# Lecture 17 Processes

CS213 – Intro to Computer Systems Branden Ghena – Spring 2021

Slides adapted from:

St-Amour, Hardavellas, Bustamente (Northwestern), Bryant, O'Hallaron (CMU), Garcia, Weaver (UC Berkeley)

#### Administrivia

- Exam on Thursday during class time
  - Same process as last time
  - Covers material through last week Thursday (x86-64 Assembly Procedures to Virtual Memory)

• SETI Lab due Tuesday next week

# Today's Goals

- Explore various mechanisms by which OS and processes interact
  - System calls and signals

Discuss operations on files from processes

#### **Outline**

Process Control Flow

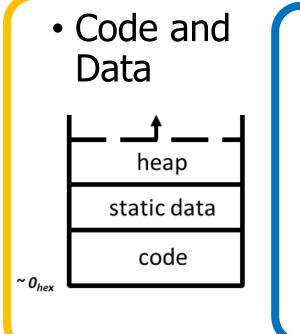
System Calls

- File I/O
  - Standard I/O

Signals

## Reminder: view of a process

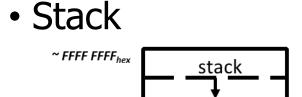
- Process: program that is being executed
- Contains code, data, and a thread
  - Thread contains registers, instruction pointer, and stack



Registers

%rax	%eax	%r8	%r8d
%rbx	%ebx	%r9	%r9d
%rcx	%ecx	%r10	%r10d
%rdx	%edx	%r11	%r11d
%rsi	%esi	%r12	%r12d
%rdi	%edi	%r13	%r13d
%rsp	%esp	%r14	%r14d
%rbp	%ebp	%r15	%r15d

- Instruction Pointer
- Condition Codes



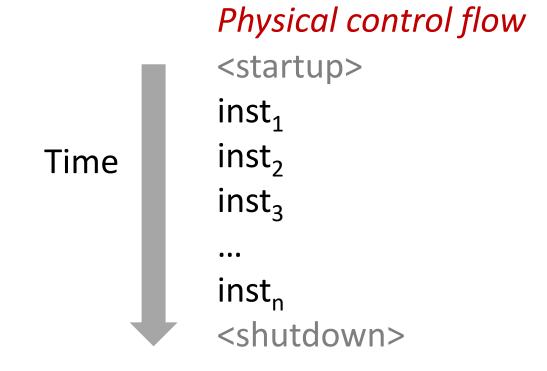
## Questions remaining about processes

- Interaction mechanisms with OS
  - How do processes make requests of the OS?
  - How does the OS inform processes of various events?

 Both answered by the same basic mechanism: exceptional control flow

#### Control flow

- Processors do only one thing:
  - From startup to shutdown, a CPU simply reads and executes (interprets) a sequence of instructions, one at a time
  - This sequence is the CPU's control flow (or flow of control)

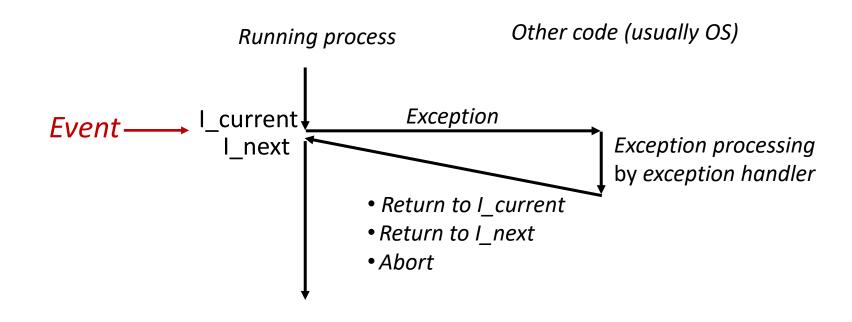


## Altering control flow

- Instructions that change control flow allow software to react changes in program state
  - Jumps/branches
  - Call/return
- Also need to react to changes in system state
  - Data arrives at network adapter
  - Instruction divides by zero
  - User hits Ctrl-C on the keyboard
  - System timer expires
- These mechanisms are known as "exceptional control flow"

#### Exceptional control flow

- Mechanisms
  - Exceptions: events cause execution to jump to OS handler
  - Context switch: request or timeout causes execution to jump to OS
  - Signals: event plus OS causes execution to jump to process handler



#### Exceptions

- Hardware detects an event that OS software needs to resolve
- Could be an error
  - Invalid memory access
  - Invalid instruction
- Could just be something the OS should handle
  - Page fault
  - USB device detected
- OS has a table of "exception handlers", which are functions that handle each exception class (also known as interrupt handlers)
  - Hardware jumps execution to the proper handler

#### **Outline**

Process Control Flow

System Calls

- File I/O
  - Standard I/O

Signals

## Things a program cannot do itself

- Print "hello world"
  - because the display is a shared resource.
- Download a web page
  - because the network card is a shared resource.
- Save or read a file
  - because the filesystem is a shared resource, and the OS wants to check file permissions first.
- Launch another program
  - because processes are managed by the OS
- Send data to another program
  - · because each program runs in isolation, one at a time

# How does a process ask the OS to do something?

- Certain things can only be accessed from kernel mode
  - All of memory, I/O devices, etc.
  - Kernel: the portion of the OS that is running and in memory
- Bad Idea to allow processes to switch into kernel mode
  - We do NOT trust processes
- Requirements
  - Switch execution to the OS kernel
  - 2. Change into kernel mode
  - 3. Inform the OS kernel what you want it to do

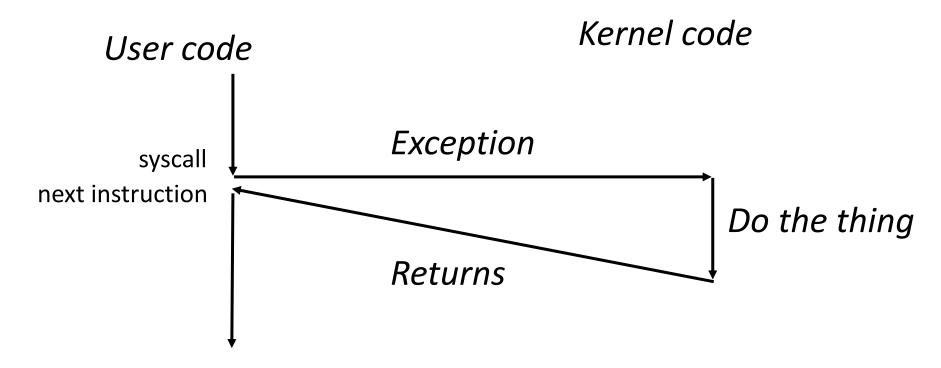
#### Hardware can save us!

- Solution: trigger an exception to run an OS handler
  - Hardware instruction: trap
- When instruction runs:
  - 1. PC is moved to a known location in the kernel
  - 2. Mode is changed to kernel mode

- Same mechanism is used for other exceptions
  - Division by zero, invalid memory access
  - Also very similar to hardware interrupts

# System call example

- System call: making a request of the OS from a process
  - Uses exceptional control flow to enter OS kernel
  - Returns back to process when complete
    - Instruction after the system call



# System call steps (simplification)

- 1. Process loads parameters into registers (just like a function call)
- 2. Process executes trap instruction (int, syscall, svc, etc.)
- 3. Hardware moves %rip to "handler" and switches to kernel mode
- 4. OS checks what the process wants to do from registers
- 5. OS decides whether the process is allowed to do so

## Returning from a system call (simplification)

- After OS finishes whatever operation it was asked to do
  - And when the process is scheduled to run again

- 1. OS places return result in a register (just like a function call)
- 2. OS sets process state to running
- 3. OS changes mode to user mode (and sets virtual memory stuff)
- 4. OS sets %rip to instruction after the system call

Process continues and can use results of system call

# Linux system calls

- Example system calls
  - <a href="https://man7.org/linux/man-pages/man2/syscalls.2.html">https://man7.org/linux/man-pages/man2/syscalls.2.html</a>

Number	Name	Description	
0	read	Read file	
1	write	Write file	
2	open	Open file	
3	close	Close file	
4	stat	Get info about file	
57	fork	Create process	
59	execve	Execute a program	
60	_exit	Terminate process	
62	kill	Send signal to process	

#### Example using system calls

Let's create new processes with system calls

- From process view:
  - Just look like regular C functions
  - Take arguments, return values
- Underneath:
  - Function uses special assembly instruction to trigger exception

#### Process management system calls

```
pid_t fork(void);
```

- Create a new process that is a copy of the current one
- Returns either PID of child process (parent) or 0 (child)

```
void _exit(int status);
```

• Exit the current process (exit(), the library call cleans things up first)

```
pid_t waitpid(pid_t pid, int *status, int options);
```

Suspends the current process until a child (pid) terminates

```
int execve(const char *filename, char *const argv[], char *const envp[]);
```

- Execute a new program, replacing the existing one
- Replaces code and data, clears registers, sets %rip to start again

#### Creating a new process

```
#include <stdio.h>
#include <unistd.h>
int main(){
  if(fork() == 0) {
    printf("Child!\n");
  } else {
    printf("Parent!\n");
  printf("Both!\n");
  return 0;
```

#### Creating a new process

```
#include <stdio.h>
#include <unistd.h>
int main(){
 if(fork() == 0) {
   } else {
   printf("Parent!\n");
 printf("Both!\n");
 return 0;
```

#### Executing a new program

```
#include <stdio.h>
#include <unistd.h>
int main(){
  if(fork() == 0) {
    execve("/bin/python3", ...);
  } else {
    printf("Parent!\n");
  printf("Only parent!\n");
  return 0;
```

#### Break + Question

 What does the following code do?

```
#include <stdio.h>
#include <sys/types.h>
int main() {
 while(1){
    fork();
  return 0;
```

#### Break + Question

 What does the following code do?

```
#include <stdio.h>
#include <sys/types.h>
int main() {
  while (1) {
    fork();
  return 0;
```

- Creates a new process
  - Then each process creates a new process
  - Then each of those creates a new process...
- Known as a Fork bomb!
  - Machine eventually runs out of memory and processing power and will stop working
- Defense: limit number of processes per user

#### **Outline**

Process Control Flow

System Calls

- File I/O
  - Standard I/O

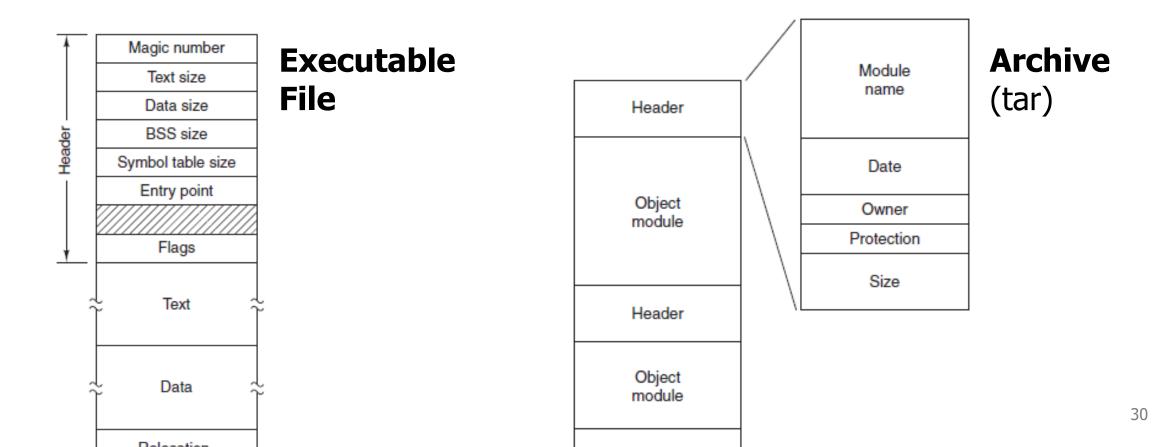
Signals

#### **Files**

- Collections of data
  - Usually in permanent storage on your computer
- Types of files
  - Regular files
    - Arbitrary data
    - Think of as a big array of bytes
  - Directories
    - Collections of regular files
  - Special files
    - Links, pipes, devices (see CS343)

# Sidebar: what about types of regular files?

- Text files versus Executables versus Tar files
  - All just differing patterns of bytes!
  - It really is just all data. The meaning is in how you interpret it.



#### Identifying regular files

- file in Linux command line can help determine the type of a file
  - https://github.com/file/file

```
arguments arguments.c
[brghena@ubuntu code] $ file arguments.c
arguments.c: C source, ASCII text
[brghena@ubuntu code] $ file arguments
arguments: ELF 64-bit LSB shared object, x86-64, version 1 (SYSV), dynamically linked, interpreter /lib64
/ld-linux-x86-64.so.2, BuildID[sha1]=8731c4961d371f4989cd1b056f796ad54b711e6f, for GNU/Linux 3.2.0, not s
tripped
[brghena@ubuntu code] $ file ./
./: directory
[brghena@ubuntu code] $ file ~/scratch/GlobalProtect_UI_deb-5.1.0.0-101.deb
/home/brghena/scratch/GlobalProtect_UI_deb-5.1.0.0-101.deb: Debian binary package (format 2.0), with cont
rol.tar.gz, data compression xz
```

```
[brghena@ubuntu code] $ ls -la
total 32
drwxrwxr-x 2 brghena brghena 4096 Nov 18 22:07 .
drwxr-xr-x 4 brghena brghena 4096 Nov 18 18:38 ..
-rwxrwxr-x 1 brghena brghena 16704 Nov 18 21:42 arguments
-rw-rw-r-- 1 brghena brghena 235 Nov 18 21:42 arguments.c
```

```
[brghena@ubuntu code] $ ls -la
total 32
drwxrwxr-x 2 brghena brghena 4096 Nov 18 22:07 .
drwxr-xr-x 4 brghena brghena 4096 Nov 18 18:38 ..
-rwxrwxr-x 1 brghena brghena 16704 Nov 18 21:42 arguments
-rw-rw-r-- 1 brghena brghena 235 Nov 18 21:42 arguments.c
```

- Permissions for the owner and name of the owner
  - Read, Write, eXecute
    - Cannot execute `arguments.c`
  - For directories: Read contents, Write new contents, Traverse directory

```
[brghena@ubuntu code] $ ls -la
total 32
drwxrwxr-x 2 brghena brghena 4096 Nov 18 22:07 .
drwxr-xr-x 4 brghena brghena 4096 Nov 18 18:38 ..
-rwxrwxr-x 1 brghena brghena 16704 Nov 18 21:42 arguments .c
```

- Permissions for the group and name of the group
  - Example: I could make a CS213 group, add you all to it, and only give that group access to some folder or file

```
[brghena@ubuntu code] $ ls -la
total 32
drwxrwxr-x 2 brghena brghena 4096 Nov 18 22:07 .
drwxr-xr-x 4 brghena brghena 4096 Nov 18 18:38 ..
-rwxrwxr-x 1 brghena brghena 16704 Nov 18 21:42 arguments
-rw-rw-r-- 1 brghena brghena 235 Nov 18 21:42 arguments.c
```

- Permissions for everyone else on the computer
  - Not the owner and not in the group
  - For my personal machine, not particularly relevant
  - For Moore, probably don't want to let others read your files...

#### How do we interact with files?

- Analogy: think of a file as a book
  - Big array of characters (bytes)

- 1. Open the book, starting at the first page
- 2. Read from the book
- 3. Write to the book
- 4. Change pages (without reading everything in between)
- 5. Close the book when finished

# System calls for interacting with files

- 1. Open the book, starting at the first page
  - open()
- 2. Read from the book
  - read()
- 3. Write to the book
  - write()
- 4. Change pages (without reading everything in between)
  - lseek()
- 5. Close the book when finished
  - close()

#### Opening files

• int open (const char \*pathname, int flags);

- pathname is the string path for the file
  - "/home/brghena/class/cs213/s21/code/arguments.c"
  - "./arguments.c"
  - "arguments.c"
- flags include access permission requests
  - Read only, Write only, Read and Write (O\_RDONLY, O\_WRONLY, O\_RDWR)
  - Also can choose to append to a file (O\_APPEND)
  - Or to create the file if it does not exist (O\_CREAT)

### Open returns a "file descriptor"

• int open (const char \*pathname, int flags);

- OS keeps track of opened files for each process
  - File descriptor is just a number referring to the opened file
  - Non-negative number. Always the lowest unused, starting at zero
    - A "handle" to the file

- File descriptor is used in other calls to reference the file
  - That way the OS doesn't have to look up pathname every time
- Negative number instead specifies an error (for all of these calls)

### Reading files

```
• ssize_t read(int fd, void *buf, size_t count);
```

- fd is the file descriptor handle
- buf is a pointer to an array of bytes to read into
- count is the number of bytes to read

- Note: nowhere do we specify where to start reading
  - OS kernel keeps track of a file offset with the descriptor
  - Updated on each read
    - First read of 100 bytes starts at zero, next starts 100 bytes in

How do we know when we finished the file?

```
• ssize_t read(int fd, void *buf, size_t count);
```

- Return from read is a "signed size", a count of bytes actually read
  - Negative means an error occurred
  - Zero means we have reached the end of the file
  - Positive number is the number of bytes read
    - Probably how many we asked for, but maybe less

### Writing files looks a lot like reading

```
• ssize_t write(int fd, const void *buf, size_t count);
```

- File descriptor, buffer to write from, count of bytes to write
- Returns number of bytes actually written

• Write occurs at the current file offset

### Moving the file offset

```
• off_t lseek(int fd, off_t offset, int whence);
```

- Moves to offset for this file descriptor based on whence:
  - SEEK\_SET set to offset (essentially start of file plus offset)
  - SEEK\_CUR current location plus the offset
  - SEEK\_END end of file plus the offset (which can be negative)
- Returns the resulting offset into the file
  - Units: bytes from the beginning of the file

### Closing a file

```
• int close (int fd);
```

Closes the file descriptor

- It is an error to keep using the file descriptor after it is closed
  - Descriptor might end up getting reused for a different file

## Sidebar: how do you figure out how these calls work?

- Manual pages
- Online: <a href="https://man7.org/linux/man-pages/man2/close.2.html">https://man7.org/linux/man-pages/man2/close.2.html</a>

```
close(2) — Linux manual page
```

int close(int fd):

```
NAME | SYNOPSIS | DESCRIPTION | RETURN VALUE | ERRORS | CONFORMING TO | NOTES | SEE ALSO | COLOPHON

CLOSE(2)

Linux Programmer's Manual

CLOSE(2)

NAME top

close - close a file descriptor

SYNOPSIS top

#include <unistd.h>
```

### Higher-level methods of file interaction

- These were all system calls to the OS
- C standard library also defines file interactions
  - fopen, fread, fwrite, fseek, fclose
  - All are wrappers on top of the actual syscalls
  - Buffers your interactions to make them more efficient
    - Reads/Writes large chunks of data at a time
    - Might collect multiple fwrite's before doing a single real write
    - fflush() guarantees that the buffer is written now

### Example: "kitten" command line tool

```
int main(int argc, char *argv[]) {
    // check argument count
    if (argc != 2) {
        printf("Usage: ./kitten FILE\n");
        return -1;
   // try opening file
    int fd = open(argv[1], 0 RDONLY);
    if (fd < 0) {
        printf("Error opening file!\n");
        return -1;
    // array to hold read data
    uint8 t read size = 10;
    uint8 t read data[read size];
```

```
while(true) {
       // read from file
       ssize_t read_length = read(fd, read_data, read_size);
       if (read length < 0) {</pre>
           printf("Error reading file!\n");
           return -1;
       if (read length == 0) {
           break;
       // print out data
       for (int i=0; i<read_length; i++) {</pre>
           printf("%c", read data[i]);
   return 0;
```

### **Outline**

Process Control Flow

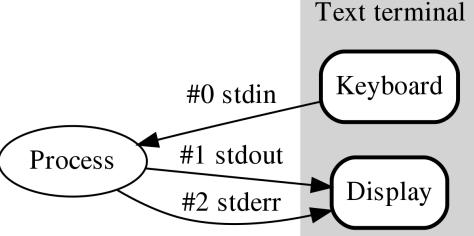
System Calls

- File I/O
  - Standard I/O

Signals

## How do programs talk to users?

- We glossed over this before in CS211
  - printf()
  - gets()



- Work through the same file mechanism
  - Three special files created for each program
  - stdin standard input (file descriptor 0)
  - stdout standard output (file descriptor 1)
  - stderr standard error (file descriptor 2)
- printf(...) -> fprintf(1, ...) -> handle arguments & write(1, ...)

### Standard I/O is a process thing, not a C thing

- You can access them in Python, for instance
  - https://docs.python.org/3/library/sys.html#sys.stdin

```
sys.stdin
sys.stdout
sys.stderr
```

File objects used by the interpreter for standard input, output and errors:

- stdin is used for all interactive input (including calls to input());
- stdout is used for the output of print() and expression statements and for the prompts of input();
- The interpreter's own prompts and its error messages go to stdern.

These streams are regular text files like those returned by the open() function. Their parameters are chosen as follows:

### Example: "kitten" write to standard output

```
int main(int argc, char *argv[]) {
    // check argument count
    if (argc != 2) {
        printf("Usage: ./kitten FILE\n");
        return -1;
    // try opening file
    int fd = open(argv[1], 0 RDONLY);
    if (fd < 0) {
        printf("Error opening file!\n");
        return -1;
    // array to hold read data
    uint8 t read size = 10;
    uint8 t read data[read size];
```

```
while(true) {
       // read from file
       ssize_t read_length = read(fd, read_data, read_size);
       if (read length < 0) {</pre>
           printf("Error reading file!\n");
           return -1;
       if (read length == 0) {
           break;
       // print out data
       ssize_t write_length = write(STDOUT_FILENO,
               read data, read length);
   return 0;
```

### Redirecting standard I/O

- Shells by default setup standard I/O to connect to the keyboard and the screen
  - But any file will work
- Shell I/O redirection commands
  - COMMAND < filename
    - Connect standard input to filename
  - COMMAND > filename
    - Connect standard output to filename (overwrite)
  - COMMAND >> filename
    - Connect standard output to filename (append)

### Piping commands

 A command shell desire is to run multiple commands where the output of the first feeds into the second

- COMMAND1 | COMMAND2
  - Connects stdout of COMMAND1 to stdin of COMMAND2
- Example: print out files and sort by size
  - Is -lah | sort -h

### Sidebar: super useful command for testing

- tee [OPTION] . . . [FILE] . . .
  - Reads from stdin and write to both stdout and file
- Example: prints out a list of files and saves results
  - Is -lah | tee results.txt

• I run this with various programs I'm testing, so I can record the results, but also seem them in real-time.

### Example: redirection with kitten

- Standard I/O redirection is handled when the process is created
  - So it does not need to be aware of it at all
- Our kitten tool works with redirection automatically!
  - ./kitten arguments.c > OUTPUT\_FILE

# Break + Open Question

How does printf() work?

### Break + Open Question

- How does printf() work?
  - 1. Read in arguments and determine what it needs to format
  - 2. Create a new string buffer and write arguments into it
  - 3. Call write() on STDOUT with the string

### **Outline**

Process Control Flow

System Calls

- File I/O
  - Standard I/O

Signals

### Alerting processes of events

- How do we let a process know there was an event?
  - Errors
  - Termination
  - User commands (like CTRL-C or CTRL-\)
- Events could happen whenever
  - Need to interrupt process control flow and run an event handler
  - Linux mechanism to do so is called "signals"

### Signals are asynchronous messages to processes

- Sometimes the OS wants to send something like an interrupt to a process
  - Your child process completed
  - You tried to use an illegal instruction
  - You accessed invalid memory
  - You are terminating now
- In POSIX systems, this idea is called "Signals"

```
1) SIGHUP 2) SIGINT 3) SIGQUIT 4) SIGILL 5) SIGTRAP
6) SIGABRT 7) SIGBUS 8) SIGFPE 9) SIGKILL 10) SIGUSR1
11) SIGSEGV 12) SIGUSR2 13) SIGPIPE 14) SIGALRM 15) SIGTERM
16) SIGSTKFLT 17) SIGCHLD 18) SIGCONT 19) SIGSTOP 20) SIGTSTP
21) SIGTTIN 22) SIGTTOU 23) SIGURG 24) SIGXCPU 25) SIGXFSZ
26) SIGVTALRM 27) SIGPROF 28) SIGWINCH 29) SIGIO 30) SIGPWR
31) SIGSYS ...
```

### Signals are asynchronous messages to processes

- Sometimes the OS wants to send something like an interrupt to a process
  - Your child process completed
  - You tried to use an illegal instruction
  - You accessed invalid memory
  - You are terminating now
- In POSIX systems, this idea is called "Signals"

```
3) SIGOUTT 4) SIGILL
                                                5) SIGTRAP
1) SIGHUP
             2) STGTNT
                                                                     Process Errors
6) STGABRT
             7) SIGBUS
                         8) SIGFPE
                                      9) SIGKILL 10) SIGUSR1
             12) SIGUSR2 13) SIGPIPE
11) SIGSEGV
                                     14) SIGALRM 15) SIGTERM
16) SIGSTKFLT 17) SIGCHLD 18) SIGCONT 19) SIGSTOP 20) SIGTSTP
21) SIGTTIN 22) SIGTTOU 23) SIGURG 24) SIGXCPU 25) SIGXFSZ
26) SIGVTALRM 27) SIGPROF 28) SIGWINCH 29) SIGIO
                                                 30) SIGPWR
31) SIGSYS
```

### Signals are asynchronous messages to processes

- Sometimes the OS wants to send something like an interrupt to a process
  - Your child process completed
  - You tried to use an illegal instruction
  - You accessed invalid memory
  - You are terminating now
- In POSIX systems, this idea is called "Signals"

```
1) SIGHUP 2) SIGINT 3) SIGQUIT 4) SIGIL 5) SIGTRAP

6) SIGABRT 7) SIGBUS 8) SIGFPE 9) SIGKILL 10) SIGUSR1

11) SIGSEGV 12) SIGUSR2 13) SIGPIPE 14) SIGALRM 15) SIGTERM

16) SIGSTKFLT 17) SIGCHLD 18) SIGCONT 19) SIGSTOP 20) SIGTSTP

21) SIGTTIN 22) SIGTTOU 23) SIGURG 24) SIGXCPU 25) SIGXFSZ

26) SIGVTALRM 27) SIGPROF 28) SIGWINCH 29) SIGIO 30) SIGPWR

31) SIGSYS ...
```

**Process Termination** 

### Sending signals

OS sends signals when it needs to

- Processes can ask the OS send signals with a system call
  - int kill(pid\_t pid, int sig);
- Users send signals through OS from command line or keyboard
  - Shell command: kill -9 pid (SIGKILL)
  - CTRL-C (SIGINT)

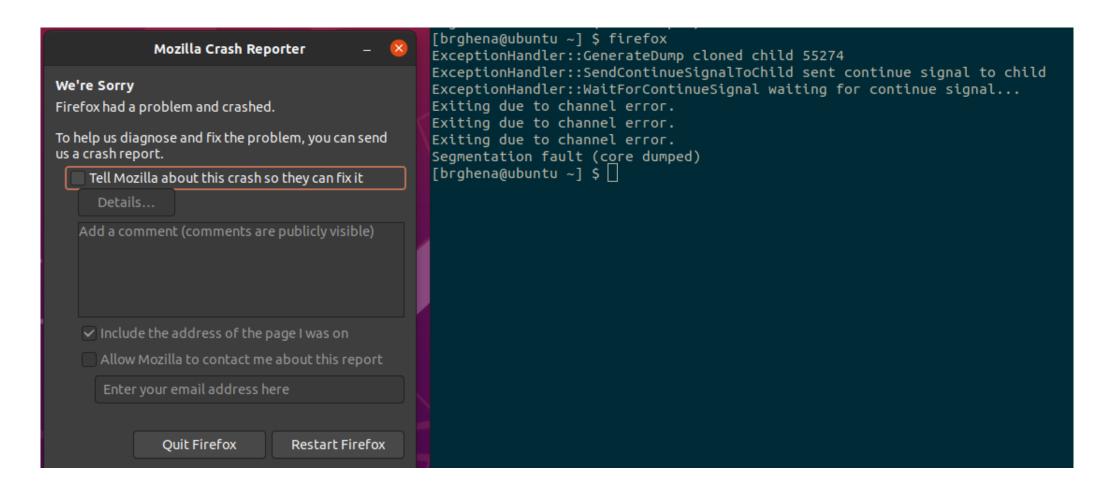
### Handling signals

- Programs can register a function to handle individual signals
  - signal(int sig, sighandler\_t handler);
- What are you supposed to do about it?
  - Do some quick processing to handle it
  - Reset the process and try again
  - Quit the process (default handler)

# Signals Examples

### Examples: sending a signal

> kill -11 *pid* (11 is SIGSEGV – a.k.a segfault)



### Example: catching a signal

```
void sighandler (int signum) {
  printf("HA HA You can't kill me!\n");
int main (void) {
  signal(SIGINT, sighandler);
   printf("Starting\n");
  while(1) {
      printf("Going to sleep for a second...\n");
      sleep(1);
  return(0);
```

```
#include <stdio.h>
#include <unistd.h>
#include <stdlib.h>
#include <signal.h>
```

### Example: catching a segfault

```
#include <stdlib.h>
int* pointer = 0x00000000;
                                                             #include <signal.h>
void sighandler (int signum) {
   printf("Oops, that pointer wasn't valid. Let's try a different one\n");
   pointer++;
   printf("About to read from pointer 0x%081X\n", (long)pointer);
int main (void) {
   signal(SIGSEGV, sighandler);
   printf("About to read from pointer 0x%081X\n", (long)pointer);
  int test = *pointer;
  return(0);
```

#include <stdio.h>

#include <unistd.h>

### **Outline**

Process Control Flow

System Calls

- File I/O
  - Standard I/O

Signals