Operating Systems Project Report

Project Number (01 / 02 / 03):	01
Name:	蕭望緯
Student ID:	0811521
YouTube link (Format youtube.com/watch?v=[key]):	https://youtu.be/Y0dAnP7n9mw
Date (YYYY-MM-DD):	2021-10-20
Names of the files uploaded to E3:	OS_Project01_0811521.pdf
Physical Machine Total RAM (Example: 8.0 GB):	16GB
Physical Machine CPU (Example: Intel i7-2600K):	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

KDGB: 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

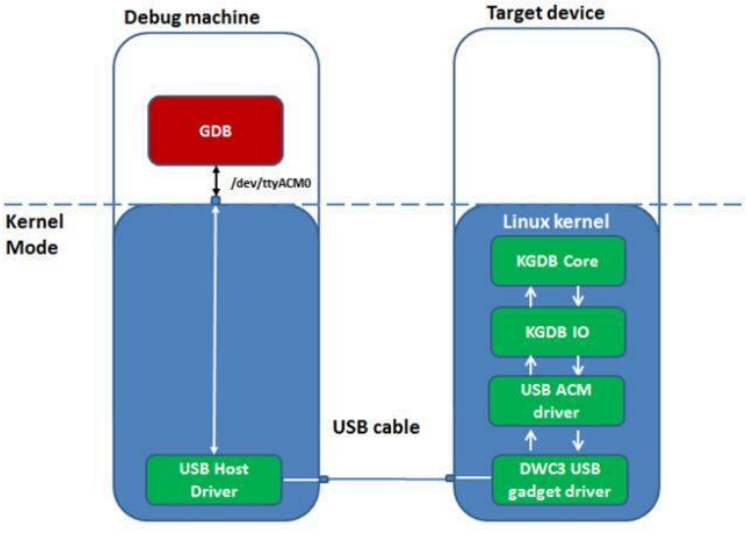
/home: personal workspace for users where the users have full access to writing and creating files

/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:

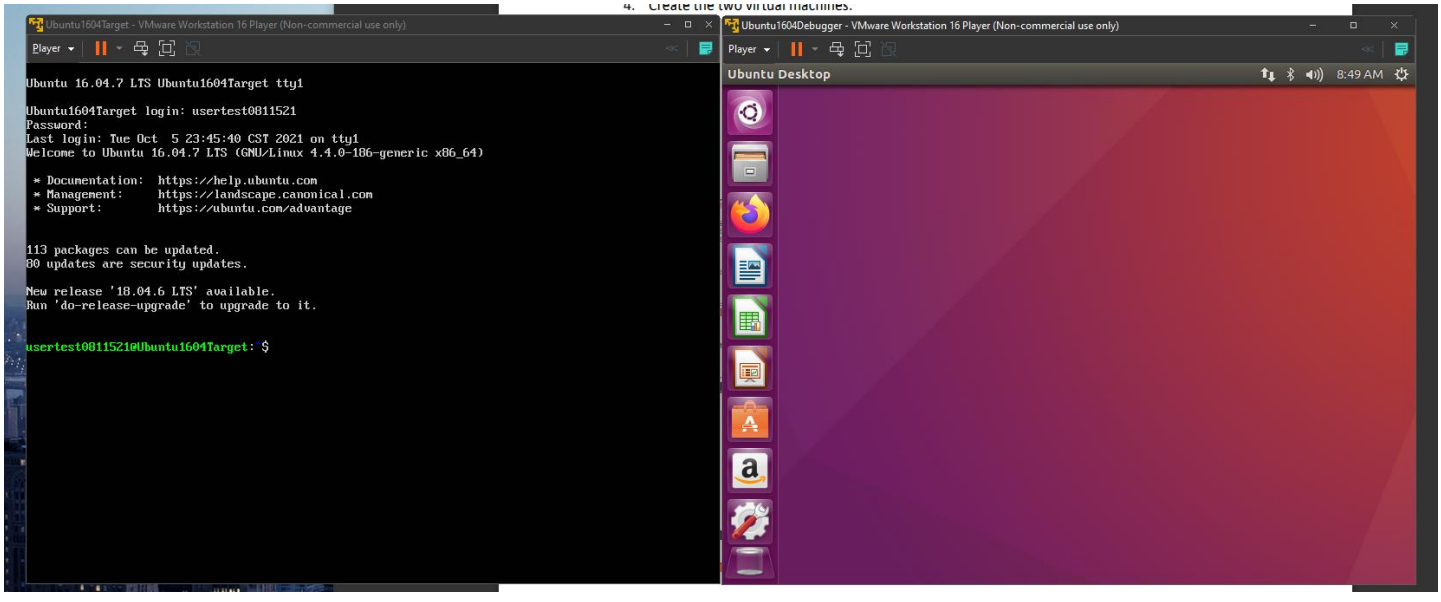
fork will call **clone** system call and do similar things like **clone.**

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

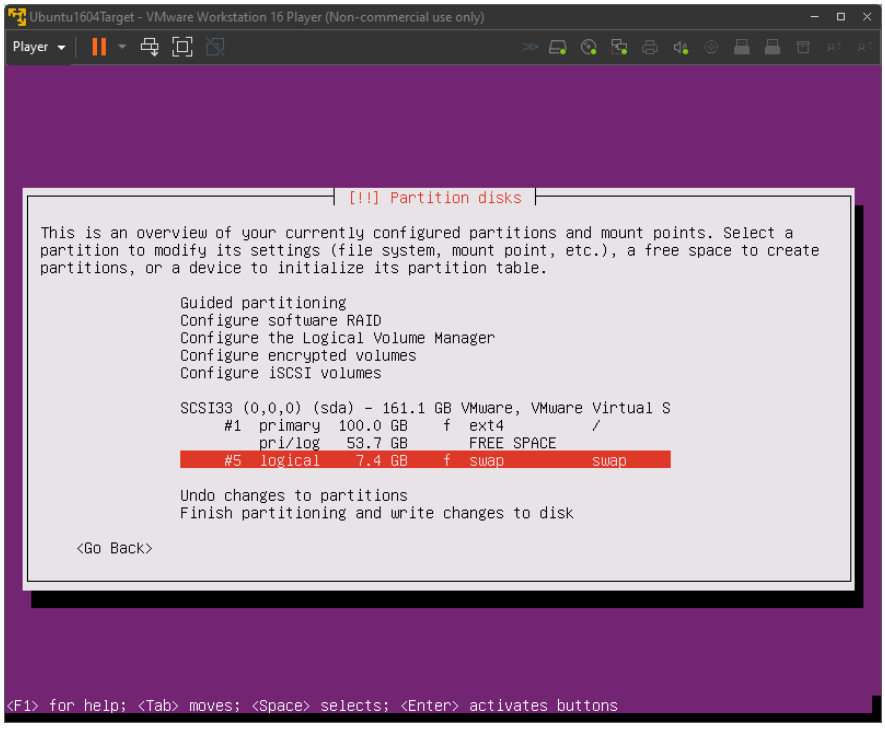
clone: the created child process can share resources with the parent

Screenshot #1

Target(left) & Debugger(right) Machine



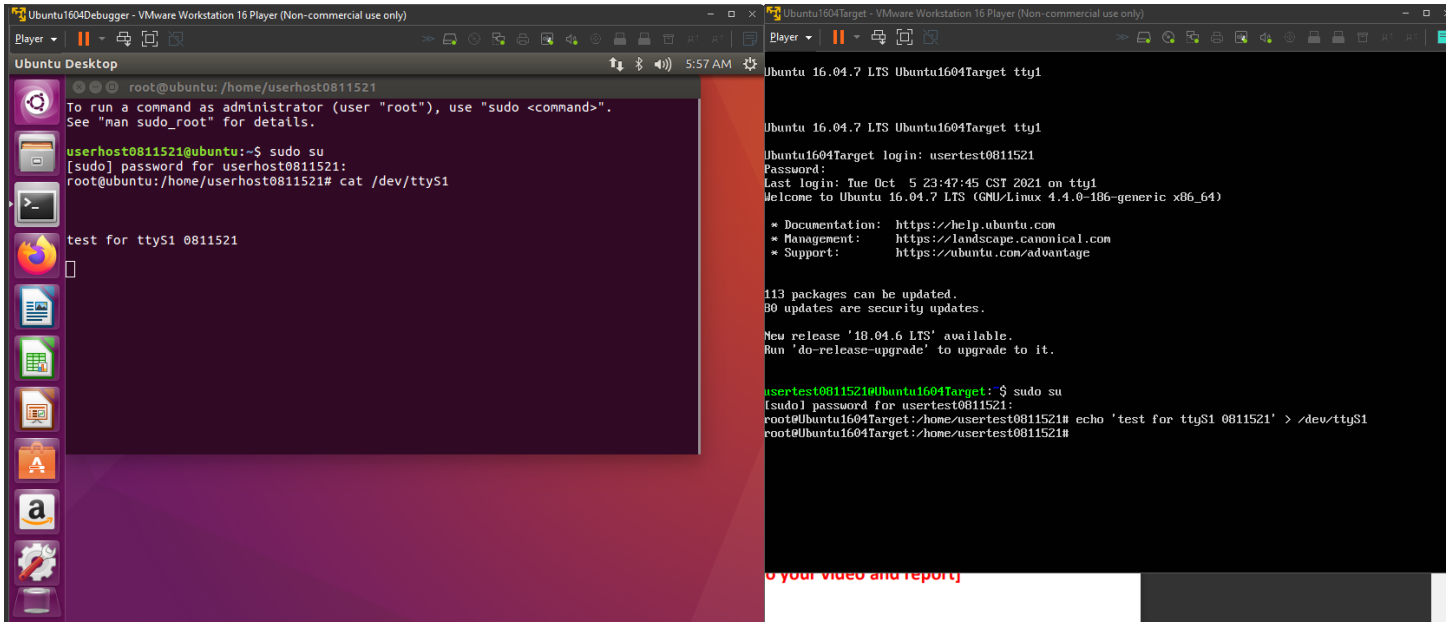
Target machine disks configuration



Screenshot #2

Serial port communication testing after some configuration

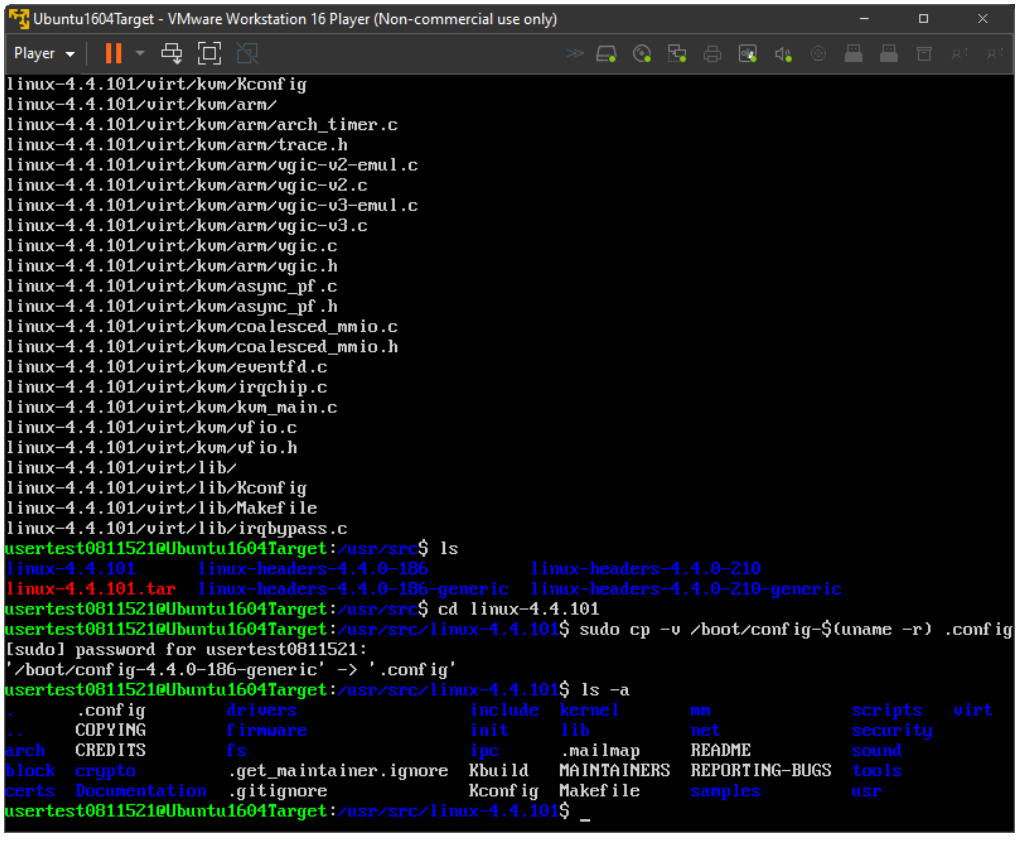
Message passed from Target to Debugger successfully through serial port communication



Screenshot #3

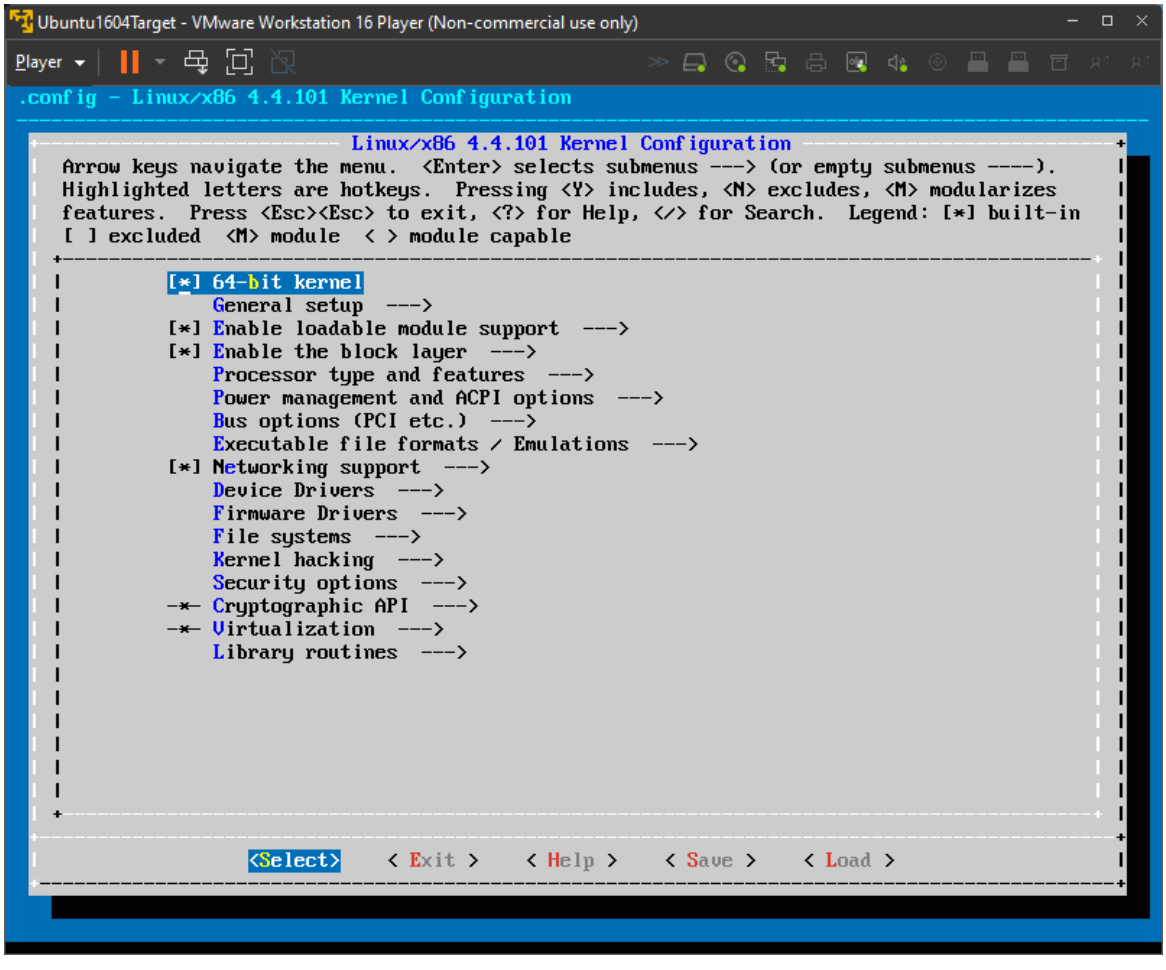
Use the predefined configuration from one of the other versions of kernel

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



Screenshot #4

Open Kernel configuration interface and examine the setup

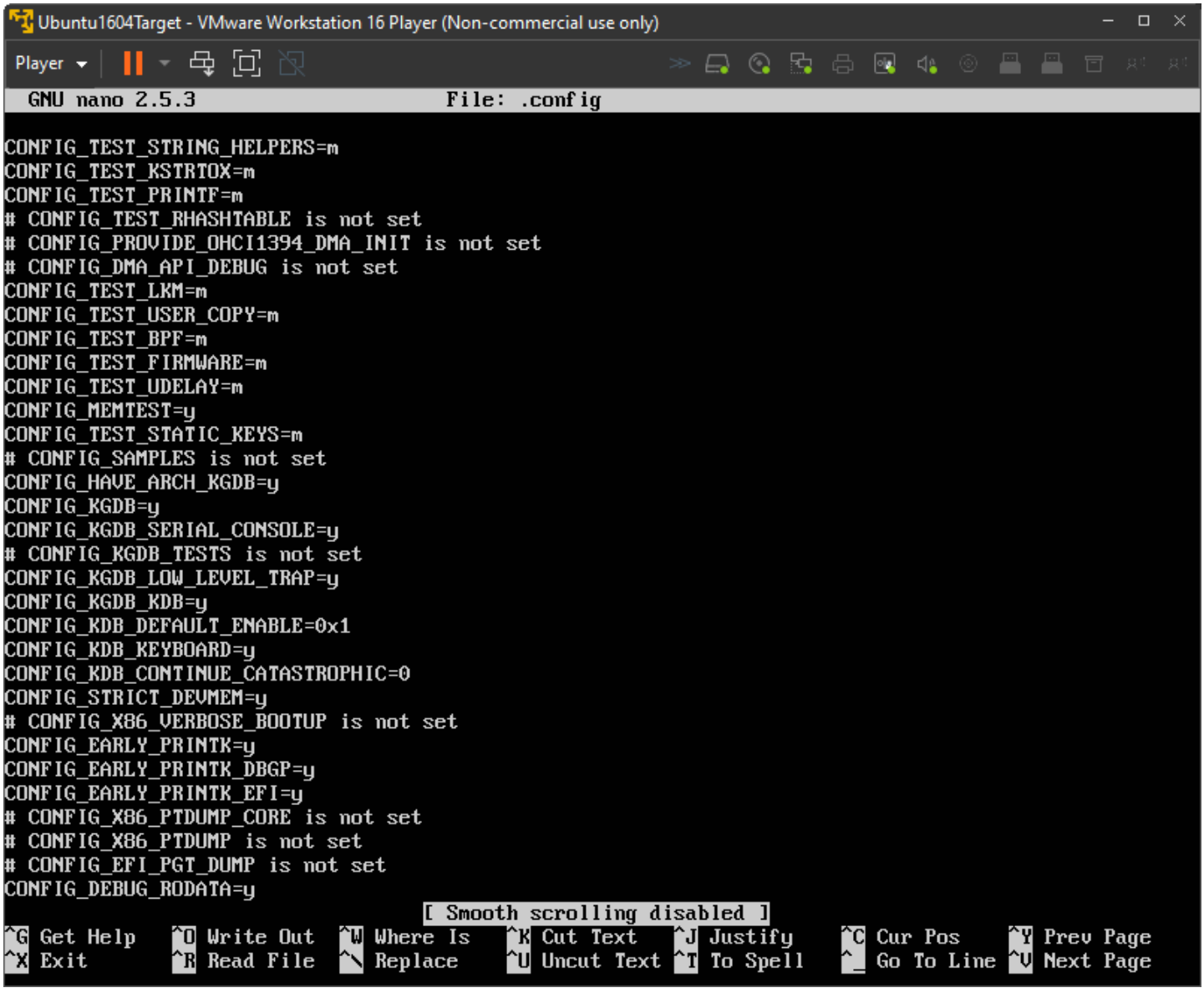


Screenshot #5

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:

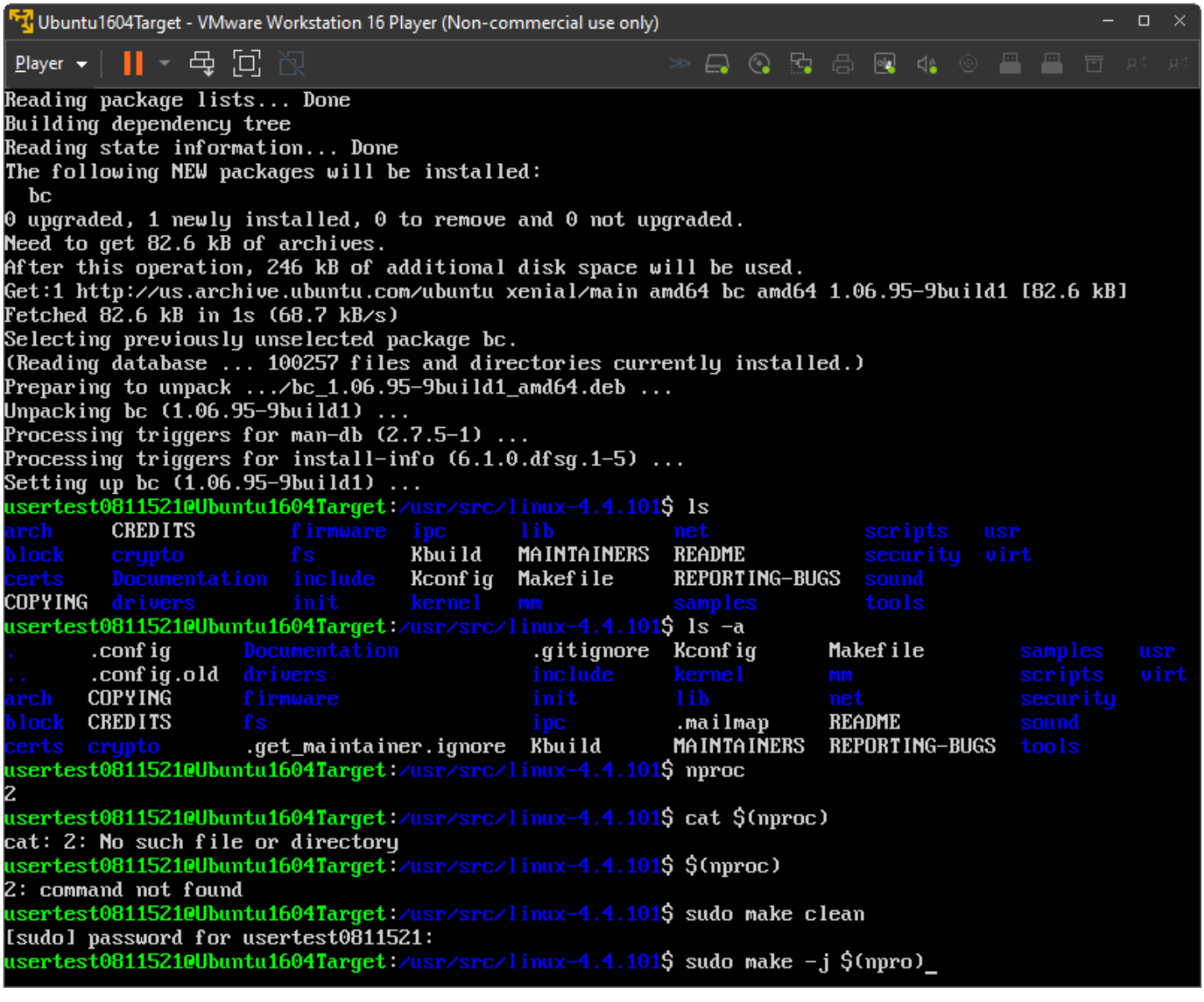
```
CONFIG_FRAME_POINTER=y
CONFIG_KGDB=y
CONFIG_KGDB_SERIAL_CONSOLE=y
CONFIG_KGDB_KDB=y
CONFIG_KDB_KEYBOARD=y
```



Screenshot #6

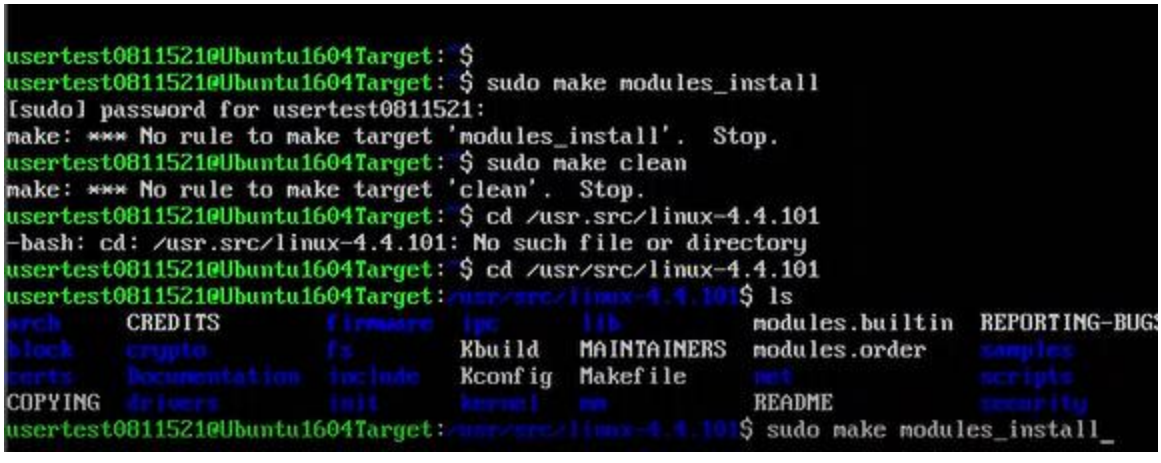
Start building the kernel.

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



Screenshot #7

After kernel is compiled, install kernel modules



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-4.4.10
1
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$
```

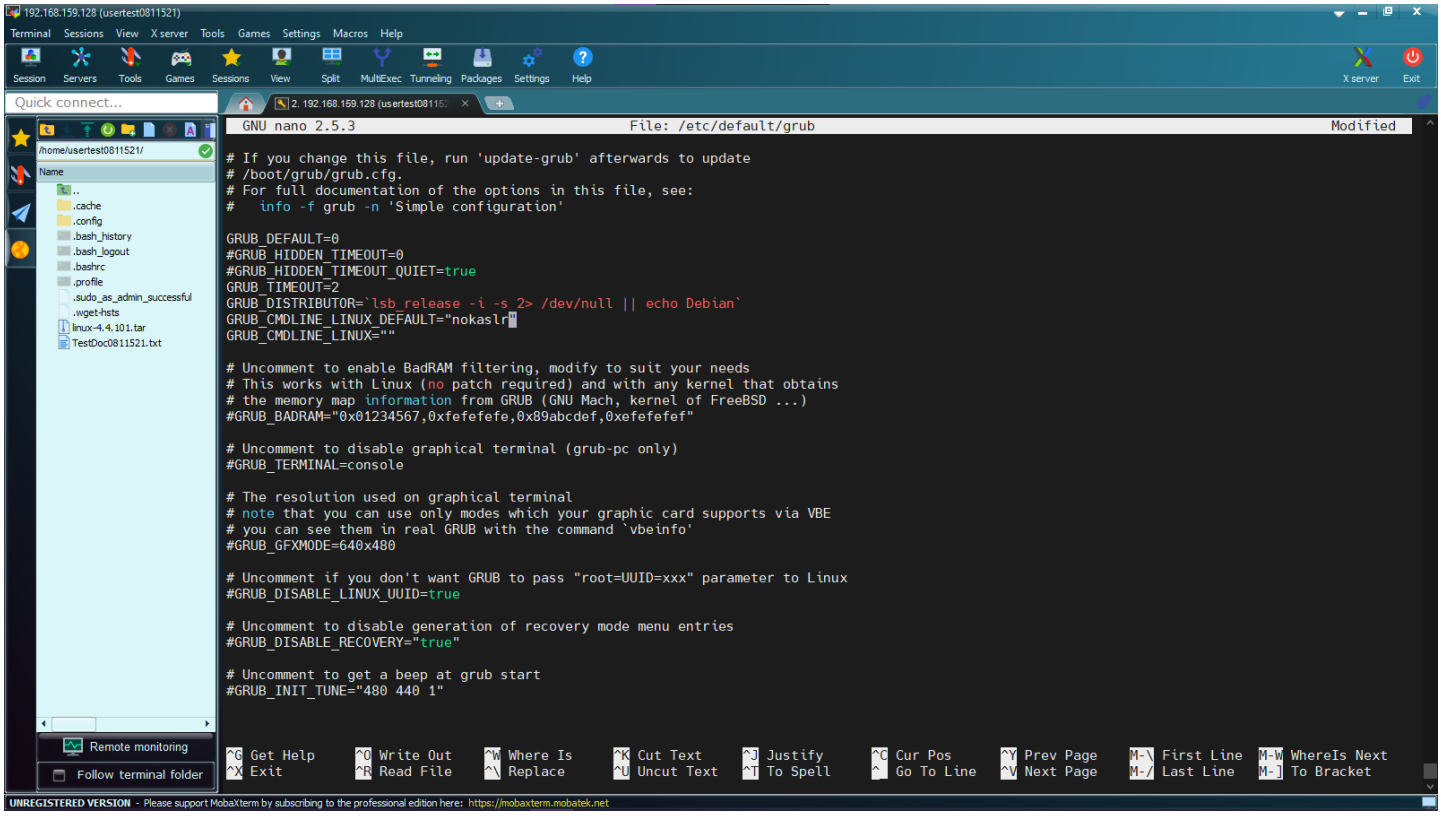
Screenshot #9

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

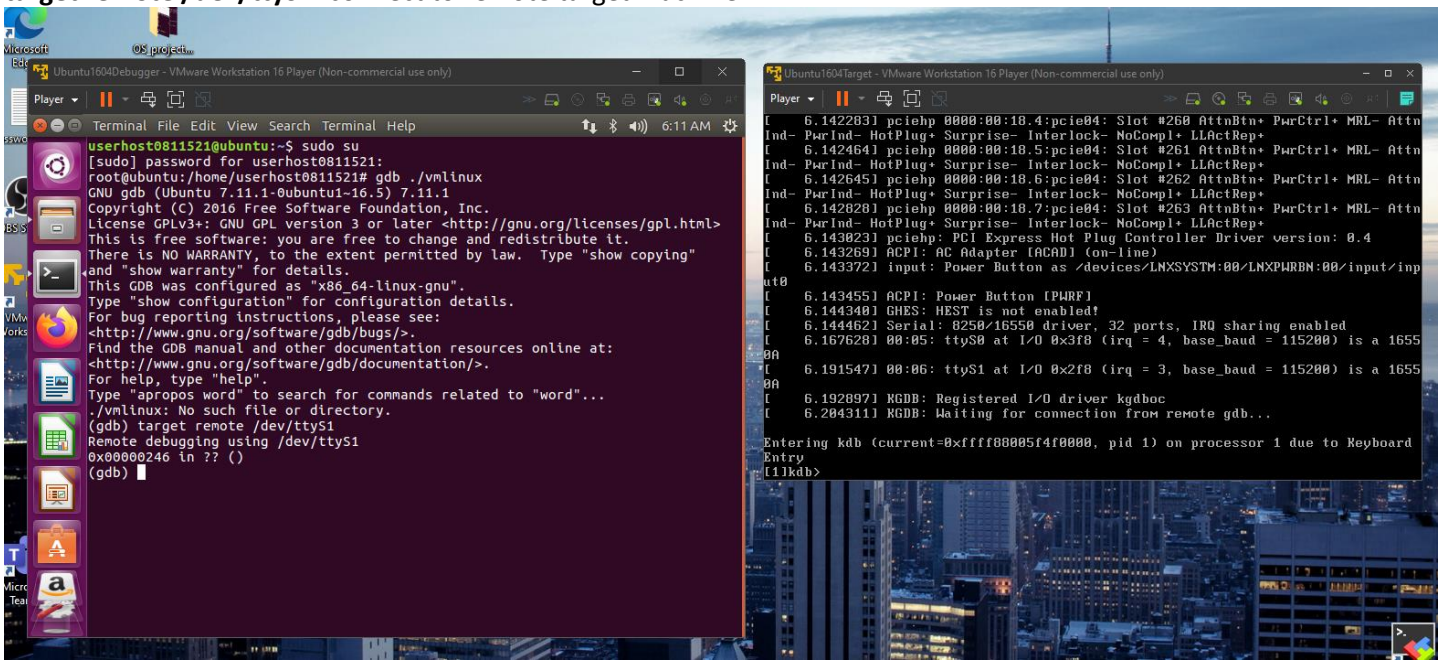


Screenshot #11

gdb ./vmlinux: open gdb

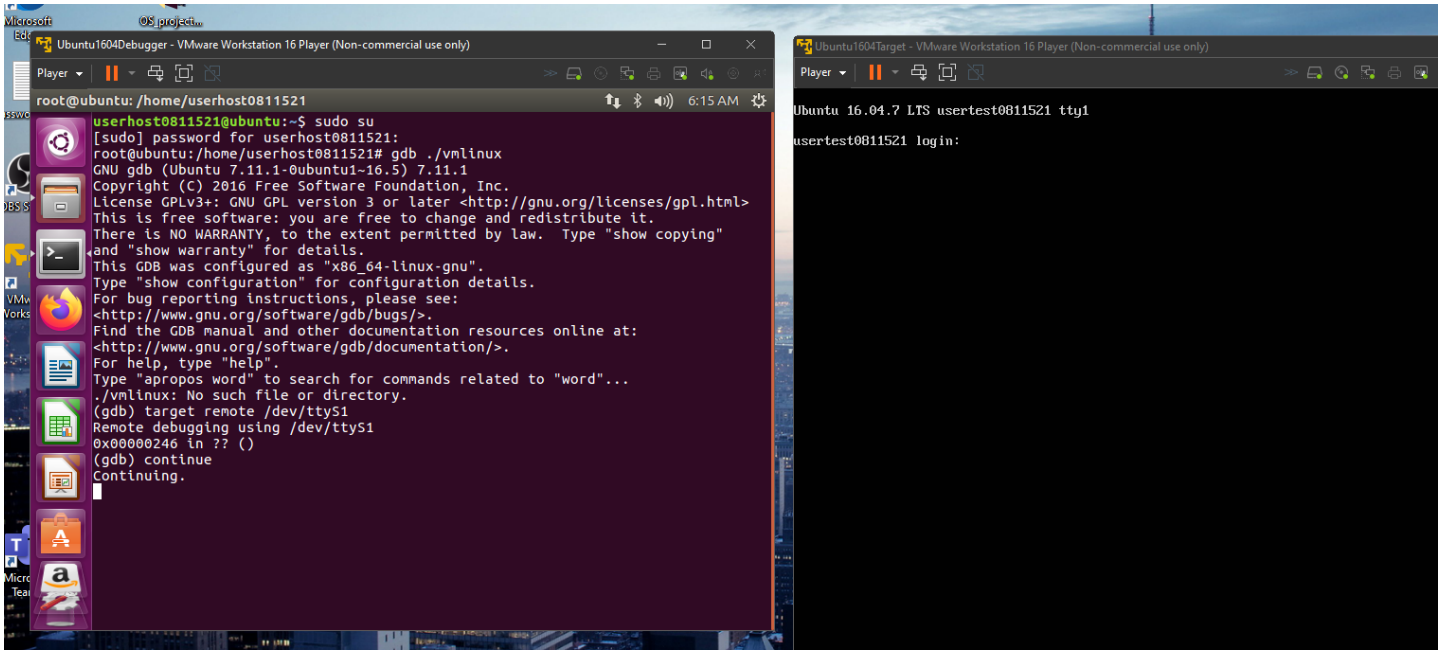
```
root@ubuntu:/boot/kgdb-image# gdb ./vmlinux
GNU gdb (Ubuntu 7.11.1-0ubuntu1~16.5) 7.11.1
Copyright (C) 2016 Free Software Foundation, Inc.
License GPLv3+: GNU GPL version 3 or later <http://gnu.org/licenses/gpl.html>
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, 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	man/ cs/	0x00	unsigned int fd	char *buf	size_t count	-	-	-
1	write	man/ cs/	0x01	unsigned int fd	const char *buf	size_t count	-	-	-
2	open	man/ cs/	0x02	const char *filename	int flags	umode_t mode	-	-	-
3	close	man/ cs/	0x03	unsigned int fd	-	-	-	-	-
4	stat	man/ cs/	0x04	const char *filename	struct __old_kernel_stat *statbuf	-	-	-	-
5	fstat	man/ cs/	0x05	unsigned int fd	struct __old_kernel_stat *statbuf	-	-	-	-
6	lstat	man/ cs/	0x06	const char *filename	struct __old_kernel_stat *statbuf	-	-	-	-
7	poll	man/ cs/	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 ma using only 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 be 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 th which has links directly to the source that is installed on your system.) Links to definitions are provided for the parameters that are typedefs or structs.

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

Linux/i386 system calls (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

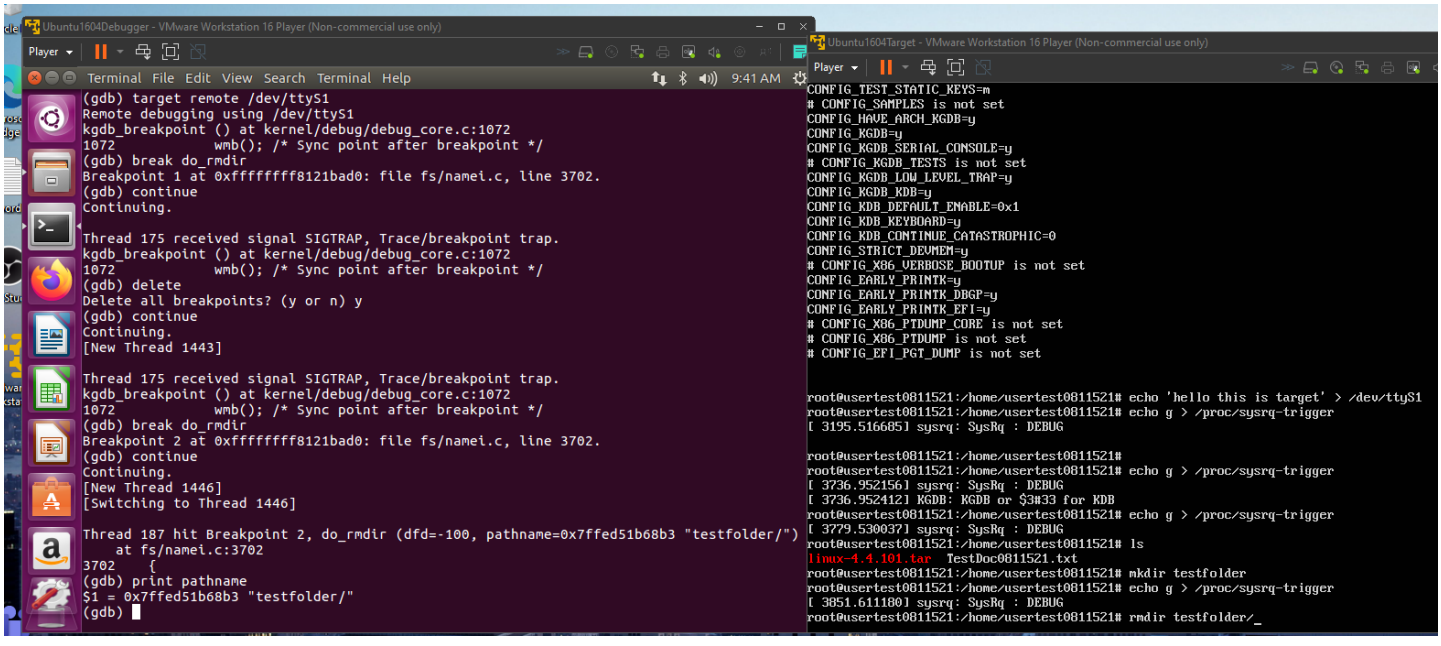
Action: read from a file descriptor

Details:

Screenshot #14

set a breakpoint for **rmdir**

trigger the breakpoint and print the parameter *pathname*



Screenshot #15

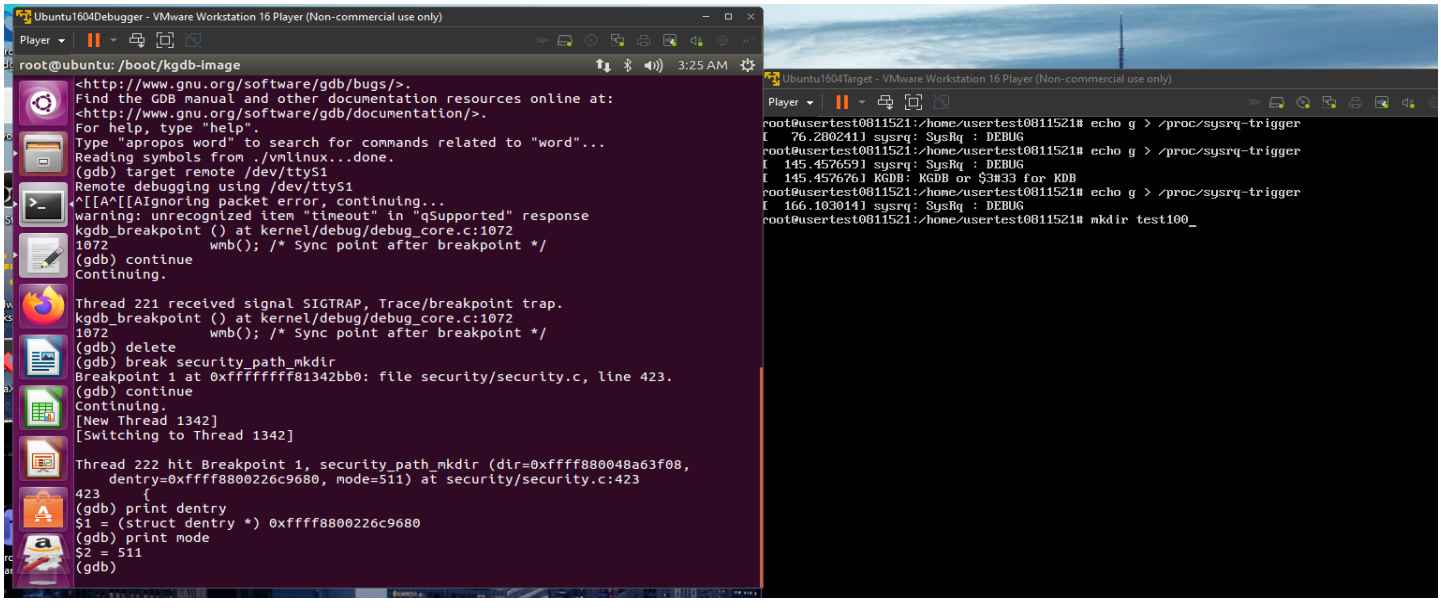
set one breakpoint for **security_path_mkdir** that will be triggered when executing **mkdir**

```
(gdb) break security_path_mkdir
Breakpoint 6 at 0xffffffff81342bb0: file security/security.c, line 423.
```

Screenshot #16

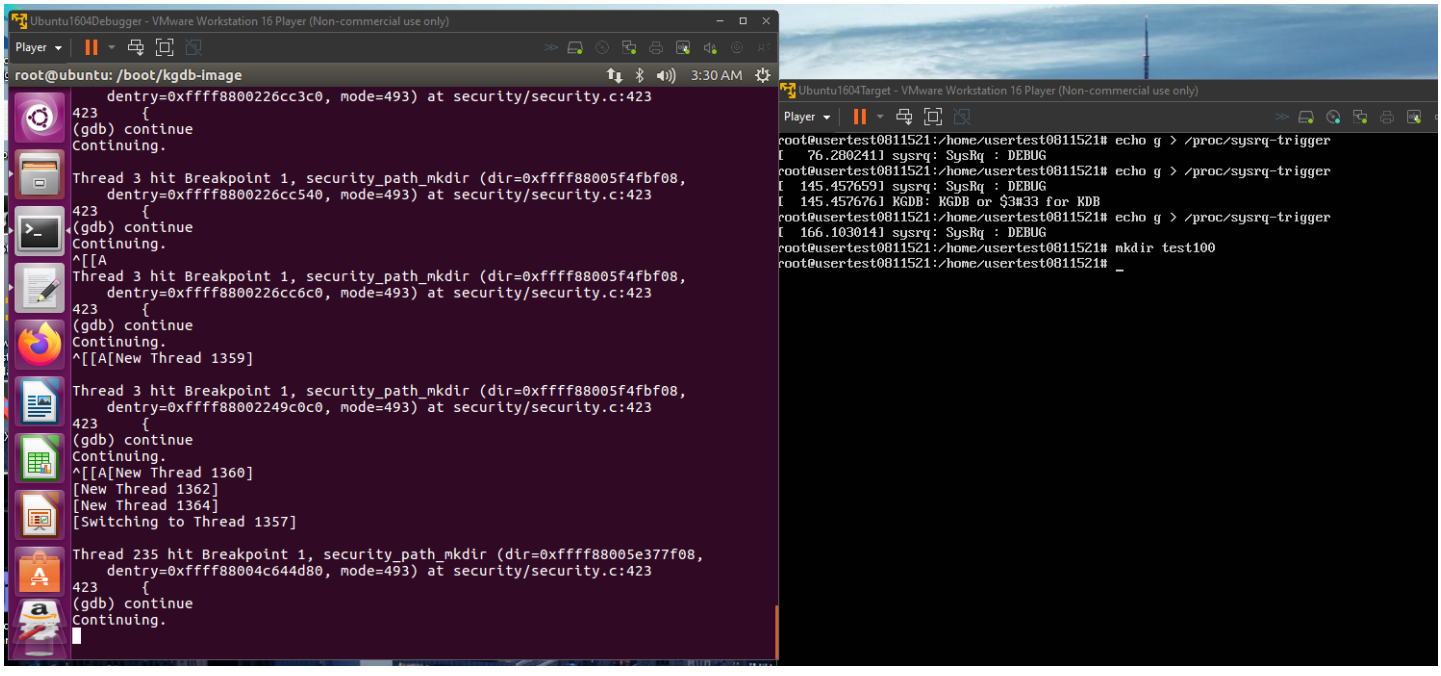
parameters of **security_path_mkdir:** path, dentry, and mode

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



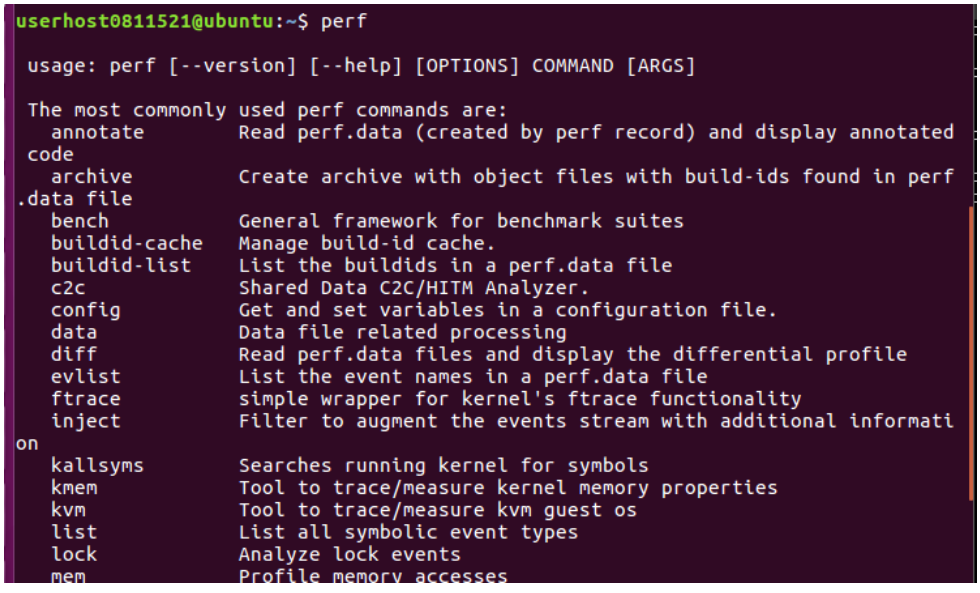
Screenshot #17

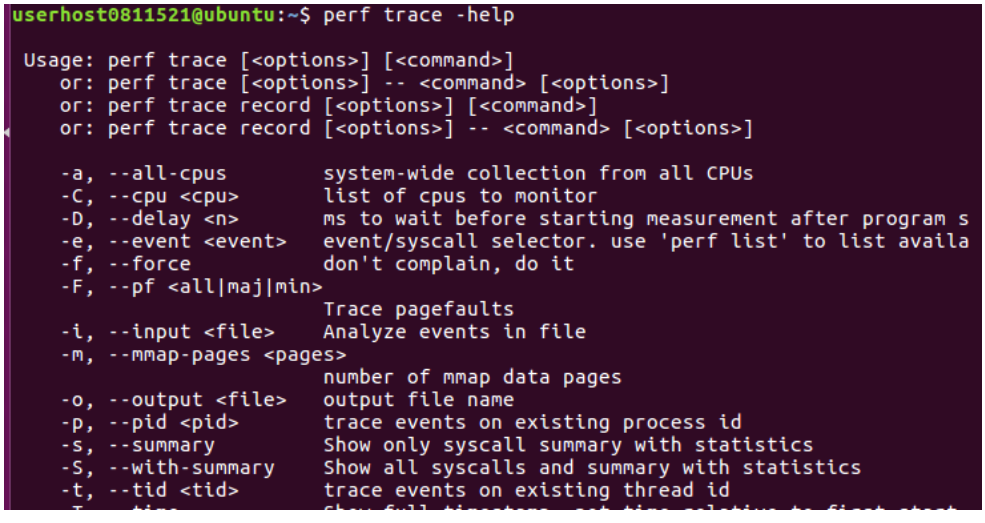
after several **continue**, the target machine regains control.



Screenshot #18

the perf command

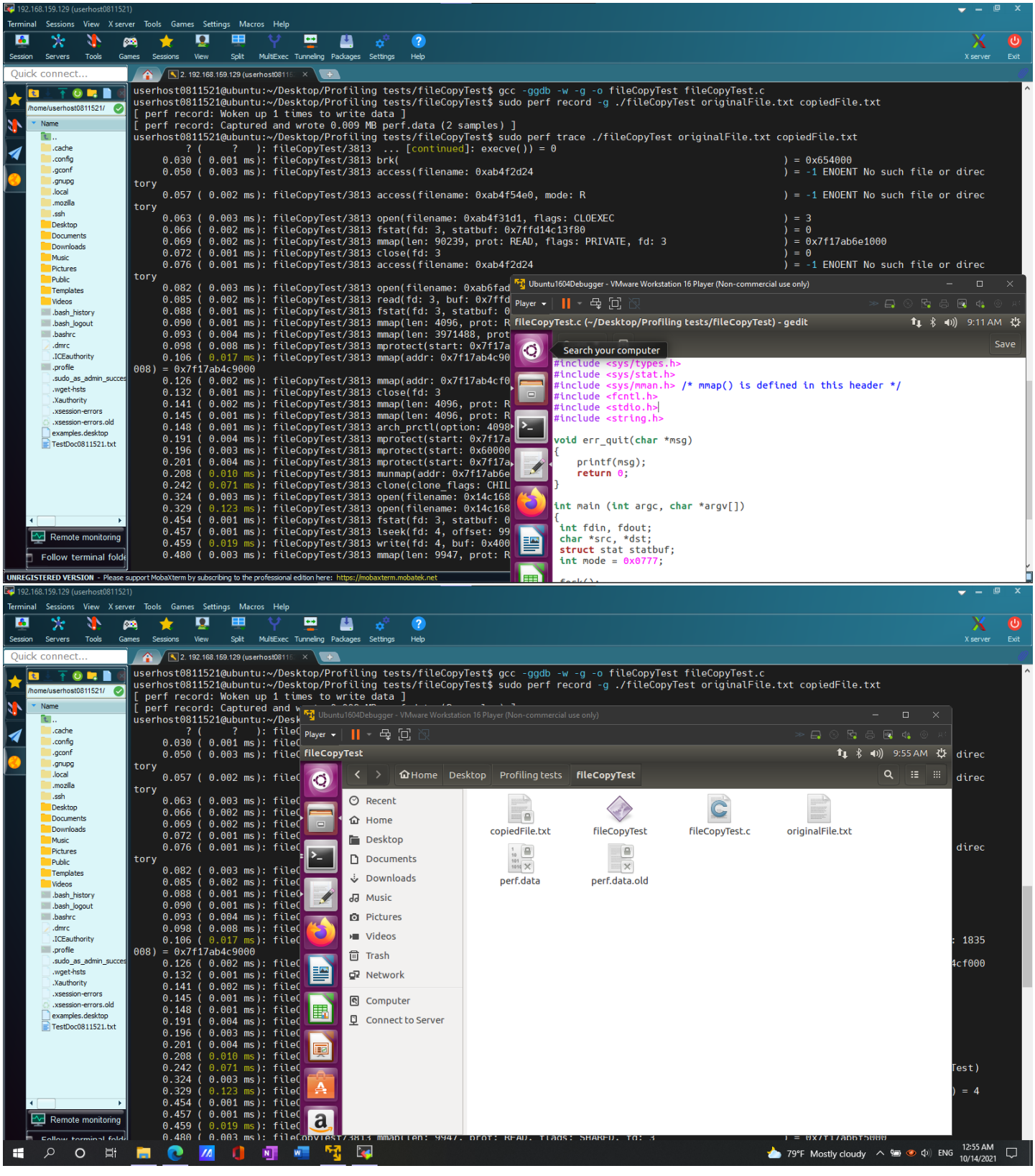




Screenshot #19

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

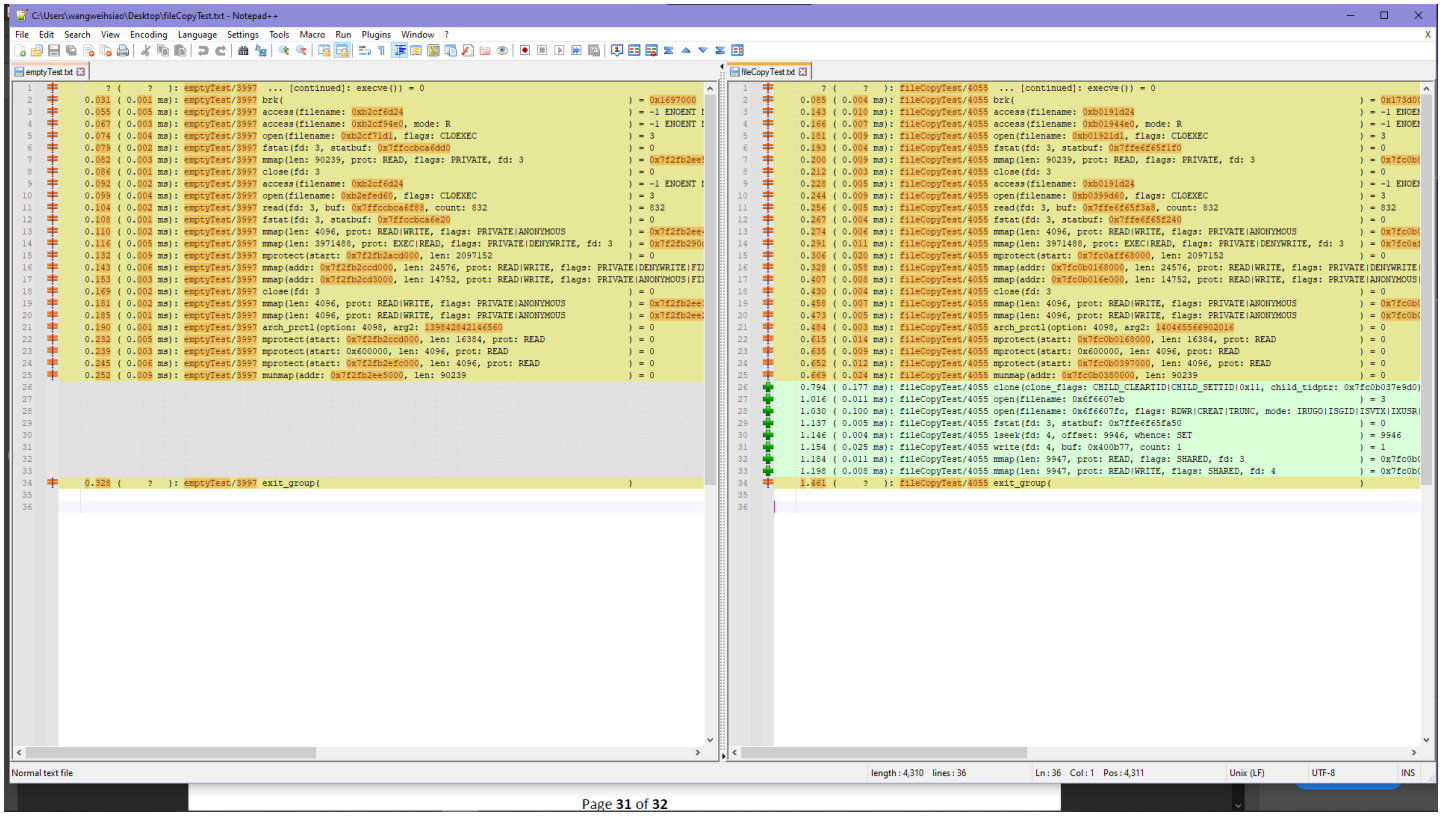
the execution result of perf, the trace result, and the generated files



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.



(p.32)

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

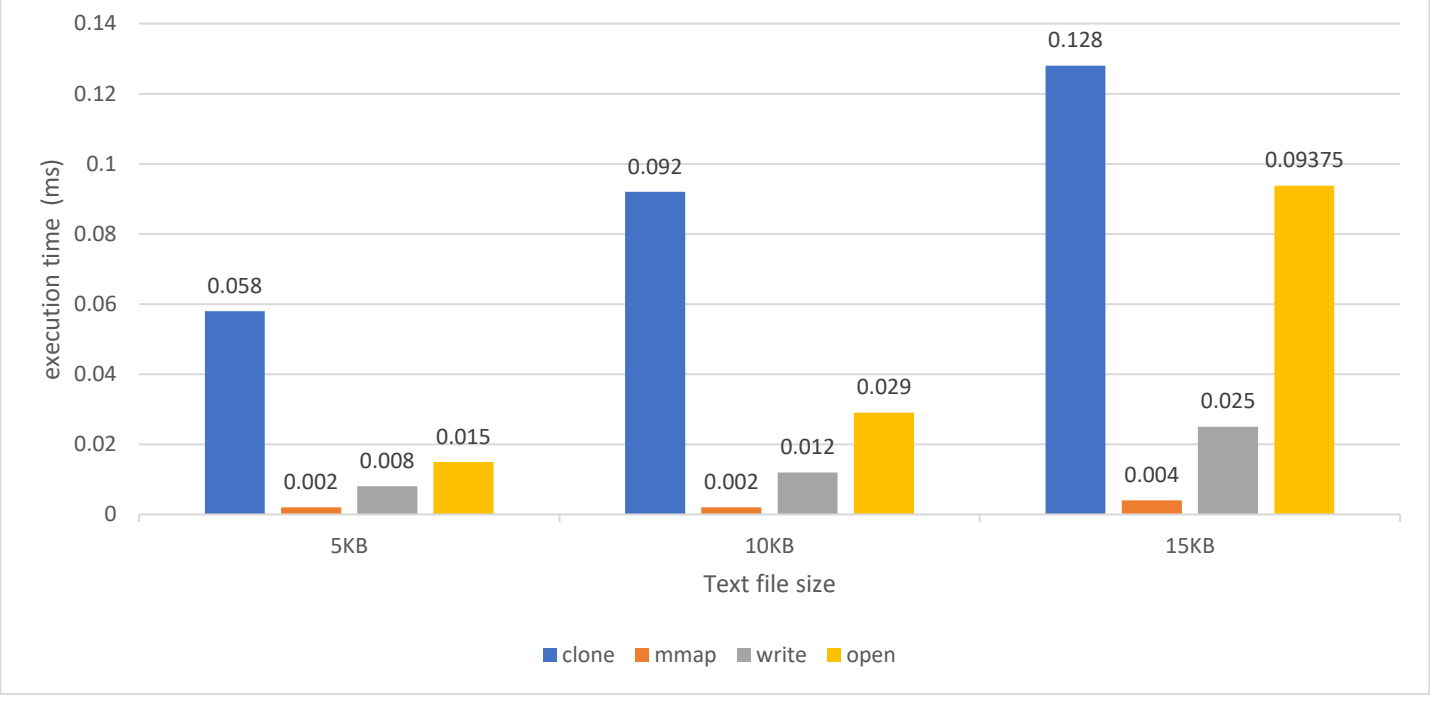
Ans: According to my experiment, the executions of **clone**, **open**, and **write** take longer time when the text file is larger. The executions of **mmap** for different file size take roughly the same amount of time. But there is some time when the executions of same function take same amount of 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.

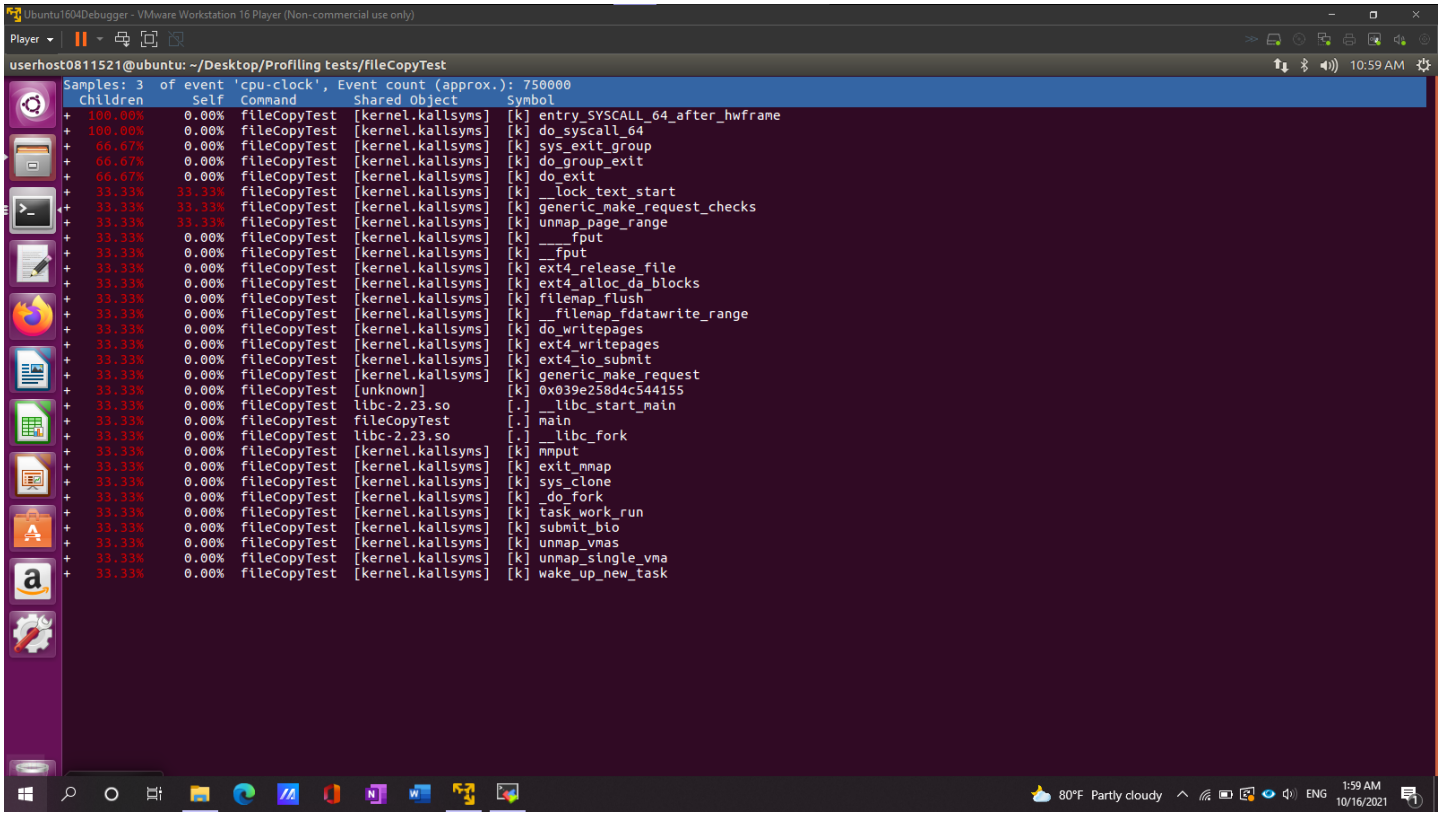
Ans: The following functions have the highest overhead:

__lock_text_start

generic_make_request_checks

unmap_page_range

Screenshot #22 [perf report]

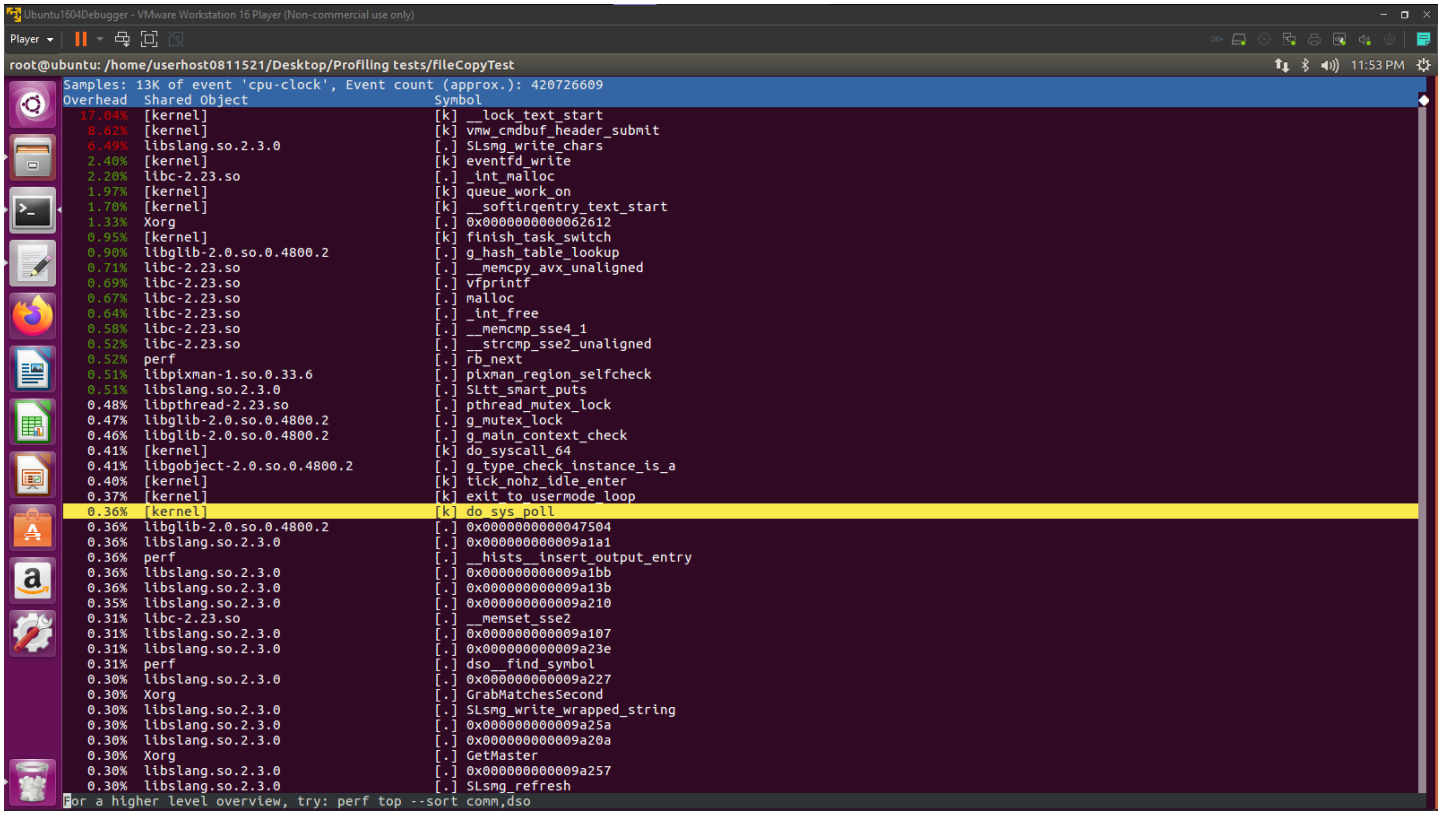


e. Perf has more commands (Section 5.1 step 4). Select another command (besides report, trace and record), explain what it is for and show how to use it. Create your own scenario.

Ans: **perf top**

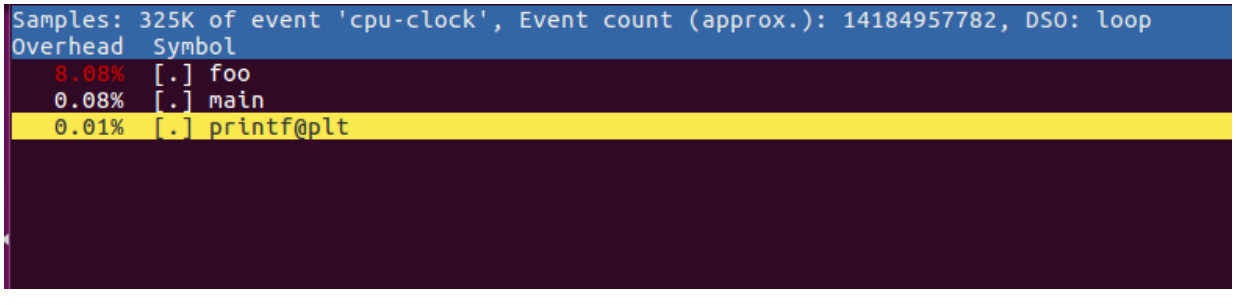
function: real-time analysis of performance

Screenshot #23 [perf top]

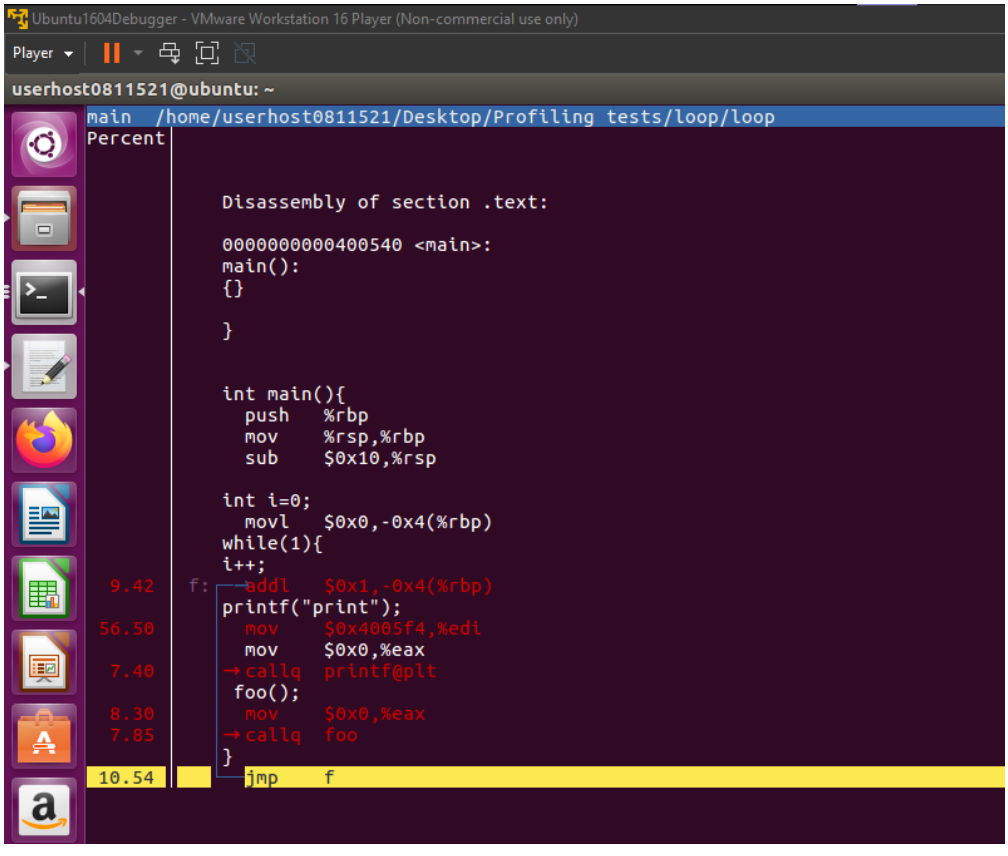


Screenshot #24 [zoom into DSO loop]

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



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



Sources:

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