_FIRMWARE=m
CONFIG_TEST_UDELAY=m
CONFIG_MEMTEST=y
CONFIG_TEST_STATIC_KEYS=m
# CONFIG_SAMPLES is not set
CONFIG_HAVE_ARCH_KGDB=y
CONFIG_KGDB=y
CONFIG_KGDB_SERIAL_CONSOLE=y
# CONFIG_KGDB_TESTS is not set
CONFIG_KGDB_LOW_LEVEL_TRAP=y
CONFIG_KGDB_KDB=y
CONFIG_KDB_DEFAULT_ENABLE=0x1
CONFIG_KDB_KEYBOARD=y
CONFIG_KDB_CONTINUE_CATASTROPHIC=0
CONFIG_STRICT_DEVMEM=y
# CONFIG_X86_VERBOSE_BOOTUP is not set
CONFIG_EARLY_PRINTK=y
CONFIG_EARLY_PRINTKDBG=y
CONFIG_EARLY_PRINTK_EFI=y
# CONFIG_X86_PTDUMP_CORE is not set
# CONFIG_X86_PTDUMP is not set
# CONFIG_EFI_PGT_DUMP is not set
CONFIG_DEBUG_RODATA=y

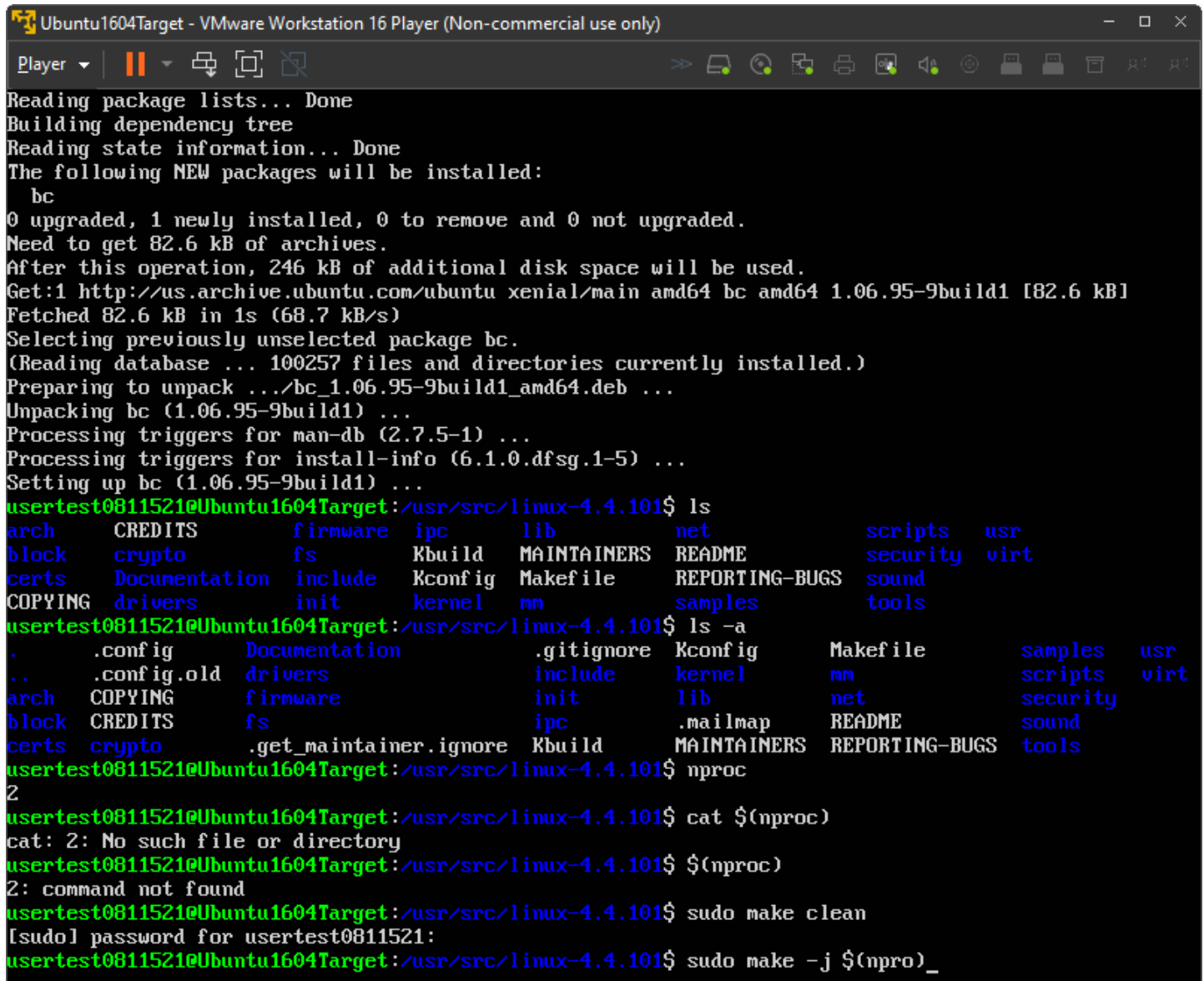
[ Smooth scrolling disabled ]

^G Get Help      ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify      ^C Cur Pos      ^Y Prev Page
^X Exit          ^R Read File    ^_ Replace      ^U Uncut Text   ^T To Spell     ^_ Go To Line   ^V Next Page
```

## Screenshot #6

Start building the kernel.

First, clean up the executable or compiled object files. Then, build kernel with the maximum available processors.



```
Ubuntu1604Target - VMware Workstation 16 Player (Non-commercial use only)
Player
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  bc
0 upgraded, 1 newly installed, 0 to remove and 0 not upgraded.
Need to get 82.6 kB of archives.
After this operation, 246 kB of additional disk space will be used.
Get:1 http://us.archive.ubuntu.com/ubuntu xenial/main amd64 bc amd64 1.06.95-9build1 [82.6 kB]
Fetched 82.6 kB in 1s (68.7 kB/s)
Selecting previously unselected package bc.
(Reading database ... 100257 files and directories currently installed.)
Preparing to unpack .../bc_1.06.95-9build1_amd64.deb ...
Unpacking bc (1.06.95-9build1) ...
Processing triggers for man-db (2.7.5-1) ...
Processing triggers for install-info (6.1.0.dfsg.1-5) ...
Setting up bc (1.06.95-9build1) ...
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ ls
arch      CREDITS      firmware     ipc          lib          net          scripts     usr
block     crypto       fs           Kbuild      MAINTAINERS  README      security    virt
certs     Documentation include      Kconfig     Makefile     REPORTING-BUGS sound
COPYING   drivers      init         kernel      mm           samples      tools
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ ls -a
.          .config      Documentation .gitignore  Kconfig     Makefile     samples     usr
..         .config.old  drivers       include     kernel      mm           scripts     virt
arch       COPYING      firmware     init        lib         net          security
block      CREDITS      fs           ipc         .mailmap    README      sound
certs      crypto       .get_maintainer.ignore Kbuild      MAINTAINERS REPORTING-BUGS tools
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ nproc
2
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ cat $(nproc)
cat: 2: No such file or directory
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ $(nproc)
2: command not found
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ sudo make clean
[sudo] password for usertest0811521:
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ sudo make -j $(nproc)_
```

## Screenshot #7

After kernel is compiled, install kernel modules

```
usertest0811521@Ubuntu1604Target: $
usertest0811521@Ubuntu1604Target: $ sudo make modules_install
[sudo] password for usertest0811521:
make: *** No rule to make target 'modules_install'. Stop.
usertest0811521@Ubuntu1604Target: $ sudo make clean
make: *** No rule to make target 'clean'. Stop.
usertest0811521@Ubuntu1604Target: $ cd /usr/src/linux-4.4.101
-bash: cd: /usr/src/linux-4.4.101: No such file or directory
usertest0811521@Ubuntu1604Target: $ cd /usr/src/linux-4.4.101
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ ls
arch      CREDITS   firmware  ipc        lib        modules.builtin  REPORTING-BUGS
block     crypto    fs         Kbuild     MAINTAINERS  modules.order    samples
certs     documenta include   Kconfig    Makefile     net              scripts
COPYING   drivers   init       kernel     mm           README           security
usertest0811521@Ubuntu1604Target: /usr/src/linux-4.4.101$ sudo make modules_install_
```

## Screenshot #8

Install the complete kernel to the system

```
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$ sudo make install
sh ./arch/x86/boot/install.sh 4.4.101 arch/x86/boot/bzImage \
    System.map "/boot"
run-parts: executing /etc/kernel/postinst.d/apt-auto-removal 4.4.101 /boot/vmlinuz-4.4.101
run-parts: executing /etc/kernel/postinst.d/initramfs-tools 4.4.101 /boot/vmlinuz-4.4.101
update-initramfs: Generating /boot/initrd.img-4.4.101
W: mdadm: /etc/mdadm/mdadm.conf defines no arrays.
run-parts: executing /etc/kernel/postinst.d/unattended-upgrades 4.4.101 /boot/vmlinuz-4.4.101
run-parts: executing /etc/kernel/postinst.d/update-notifier 4.4.101 /boot/vmlinuz-4.4.101
run-parts: executing /etc/kernel/postinst.d/x-grub-legacy-ec2 4.4.101 /boot/vmlinuz-4.4.101
Searching for GRUB installation directory ... found: /boot/grub
Searching for default file ... found: /boot/grub/default
Testing for an existing GRUB menu.lst file ... found: /boot/grub/menu.lst
Searching for splash image ... none found, skipping ...
Found kernel: /boot/vmlinuz-4.4.0-210-generic
Found kernel: /boot/vmlinuz-4.4.0-186-generic
Found kernel: /boot/vmlinuz-4.4.101
Found kernel: /boot/vmlinuz-4.4.0-210-generic
Found kernel: /boot/vmlinuz-4.4.0-186-generic
Replacing config file /run/grub/menu.lst with new version
Updating /boot/grub/menu.lst ... done

run-parts: executing /etc/kernel/postinst.d/zz-update-grub 4.4.101 /boot/vmlinuz-4.4.101
Generating grub configuration file ...
Found linux image: /boot/vmlinuz-4.4.101
Found initrd image: /boot/initrd.img-4.4.101
Found linux image: /boot/vmlinuz-4.4.0-210-generic
Found initrd image: /boot/initrd.img-4.4.0-210-generic
Found linux image: /boot/vmlinuz-4.4.0-186-generic
Found initrd image: /boot/initrd.img-4.4.0-186-generic
done
```

In the /boot directory, the four files of the kernel 4.4.101 are present.

a. initrd.img-4.4.101

b. vmlinuz-4.4.101

c. system.map-4.4.101

d. config-4.4.101

```
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$ ls -al /boot
total 420916
drwxr-xr-x  3 root root    4096 Oct 11 16:34 .
drwxr-xr-x 23 root root    4096 Oct 11 12:54 ..
-rw-r--r--  1 root root 191087 Jul  1 2020 config-4.4.0-186-generic
-rw-r--r--  1 root root 191002 Apr 16 19:34 config-4.4.0-210-generic
-rw-r--r--  1 root root 188361 Oct 11 16:33 config-4.4.101
drwxr-xr-x  5 root root    4096 Oct 11 16:34 grub
-rw-r--r--  1 root root 41808101 Oct 11 12:54 initrd.img-4.4.0-186-generic
-rw-r--r--  1 root root 41821238 Oct 11 12:55 initrd.img-4.4.0-210-generic
-rw-r--r--  1 root root 313634440 Oct 11 16:34 initrd.img-4.4.101
-rw-----  1 root root 3920886 Jul  1 2020 System.map-4.4.0-186-generic
-rw-----  1 root root 3925753 Apr 16 19:34 System.map-4.4.0-210-generic
-rw-r--r--  1 root root 3837328 Oct 11 16:33 System.map-4.4.101
-rw-----  1 root root 7218016 Jul  6 2020 vmlinuz-4.4.0-186-generic
-rw-----  1 root root 7225568 Apr 17 14:03 vmlinuz-4.4.0-210-generic
-rw-r--r--  1 root root 7017600 Oct 11 16:33 vmlinuz-4.4.101
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$
```

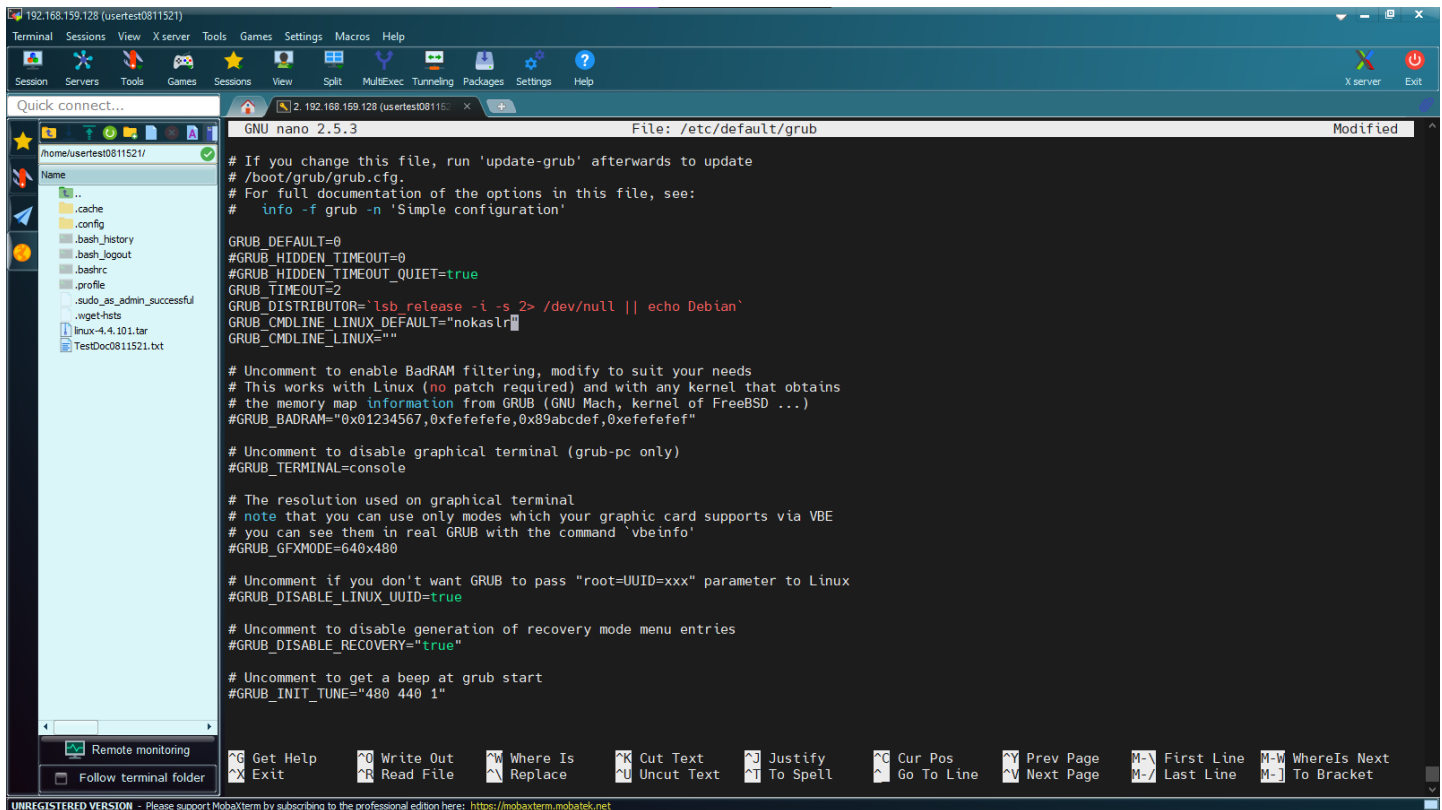
### Screenshot #9

update initrd.img. Then, update GRUB

```
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$ sudo update-initramfs -c -k 4.4.101
update-initramfs: Generating /boot/initrd.img-4.4.101
W: mdadm: /etc/mdadm/mdadm.conf defines no arrays.
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$ sudo update-grub
Generating grub configuration file ...
Found linux image: /boot/vmlinuz-4.4.101
Found initrd image: /boot/initrd.img-4.4.101
Found linux image: /boot/vmlinuz-4.4.0-210-generic
Found initrd image: /boot/initrd.img-4.4.0-210-generic
Found linux image: /boot/vmlinuz-4.4.0-186-generic
Found initrd image: /boot/initrd.img-4.4.0-186-generic
done
usertest0811521@usertest0811521:/usr/src/linux-4.4.101$
```

## Screenshot #10

set up parameters for GRUB so that we can see the menu for different kernels and set nokaslr. Then, update GRUB



The screenshot shows a terminal window with the GNU nano 2.5.3 editor open, editing the file /etc/default/grub. The left sidebar shows a file explorer with various files and folders. The main editor area contains the following configuration:

```
# If you change this file, run 'update-grub' afterwards to update
# /boot/grub/grub.cfg.
# For full documentation of the options in this file, see:
# info -f grub -n 'Simple configuration'

GRUB_DEFAULT=0
#GRUB_HIDDEN_TIMEOUT=0
#GRUB_HIDDEN_TIMEOUT_QUIET=true
GRUB_TIMEOUT=2
GRUB_DISTRIBUTOR=`lsb_release -i -s 2> /dev/null || echo Debian`
GRUB_CMDLINE_LINUX_DEFAULT="nokaslr"
GRUB_CMDLINE_LINUX=""

# Uncomment to enable BadRAM filtering, modify to suit your needs
# This works with Linux (no patch required) and with any kernel that obtains
# the memory map information from GRUB (GNU Mach, kernel of FreeBSD ...)
#GRUB_BADRAM="0x01234567,0xfefefefe,0x89abcdef,0xefefefef"

# Uncomment to disable graphical terminal (grub-pc only)
#GRUB_TERMINAL=console

# The resolution used on graphical terminal
# note that you can use only modes which your graphic card supports via VBE
# you can see them in real GRUB with the command 'vbeinfo'
#GRUB_GFXMODE=640x480

# Uncomment if you don't want GRUB to pass "root=UUID=xxx" parameter to Linux
#GRUB_DISABLE_LINUX_UUID=true

# Uncomment to disable generation of recovery mode menu entries
#GRUB_DISABLE_RECOVERY="true"

# Uncomment to get a beep at grub start
#GRUB_INIT_TUNE="480 440 1"
```

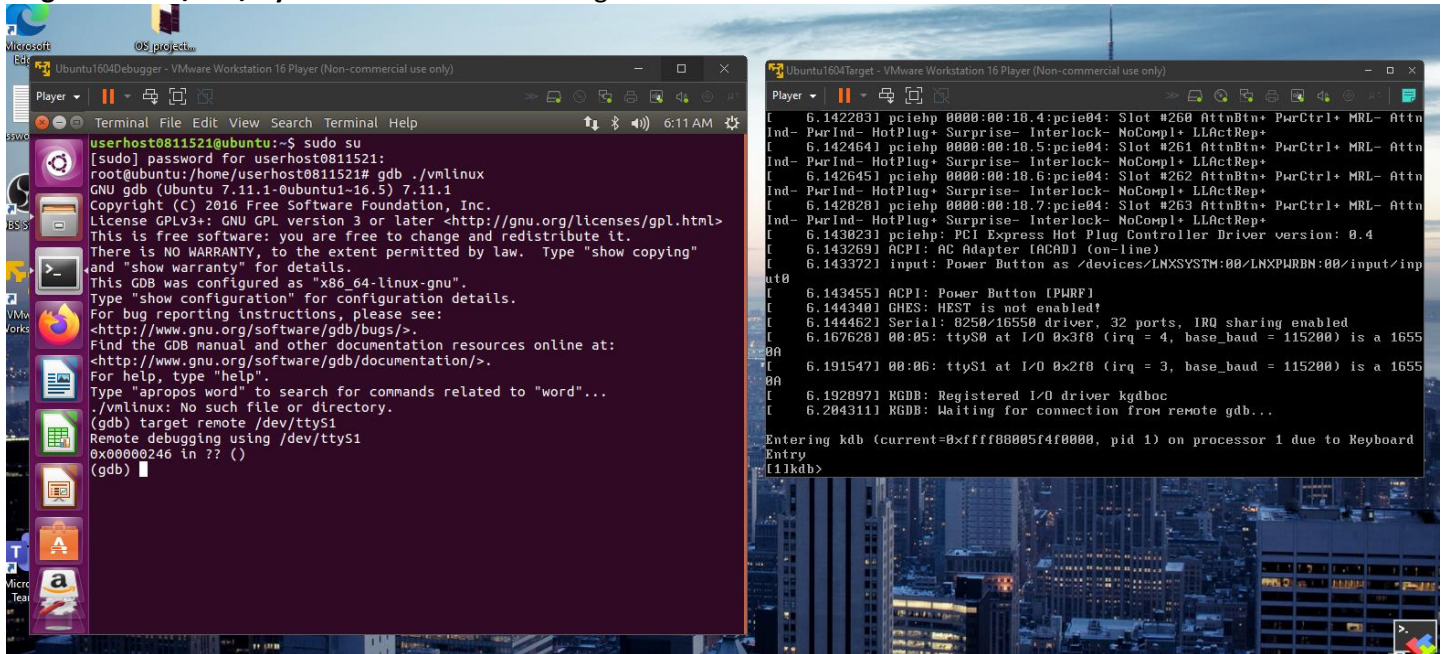
## Screenshot #11

**gdb ./vmlinux:** open gdb

```
root@ubuntu:/boot/kgdb-image# gdb ./vmlinux
GNU gdb (Ubuntu 7.11.1-0ubuntu1~16.5) 7.11.1
Copyright (C) 2016 Free Software Foundation, Inc.
License GPLv3+: GNU GPL version 3 or later <http://gnu.org/licenses/gpl.html>
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, to the extent permitted by law. Type "show copying"
and "show warranty" for details.
This GDB was configured as "x86_64-linux-gnu".
Type "show configuration" for configuration details.
For bug reporting instructions, please see:
<http://www.gnu.org/software/gdb/bugs/>.
Find the GDB manual and other documentation resources online at:
<http://www.gnu.org/software/gdb/documentation/>.
For help, type "help".
Type "apropos word" to search for commands related to "word"...
Reading symbols from ./vmlinux...done.
(gdb)
```

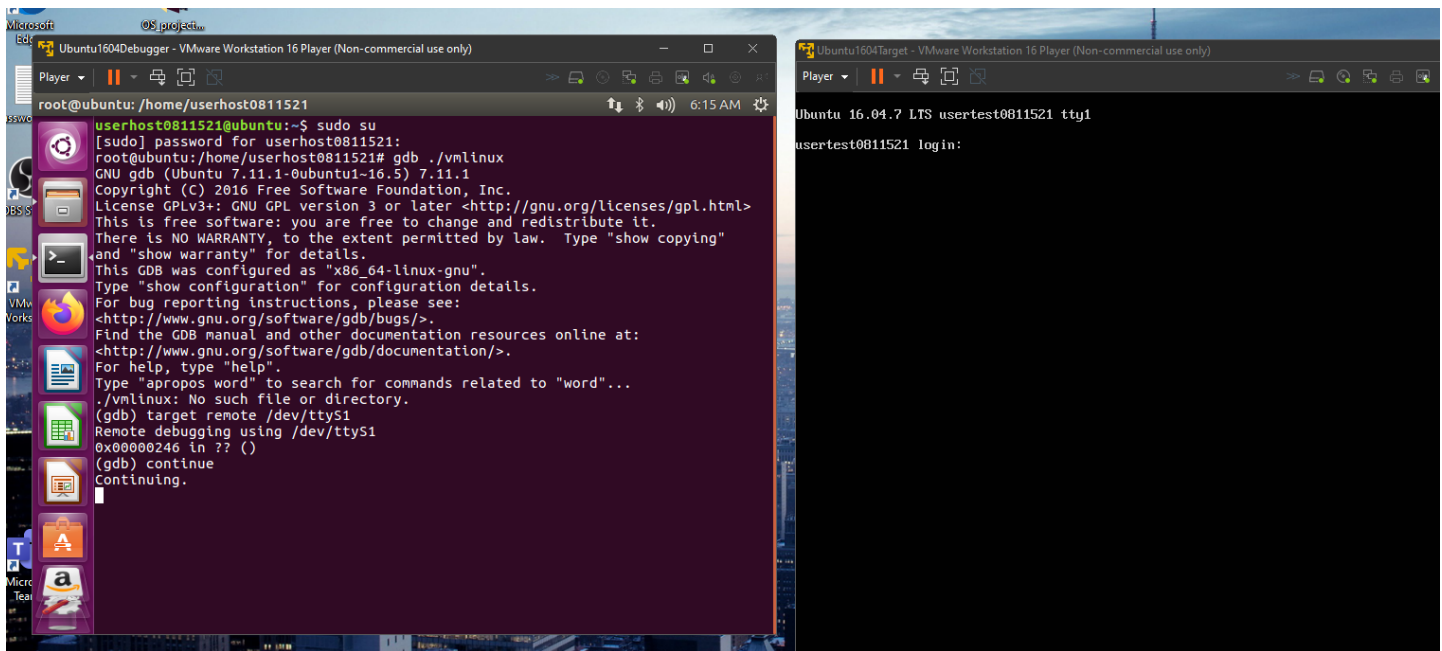


target remote /dev/ttyS1: connect to remote target machine



Screenshot #12

After I enter command **continue** in the debugger machine, the target machine resumes booting its system.



## Screenshot #13

### [Chromium OS Docs - Linux System Call Table \(googlesource.com\)](#)

#### Tables

x86\_64 (64-bit)

Compiled from [Linux 4.14.0 headers](#).

NR	syscall name	references	%rax	arg0 (%rdi)	arg1 (%rsi)	arg2 (%rdx)	arg3 (%r10)	arg4 (%r8)	arg5 (%r9)
0	read	<a href="#">man/ cs/</a>	0x00	unsigned int fd	char *buf	size_t count	-	-	-
1	write	<a href="#">man/ cs/</a>	0x01	unsigned int fd	const char *buf	size_t count	-	-	-
2	open	<a href="#">man/ cs/</a>	0x02	const char *filename	int flags	umode_t mode	-	-	-
3	close	<a href="#">man/ cs/</a>	0x03	unsigned int fd	-	-	-	-	-
4	stat	<a href="#">man/ cs/</a>	0x04	const char *filename	struct __old_kernel_stat *statbuf	-	-	-	-
5	fstat	<a href="#">man/ cs/</a>	0x05	unsigned int fd	struct __old_kernel_stat *statbuf	-	-	-	-
6	lstat	<a href="#">man/ cs/</a>	0x06	const char *filename	struct __old_kernel_stat *statbuf	-	-	-	-
7	poll	<a href="#">man/ cs/</a>	0x07	struct pollfd *ufds	unsigned int nfds	int timeout	-	-	-

### [Linux System Call Table \(nps.edu\)](#)

#### Linux System Call Table

The following table lists the **system calls** for the **Linux 2.2** kernel. It could also be thought of as an API for the interface between user space and kernel space. My motivation for making this table was to provide a reference for **system calls** and not the C library (for more information on this topic, go to <http://www.linuxassembly.org>). On the left are the numbers of the **system calls**. This number is put into the remaining registers before calling the software interrupt 'int 0x80'. After each syscall, an integer is returned in %eax.

For convenience, links go from the "Name" column to the man page for most of the **system calls**. Links to the kernel source file where each **system** call is located are linked to in the "Source" column. Links to definitions are provided for the parameters that are typedefs or structs.

%eax	Name	Source	%ebx	%ecx	%edx	%esx	%edi
1	<a href="#">sys_exit</a>	<a href="#">kernel/exit.c</a>	int	-	-	-	-
2	<a href="#">sys_fork</a>	<a href="#">arch/i386/kernel/process.c</a>	<a href="#">struct pt_regs</a>	-	-	-	-
3	<a href="#">sys_read</a>	<a href="#">fs/read_write.c</a>	unsigned int	char *	<a href="#">size_t</a>	-	-
4	<a href="#">sys_write</a>	<a href="#">fs/read_write.c</a>	unsigned int	const char *	<a href="#">size_t</a>	-	-
5	<a href="#">sys_open</a>	<a href="#">fs/open.c</a>	const char *	int	int	-	-
6	<a href="#">sys_close</a>	<a href="#">fs/open.c</a>	unsigned int	-	-	-	-
7	<a href="#">sys_waitpid</a>	<a href="#">kernel/exit.c</a>	pid_t	unsigned int *	int	-	-
8	<a href="#">sys_creat</a>	<a href="#">fs/open.c</a>	const char *	int	-	-	-
9	<a href="#">sys_link</a>	<a href="#">fs/namei.c</a>	const char *	const char *	-	-	-
10	<a href="#">sys_unlink</a>	<a href="#">fs/namei.c</a>	const char *	-	-	-	-
11	<a href="#">sys_execve</a>	<a href="#">arch/i386/kernel/process.c</a>	<a href="#">struct pt_regs</a>	-	-	-	-
12	<a href="#">sys_chdir</a>	<a href="#">fs/open.c</a>	const char *	-	-	-	-
13	<a href="#">sys_time</a>	<a href="#">kernel/time.c</a>	int *	-	-	-	-
14	<a href="#">sys_mknod</a>	<a href="#">fs/namei.c</a>	const char *	int	<a href="#">dev_t</a>	-	-
15	<a href="#">sys_chmod</a>	<a href="#">fs/open.c</a>	const char *	<a href="#">mode_t</a>	-	-	-
16	<a href="#">sys_lchown</a>	<a href="#">fs/open.c</a>	const char *	<a href="#">uid_t</a>	<a href="#">gid_t</a>	-	-

## [Linux/i386 system calls \(sourceforge.net\)](http://sourceforge.net)

### **1. *sys\_exit***

Syntax: `int sys_exit(int status)`

Source: `kernel/exit.c`

Action: terminate the current process

Details: `status` is return code

### **2. *sys\_fork***

Syntax: `int sys_fork()`

Source: `arch/i386/kernel/process.c`

Action: create a child process

Details:

### **3. *sys\_read***

Syntax: `ssize_t sys_read(unsigned int fd, char * buf, size_t count)`

Source: `fs/read_write.c`

Action: read from a file descriptor

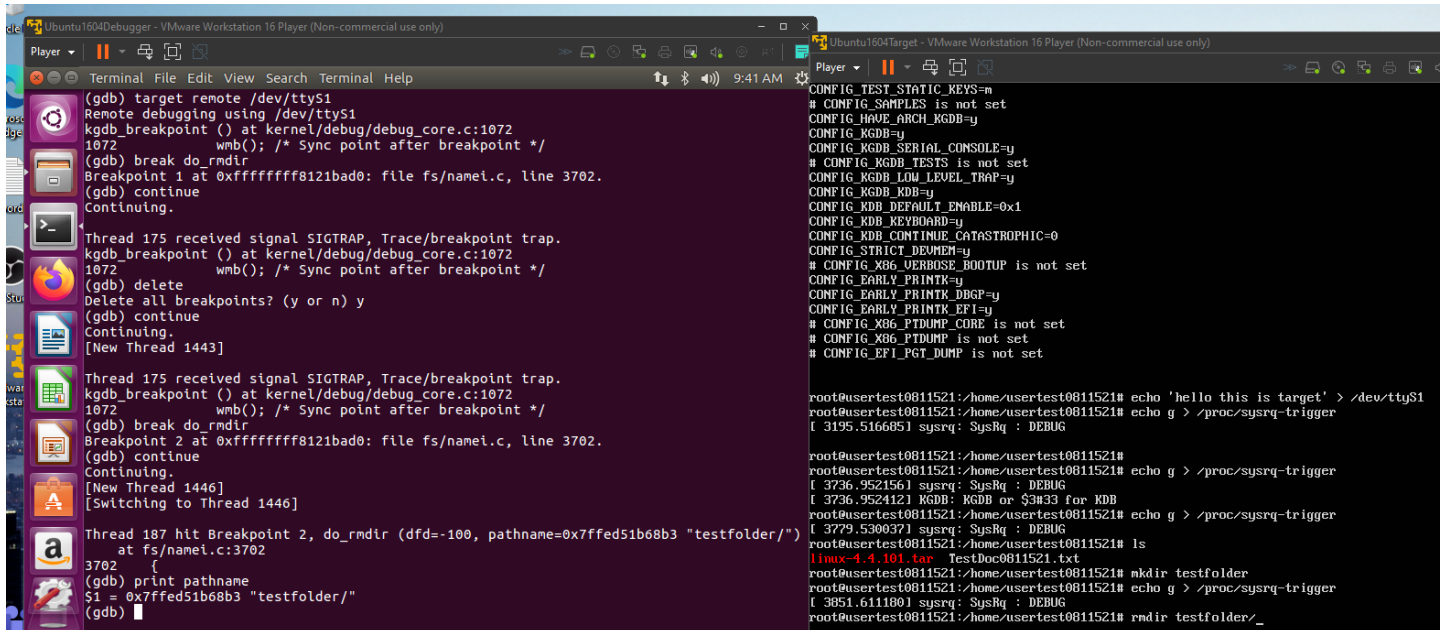
Details:



## Screenshot #14

set a breakpoint for **rmdir**

trigger the breakpoint and print the parameter *pathname*



```
(gdb) target remote /dev/ttyS1
Remote debugging using /dev/ttyS1
kgdb_breakpoint () at kernel/debug/debug_core.c:1072
1072      wmb(); /* Sync point after breakpoint */
(gdb) break do_rmdir
Breakpoint 1 at 0xfffffff8121bad0: file fs/namei.c, line 3702.
(gdb) continue
Continuing.

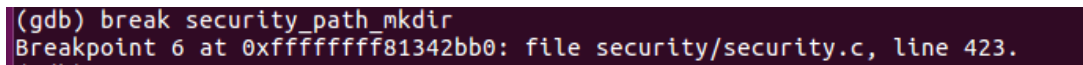
Thread 175 received signal SIGTRAP, Trace/breakpoint trap.
kgdb_breakpoint () at kernel/debug/debug_core.c:1072
1072      wmb(); /* Sync point after breakpoint */
(gdb) delete
Delete all breakpoints? (y or n) y
(gdb) continue
Continuing.
[New Thread 1443]

Thread 175 received signal SIGTRAP, Trace/breakpoint trap.
kgdb_breakpoint () at kernel/debug/debug_core.c:1072
1072      wmb(); /* Sync point after breakpoint */
(gdb) break do_rmdir
Breakpoint 2 at 0xfffffff8121bad0: file fs/namei.c, line 3702.
(gdb) continue
Continuing.
[New Thread 1446]
[Switching to Thread 1446]

Thread 187 hit Breakpoint 2, do_rmdir (dfd=-100, pathname=0x7ffed51b68b3 "testfolder/")
3702 {
(gdb) print pathname
$1 = 0x7ffed51b68b3 "testfolder/"
(gdb)
```

## Screenshot #15

set one breakpoint for **security\_path\_mkdir** that will be triggered when executing **mkdir**

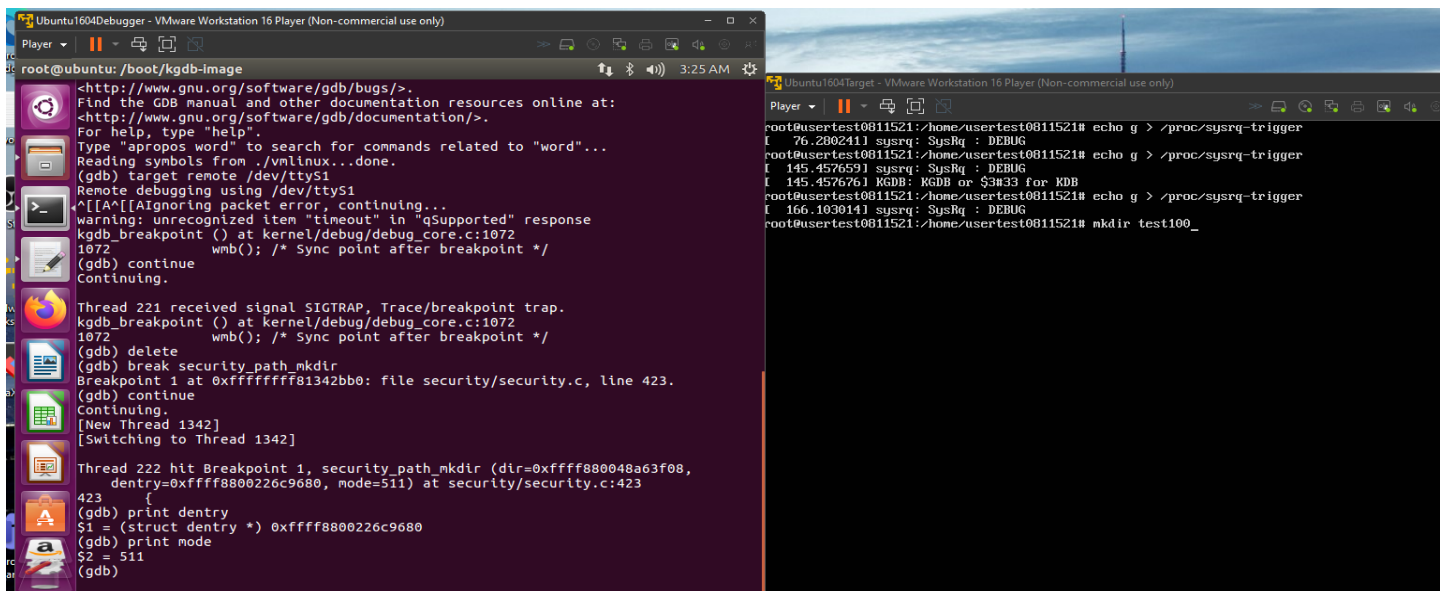


```
(gdb) break security_path_mkdir
Breakpoint 6 at 0xfffffff81342bb0: file security/security.c, line 423.
```

## Screenshot #16

parameters of **security\_path\_mkdir**: path, dentry, and mode

print some parameters of the function to gain a closer look into the information of the breakpoint



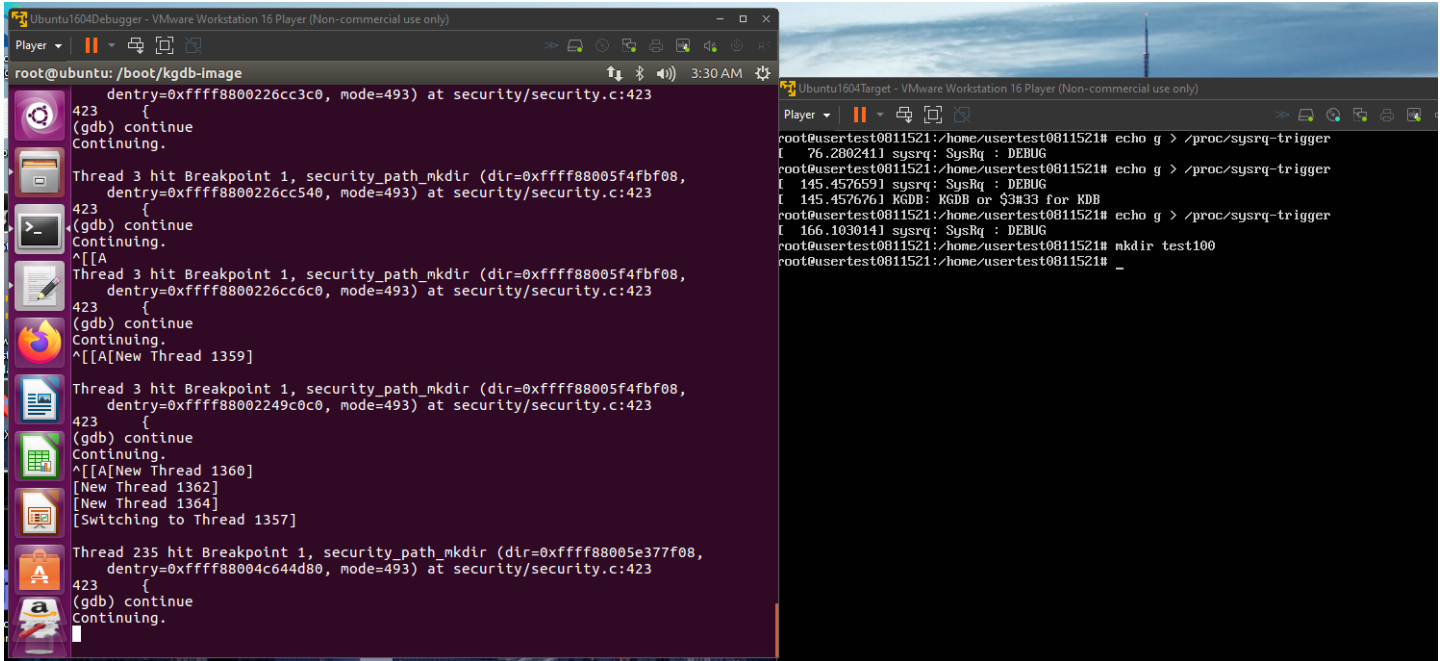
```
root@ubuntu:/boot/kgdb-image
<http://www.gnu.org/software/gdb/bugs/>.
Find the GDB manual and other documentation resources online at:
<http://www.gnu.org/software/gdb/documentation/>.
For help, type "help".
Type "apropos word" to search for commands related to "word"...
Reading symbols from /vmlinuz...done.
(gdb) target remote /dev/ttyS1
Remote debugging using /dev/ttyS1
warning: unrecognized item "timeout" in "qSupported" response
kgdb_breakpoint () at kernel/debug/debug_core.c:1072
1072      wmb(); /* Sync point after breakpoint */
(gdb) continue
Continuing.

Thread 221 received signal SIGTRAP, Trace/breakpoint trap.
kgdb_breakpoint () at kernel/debug/debug_core.c:1072
1072      wmb(); /* Sync point after breakpoint */
(gdb) delete
(gdb) break security_path_mkdir
Breakpoint 1 at 0xfffffff81342bb0: file security/security.c, line 423.
(gdb) continue
Continuing.
[New Thread 1342]
[Switching to Thread 1342]

Thread 222 hit Breakpoint 1, security_path_mkdir (dir=0xfffff880048a63f08,
dentry=0xfffff8800226c9680, mode=511) at security/security.c:423
423 {
(gdb) print dentry
$1 = (struct dentry *) 0xfffff8800226c9680
(gdb) print mode
$2 = 511
(gdb)
```

## Screenshot #17

after several **continue**, the target machine regains control.



## Screenshot #18

the perf command

```
userhost0811521@ubuntu:~$ perf

usage: perf [--version] [--help] [OPTIONS] COMMAND [ARGS]

The most commonly used perf commands are:
  annotate      Read perf.data (created by perf record) and display annotated
code
  archive       Create archive with object files with build-ids found in perf
.data file
  bench         General framework for benchmark suites
  buildid-cache Manage build-id cache.
  buildid-list  List the buildids in a perf.data file
  c2c           Shared Data C2C/HITM Analyzer.
  config        Get and set variables in a configuration file.
  data          Data file related processing
  diff          Read perf.data files and display the differential profile
  evlist        List the event names in a perf.data file
  ftrace        Simple wrapper for kernel's ftrace functionality
  inject        Filter to augment the events stream with additional informati
on
  kallsyms      Searches running kernel for symbols
  kmem          Tool to trace/measure kernel memory properties
  kvm           Tool to trace/measure kvm guest os
  list          List all symbolic event types
  lock          Analyze lock events
  mem           Profile memory accesses
```

```
userhost0811521@ubuntu:~$ perf trace -help

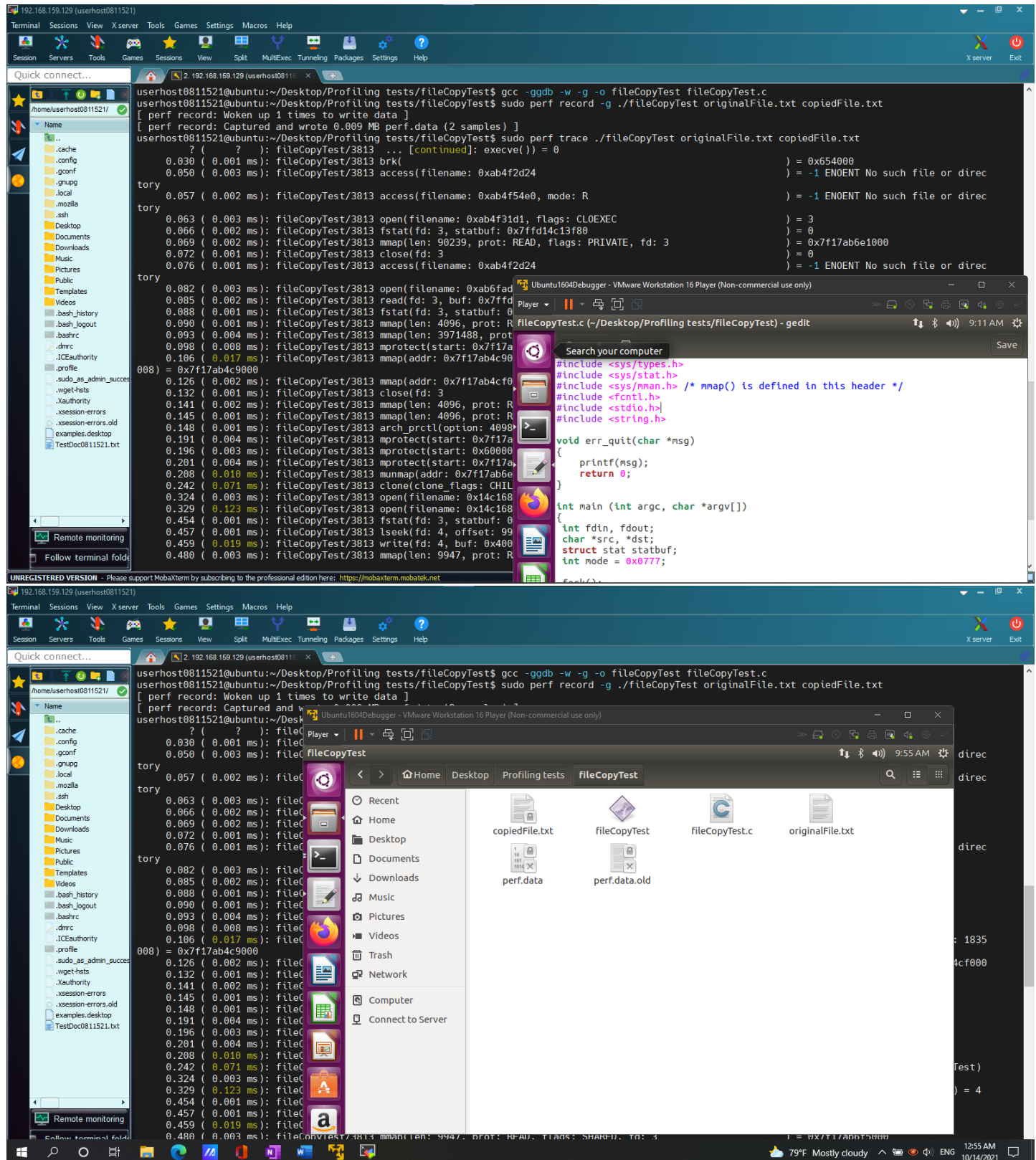
Usage: perf trace [<options>] [<command>]
or: perf trace [<options>] -- <command> [<options>]
or: perf trace record [<options>] [<command>]
or: perf trace record [<options>] -- <command> [<options>]

-a, --all-cpus      system-wide collection from all CPUs
-C, --cpu <cpu>    list of cpus to monitor
-D, --delay <n>    ms to wait before starting measurement after program s
-e, --event <event> event/syscall selector. use 'perf list' to list availa
-f, --force         don't complain, do it
-F, --pf <all|maj|min> Trace pagefaults
-i, --input <file>  Analyze events in file
-m, --mmap-pages <pages> number of mmap data pages
-o, --output <file> output file name
-p, --pid <pid>    trace events on existing process id
-s, --summary       Show only syscall summary with statistics
-S, --with-summary  Show all syscalls and summary with statistics
-t, --tid <tid>    trace events on existing thread id
-T, --time         show full timestamps, not time relative to first event
```

## Screenshot #19

This is the execution of **fileCopyTest.c**.

the execution result of perf, the trace result, and the generated files



## Screenshot #20

compare the trace results of **emptyTest.c** and **fileCopyTest.c**

Since **fileCopyTest.c** executes a few more commands such as calling **fork()**, so its trace result shows additional events.

```
1 0.000 (0.000 ms): emptyTest/3997 ... [continued]: execve() = 0
2 0.031 (0.001 ms): emptyTest/3997 brk( ) = 0x1497000
3 0.055 (0.005 ms): emptyTest/3997 access(filename: 0x2c2cf6d4) = -1 EACCES
4 0.067 (0.003 ms): emptyTest/3997 access(filename: 0x2c2cf6d4, mode: R) = -1 EACCES
5 0.074 (0.004 ms): emptyTest/3997 open(filename: 0x2c2cf6d4, flags: CLOEXEC) = 3
6 0.079 (0.002 ms): emptyTest/3997 fstat(fds: 3, statbuf: 0x7f6c0a4e60) = 0
7 0.082 (0.003 ms): emptyTest/3997 mmap(len: 90239, prot: READ, flags: PRIVATE, fd: 3) = 0x7f2fb2ee
8 0.084 (0.001 ms): emptyTest/3997 close(fds: 3) = 0
9 0.092 (0.002 ms): emptyTest/3997 access(filename: 0x2c2cf6d4) = -1 EACCES
10 0.099 (0.004 ms): emptyTest/3997 open(filename: 0x2c2cf6d4, flags: CLOEXEC) = 3
11 0.104 (0.002 ms): emptyTest/3997 read(fds: 3, buf: 0x7f6c0a4e60, count: 832) = 832
12 0.108 (0.001 ms): emptyTest/3997 fstat(fds: 3, statbuf: 0x7f6c0a4e60) = 0
13 0.110 (0.002 ms): emptyTest/3997 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f2fb2ee
14 0.114 (0.005 ms): emptyTest/3997 mmap(len: 3971480, prot: EXEC|READ, flags: PRIVATE|DENYWRITE, fd: 3) = 0x7f2fb290
15 0.132 (0.009 ms): emptyTest/3997 mprotect(start: 0x7f2fb29000, len: 2097152) = 0
16 0.143 (0.006 ms): emptyTest/3997 mmap(addr: 0x7f2fb29000, len: 24576, prot: READ|WRITE, flags: PRIVATE|DENYWRITE|FD) = 0x7f2fb290
17 0.153 (0.003 ms): emptyTest/3997 mmap(addr: 0x7f2fb29000, len: 14752, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS|FD) = 0
18 0.169 (0.002 ms): emptyTest/3997 close(fds: 3) = 0
19 0.181 (0.002 ms): emptyTest/3997 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f2fb2ee
20 0.185 (0.001 ms): emptyTest/3997 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f2fb2ee
21 0.190 (0.001 ms): emptyTest/3997 arch_prctl(option: 4096, arg2: 13984282146560) = 0
22 0.232 (0.005 ms): emptyTest/3997 mprotect(start: 0x7f2fb29000, len: 16384, prot: READ) = 0
23 0.239 (0.003 ms): emptyTest/3997 mprotect(start: 0x600000, len: 4096, prot: READ) = 0
24 0.245 (0.006 ms): emptyTest/3997 mprotect(start: 0x7f2fb29000, len: 4096, prot: READ) = 0
25 0.252 (0.009 ms): emptyTest/3997 munmap(addr: 0x7f2fb2ee5000, len: 90239) = 0
26
27
28
29
30
31
32
33
34 0.328 (0.000 ms): emptyTest/3997 exit_group( )
35
36
```

```
1 0.085 (0.008 ms): fileCopyTest/4055 ... [continued]: execve() = 0
2 0.143 (0.010 ms): fileCopyTest/4055 brk( ) = 0x173400
3 0.166 (0.007 ms): fileCopyTest/4055 access(filename: 0x2b0191d24) = -1 EACCES
4 0.181 (0.008 ms): fileCopyTest/4055 access(filename: 0x2b0191d24, mode: R) = -1 EACCES
5 0.193 (0.008 ms): fileCopyTest/4055 open(filename: 0x2b0191d24, flags: CLOEXEC) = 3
6 0.200 (0.008 ms): fileCopyTest/4055 fstat(fds: 3, statbuf: 0x7f6c0a4e60) = 0
7 0.212 (0.003 ms): fileCopyTest/4055 mmap(len: 90239, prot: READ, flags: PRIVATE, fd: 3) = 0x7f6c0b
8 0.220 (0.006 ms): fileCopyTest/4055 close(fds: 3) = 0
9 0.224 (0.008 ms): fileCopyTest/4055 access(filename: 0x2b0191d24) = -1 EACCES
10 0.244 (0.008 ms): fileCopyTest/4055 open(filename: 0x2b0191d24, flags: CLOEXEC) = 3
11 0.256 (0.006 ms): fileCopyTest/4055 read(fds: 3, buf: 0x7f6c0a4e60, count: 832) = 832
12 0.267 (0.004 ms): fileCopyTest/4055 fstat(fds: 3, statbuf: 0x7f6c0a4e60) = 0
13 0.274 (0.004 ms): fileCopyTest/4055 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f6c0b
14 0.291 (0.011 ms): fileCopyTest/4055 mmap(len: 3971480, prot: EXEC|READ, flags: PRIVATE|DENYWRITE, fd: 3) = 0x7f6c0b
15 0.306 (0.020 ms): fileCopyTest/4055 mprotect(start: 0x7f6c0b0000, len: 2097152) = 0
16 0.328 (0.058 ms): fileCopyTest/4055 mmap(addr: 0x7f6c0b0000, len: 24576, prot: READ|WRITE, flags: PRIVATE|DENYWRITE) = 0x7f6c0b
17 0.407 (0.008 ms): fileCopyTest/4055 mmap(addr: 0x7f6c0b0000, len: 14752, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0
18 0.430 (0.004 ms): fileCopyTest/4055 close(fds: 3) = 0
19 0.458 (0.007 ms): fileCopyTest/4055 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f6c0b
20 0.473 (0.005 ms): fileCopyTest/4055 mmap(len: 4096, prot: READ|WRITE, flags: PRIVATE|ANONYMOUS) = 0x7f6c0b
21 0.484 (0.003 ms): fileCopyTest/4055 arch_prctl(option: 4096, arg2: 14048566602016) = 0
22 0.615 (0.014 ms): fileCopyTest/4055 mprotect(start: 0x7f6c0b0000, len: 16384, prot: READ) = 0
23 0.635 (0.009 ms): fileCopyTest/4055 mprotect(start: 0x600000, len: 4096, prot: READ) = 0
24 0.652 (0.012 ms): fileCopyTest/4055 mprotect(start: 0x7f6c0b397000, len: 4096, prot: READ) = 0
25 0.669 (0.024 ms): fileCopyTest/4055 munmap(addr: 0x7f6c0b397000, len: 90239) = 0
26 0.754 (0.177 ms): fileCopyTest/4055 clone(clone_flags: CLD_CLEARTID|CLD_SETTID(0x1), child_tidptr: 0x7f6c0b397e40) = 3
27 1.016 (0.011 ms): fileCopyTest/4055 open(filename: 0x6f6607fc, flags: RDWR|CREAT|TRUNC, mode: IRUGO|ISGID|SVTX|IXUSR) = 0
28 1.030 (0.100 ms): fileCopyTest/4055 open(filename: 0x6f6607fc, flags: RDWR|CREAT|TRUNC, mode: IRUGO|ISGID|SVTX|IXUSR) = 0
29 1.137 (0.005 ms): fileCopyTest/4055 fstat(fds: 3, statbuf: 0x7f6c0b397e40) = 0
30 1.146 (0.004 ms): fileCopyTest/4055 lseek(fds: 4, offset: 8946, whence: SET) = 8946
31 1.154 (0.025 ms): fileCopyTest/4055 write(fds: 4, buf: 0x400b77, count: 1) = 1
32 1.184 (0.011 ms): fileCopyTest/4055 mmap(len: 9947, prot: READ, flags: SHARED, fd: 3) = 0x7f6c0b
33 1.198 (0.008 ms): fileCopyTest/4055 mmap(len: 9947, prot: READ|WRITE, flags: SHARED, fd: 4) = 0x7f6c0b
34 1.861 (0.000 ms): fileCopyTest/4055 exit_group( )
35
36
```

(p.32)

a. Will the functions' execution time be longer if the file is bigger?

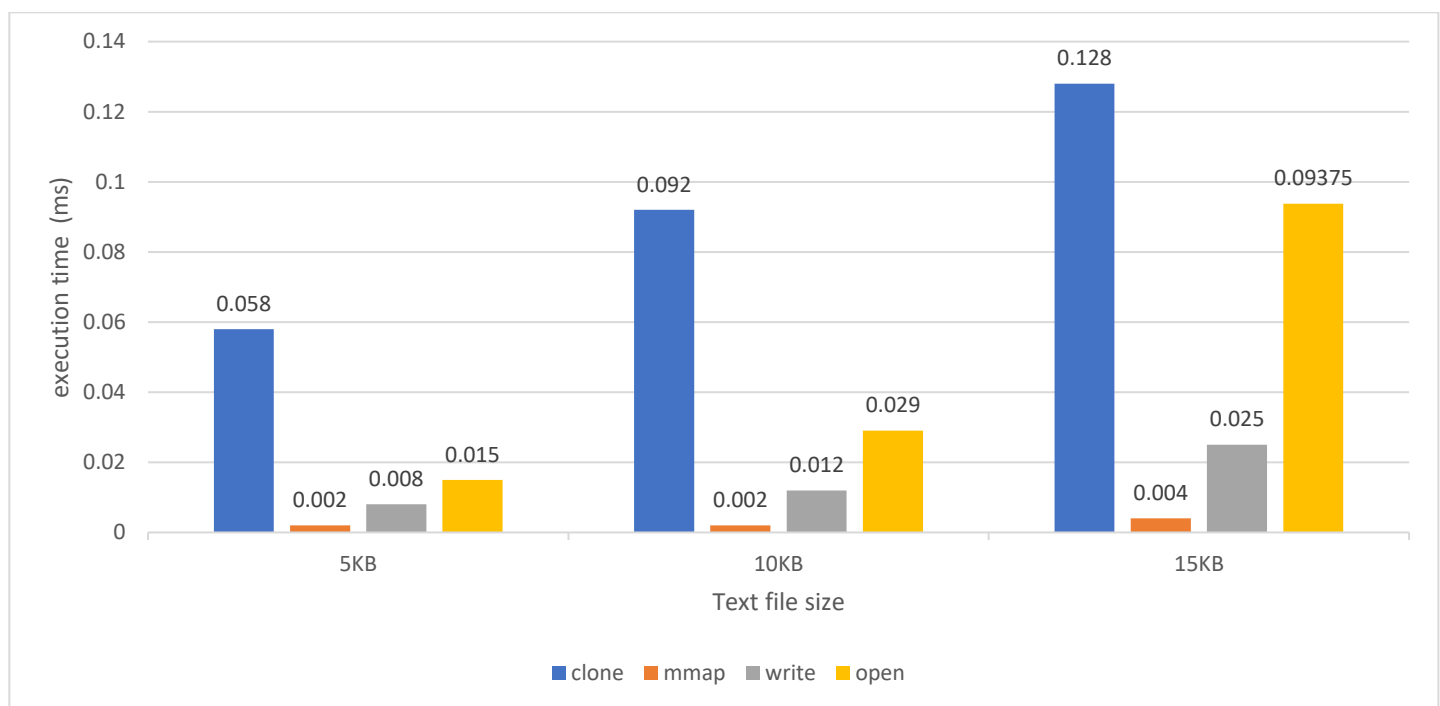
Ans: According to my experiment, the executions of **clone**, **open**, and **write** take longer time when the text file is larger. The executions of **mmap** for different file size take roughly the same amount of time. But there is some time when the executions of same function take same amount of time for different file sizes; I guess the file sizes are not large enough to generate distinct differences.

b. How is the behavior of each function? Sort them from slowest to fastest.

Ans: from fastest to slowest: **mmap**, **write**, **open**, **clone**

c. Create a graph of file size (in bytes) vs. execution time (ms) of these four functions, using 3 different file sizes.

Screenshot #21



d. Perf also has the report command:

i. What is it for?

Ans: Show details of performance of processes and overhead of each function



ii. For fileCopyTest, show and interpret the results.

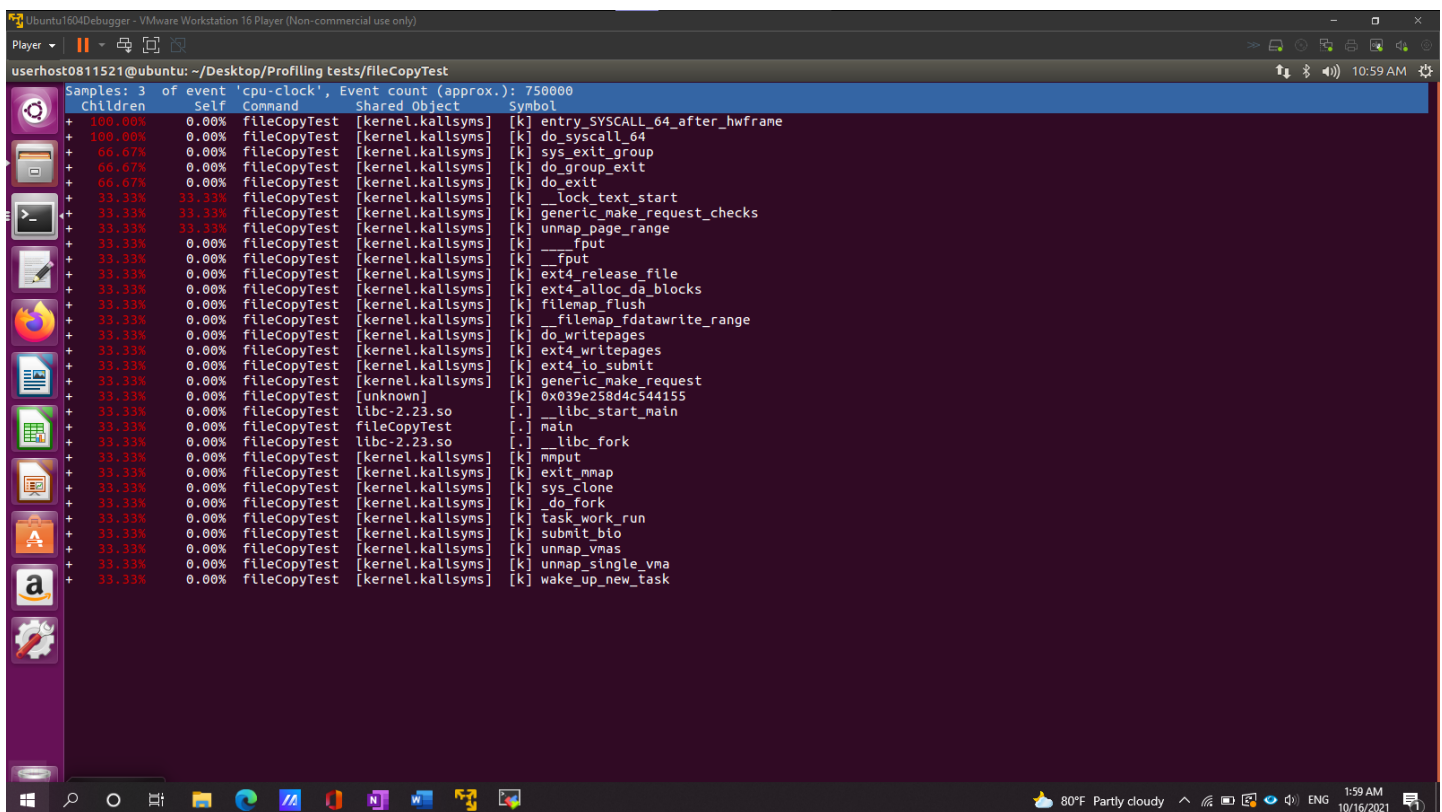
Ans: The following functions have the highest overhead:

\_\_lock\_text\_start

generic\_make\_request\_checks

unmap\_page\_range

## Screenshot #22 [perf report]



e. Perf has more commands (Section 5.1 step 4). Select another command (besides report, trace and record), explain what it is for and show how to use it. Create your own scenario.

Ans: **perf top**

function: real-time analysis of performance

## Screenshot #23 [perf top]

```
root@ubuntu: /home/userhost0811521/Desktop/Profiling tests/fileCopyTest
Samples: 13K of event 'cpu-clock', Event count (approx.): 420726609
Overhead Shared Object Symbol
17.04% [kernel] [k] __lock_text_start
8.62% [kernel] [k] vmw_cndbuf_header_submit
6.49% libslang.so.2.3.0 [.] SLsng_write_chars
2.40% [kernel] [k] eventfd_write
2.20% libc-2.23.so [.] _int_malloc
1.97% [kernel] [k] queue_work_on
1.70% [kernel] [k] _softirqentry_text_start
1.33% Xorg [.] 0x00000000000062612
0.95% [kernel] [k] finish_task_switch
0.90% libglib-2.0.so.0.4800.2 [.] g_hash_table_lookup
0.71% libc-2.23.so [.] __memcpy_avx_unaligned
0.69% libc-2.23.so [.] vfprintf
0.67% libc-2.23.so [.] malloc
0.64% libc-2.23.so [.] _int_free
0.58% libc-2.23.so [.] __memcpy_sse4_1
0.52% libc-2.23.so [.] __strcmp_sse2_unaligned
0.52% perf [.] rb_next
0.51% libpixman-1.so.0.33.6 [.] pixman_region_selfcheck
0.51% libslang.so.2.3.0 [.] SLtt_smart_puts
0.48% libpthread-2.23.so [.] pthread_mutex_lock
0.47% libglib-2.0.so.0.4800.2 [.] g_mutex_lock
0.46% libglib-2.0.so.0.4800.2 [.] g_main_context_check
0.41% [kernel] [k] do_syscall_64
0.41% libgobject-2.0.so.0.4800.2 [.] g_type_check_instance_is_a
0.40% [kernel] [k] tick_nohz_idle_enter
0.37% [kernel] [k] exit_to_usermode_loop
0.36% [kernel] [k] do_sys_poll
0.36% libglib-2.0.so.0.4800.2 [.] 0x00000000000047504
0.36% libslang.so.2.3.0 [.] 0x000000000000a1a1
0.36% perf [.] _hists_insert_output_entry
0.36% libslang.so.2.3.0 [.] 0x00000000000091bb
0.36% libslang.so.2.3.0 [.] 0x000000000000913b
0.35% libslang.so.2.3.0 [.] 0x0000000000009a210
0.31% libc-2.23.so [.] __memset_sse2
0.31% libslang.so.2.3.0 [.] 0x0000000000009a107
0.31% libslang.so.2.3.0 [.] 0x0000000000009a23e
0.31% perf [.] dso_find_symbol
0.30% libslang.so.2.3.0 [.] 0x0000000000009a227
0.30% Xorg [.] GrabMatchesSecond
0.30% libslang.so.2.3.0 [.] SLsng_write_wrapped_string
0.30% libslang.so.2.3.0 [.] 0x0000000000009a25a
0.30% libslang.so.2.3.0 [.] 0x0000000000009a20a
0.30% Xorg [.] GetMaster
0.30% libslang.so.2.3.0 [.] 0x0000000000009a257
0.30% libslang.so.2.3.0 [.] SLsng_refresh
For a higher level overview, try: perf top --sort comm,dso
```

## Screenshot #24 [zoom into DSO loop]

I execute a file with infinite while loops that keep calling a function foo().

```
Samples: 325K of event 'cpu-clock', Event count (approx.): 14184957782, DSO: loop
Overhead Symbol
8.08% [.] foo
0.08% [.] main
0.01% [.] printf@plt
```



(Below) In the main function, we can inspect the parts that occupy the resources.

```
Ubuntu1604Debugger - VMware Workstation T6 Player (Non-commercial use only)
Player
userhost0811521@ubuntu: ~
main /home/userhost0811521/Desktop/Profiling tests/loop/loop
Percent

Disassembly of section .text:
0000000000400540 <main>:
main():
{}

int main(){
    push    %rbp
    mov     %rsp,%rbp
    sub     $0x10,%rsp

    int i=0;
    movl    $0x0,-0x4(%rbp)
    while(1){
        i++;
        f:  addl    $0x1,-0x4(%rbp)
        printf("print");
        mov     $0x4005f4,%edi
        mov     $0x0,%eax
        → callq  printf@plt
        foo();
        mov     $0x0,%eax
        → callq  foo
    }
    jmp     f

9.42 f:  addl    $0x1,-0x4(%rbp)
56.50 printf("print");
7.40 mov     $0x4005f4,%edi
8.30 mov     $0x0,%eax
7.85 → callq  printf@plt
10.54 → callq  foo
```

Sources:

[1]: [https://www.trendmicro.com/en\\_us/research/17/a/practical-android-debugging-via-kgdb.html](https://www.trendmicro.com/en_us/research/17/a/practical-android-debugging-via-kgdb.html)