

Due: See website for due date.

What to submit: A tar file that should contain the file `answers.txt`, which must be an ASCII file with answers for questions 1 and 2. For question 3, which asks for code, include a file named `dpipe.c` in your tar file.

Use the `submit.py` script in `~cs3214/bin/submit.py` to submit your tar file from the command line, or use the submit website. Use the identifier 'ex1.'

The assignment will be graded on the `rlogin` cluster, so your answers must match this environment. In addition, every student will use a variant of `dpipe` that is unique; please visit `cs3214/fall2020/#!/ex1` to find your variant.

Understanding Processes and Pipes

In this exercise we consider a program 'dpipe,' which, in connection with a netcat utility program, can be used to turn any ordinary program into a network server. The goal of this exercise is to understand how a process's standard input and standard output file streams can be manipulated by a control program using standard Unix system call facilities. A very similar file descriptor manipulation is performed by a shell when it sets up I/O redirection or pipes.

In this exercise, you are asked to observe, reverse-engineer, and then reimplement the `dpipe` program. The `dpipe` binary is provided in `~cs3214/bin/dpipe.variants/dpipe-ABC` on our machines, where `ABC` is the unique id created for you. For netcat, use the `~cs3214/bin/gnetcat` binary.

To allow you to invoke those commands directly, make sure that `~cs3214/bin` and `~cs3214/bin/dpipe.variants` are included in your `PATH` environment variable.¹

Here's an example session of how to use it. Say you're logged in to the machine 'locust' and run the command `dpipe-ABC wc gnetcat -l 15999`². This will produce output as follows:

```
[cs3214@locust sys1]$ ~cs3214/bin/dpipe.variants/dpipe-ABC wc gnetcat -l 15999
Starting: gnetcat
Starting: wc
```

Note that the shell is still waiting for `dpipe` to complete. The questions listed in part 1 must be answered at this point, after you have started `dpipe`, but before the next step. You may wish to continue reading and then return to this point, after opening a second terminal on the same machine.³

¹This is, however, not necessary - you can also replace `dpipe-ABC` with its full path `~cs3214/bin/dpipe.variants/dpipe-ABC`

²When you try this out, please replace `15999` with a number that with high likelihood is unique so you do not conflict with other students using this machine, let's agree to using `10000 + last 4 digits of your VT PID` here.

³Hint: `ssh` to `portal.cs.vt.edu`, then choose the machine on which your other shell session runs.

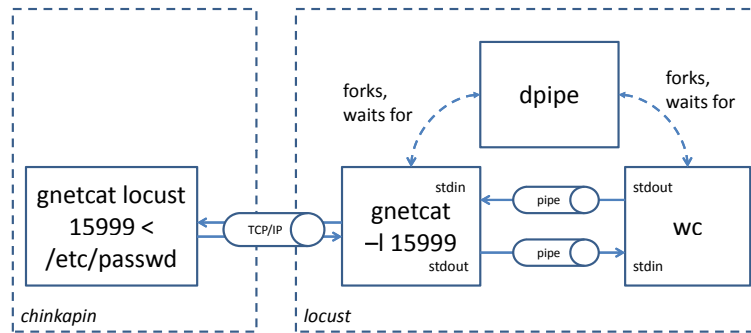


Figure 1: Using dpipe to cross-link two programs' standard input and output streams.

The next command can be run on *any* rlogin machine:

```
[cs3214@chinkapin ~]$ gnetcat locust 15999 < /etc/passwd
38      97      2170
[cs3214@chinkapin ~]$
```

The gnetcat instance started here connected to the gnetcat instance running on locust, and sent its input there (in this case, the content of file `/etc/passwd`). Gnetcat (running on locust) outputs this data to its standard output stream, which is connected via a pipe to the standard input stream of the 'wc' program. The standard output stream of 'wc' is connected to the standard input stream of locust's gnetcat instance. Anything output by 'wc' will then be forwarded to chinkapin's gnetcat instance, which outputs it to its standard output, thus appearing on the user's terminal on chinkapin. You have just turned the ordinary 'wc' program into a network service! The entire scenario is shown in Figure 1.

Finally, dpipe exits when both of the processes it started have terminated:

```
[cs3214@locust sys1]$
```

1. **Using /proc** The `/proc` file system, originally introduced as part of the Solaris OS, is a mechanism by which the Linux kernel provides information about processes to users. This information is organized in directories and files that can be accessed using ordinary tools such as `ls` or `cat`. However, they do not represent actual directories or files – rather, their content is created dynamically when a user accesses them, based on the current state of the system and its processes. The kernel creates a directory for each process, which is named for its pid. Within the directory, there are additional files and directories that describe the current state of the process.

After you have started dpipe with 'wc' and 'gnetcat -l', use the `ps (1)` command to find the process ids of these three processes, then use `ps` or `/proc` to find the parent process id, the process group id, and the current status of these processes.

(a) Provide the following information in table form:

Process	PID	Parent PID	Process Group ID	State
dpipe-ABC				
wc				
gnetcat				

You may find the `proc(5)` man page useful.⁴

Use the single letter Linux shorthand for the process state.

(b) You should find that all three processes are in the same state. To which of the 3 states of the simplified process state diagram discussed in lecture does this state correspond?

(c) Now examine the `/proc/nnnnn/fd` directories for each process. Use `ls -l`. For each process's open file descriptors, enter in the table below the device or object to which it refers. Enter N/O if a file descriptor is not currently in use.

File Descriptor	dpipe-ABC	wc	gnetcat
(stdin) 0			
(stdout) 1			
(stderr) 2			
3			

Look at Figure 1 and convince yourself that dpipe indeed creates the arrangement shown there (minus the network connection, which has not yet been created).

2. **Using strace.** Processes use system calls to obtain services from the kernel. The program 'strace' is a utility that can display which system calls a process is executing. strace can trace individual processes from start to end; on request, it can trace child and ancestor processes as well, and it can even trace already running processes. Familiarize yourselves with the 'strace' command (use 'man strace', for instance). Frequently used switches include `-f`, `-e`, `-p`, `-ff`, and `-o` which you should memorize.

Run dpipe-ABC under strace, creating again the constellation of processes described in the introduction (wc and gnetcat), and use gnetcat on another machine to send data until dpipe exits. Use the `-ff` and `-o` switches to strace to obtain separate listings for dpipe and its children.

(a) Identify which system calls the dpipe process made. From the relevant strace output, copy the entries corresponding to the system calls `pipe()`, `close()`, `fork()`, `wait()`, and `exit()`. Hint: Linux's implementation of the `fork()` system call is actually called `clone()`, `wait()` is implemented using `wait4()`, and `exit()` corresponds to `exit_group()`.

(b) Identify which system calls the process executing 'wc' made. Include the entries

⁴The number (5) in parentheses denotes the section of the man pages in which the information is found. On the command line, you would type `man 5 proc` to access this page in section 5. `man proc` works too, but only because only section 5 of the manual has an entry for `proc`.

corresponding to `write()`, `dup2()`, `close()`, and `execve()`. Stop at the first `execve()` call that was successful (listed as returning 0 in `strace`'s output.)

(c) Like part b), except for the process executing 'gnetcat'.

Note that the traces for part (b) and (c) correspond to the system calls made by the child processes spawned by `dpipe-ABC` between `fork()` and `exec()`.

(d) Now, rerun `strace` with the switches `-f -e wait4,accept,read` without connecting the gnetcat client (like in part 1). All three processes will be inside the kernel, in one of these three system calls. Answer the following questions for each process:

- What is the state for each process, using the nomenclature of the simplified process state diagram discussed in lecture?
- Which process is in which system call?
- How can each process's state be explained from the semantics of that system call?

Use the following example format for your answer: Process 'dpipe' is in the (insert state here) state in the (insert system call here) system call because it is (describe what caused it to be in this state).

3. Implement `dpipe-ABC`.

Armed with the knowledge gained in the previous part, implement `dpipe-ABC` so that it issues **exactly** the system calls you've observed in the traces, in the same order.

Although the OS will use predictable file descriptors in `pipe()`, your program should use variables and not constants, e.g. calling `close(3)` would not be an acceptable solution.

Use the `execvp()` frontend for the `execve()` system call. Make sure you implement argument passing correctly. `dpipe` should first invoke the program whose name is given as its first argument, without passing any additional arguments to it. Then it should interpret the second argument as the name of a second program and pass the third argument as the first argument to the second program, the fourth argument as the second, and so on. It is customary to set `argv[0]` to the name of the program itself, and your invocations of `execvp()` should do the same. Hint: it is possible to achieve this without creating a complete copy of the `argv[]` array passed to `dpipe`'s main program.

If done correctly, the `strace` output for the relevant system calls⁵ in your `dpipe-ABC` implementation must be identical to the output obtained from the provided `dpipe-ABC`; in fact, this is what the TAs' software will check.

⁵these include the pipe, clone, and wait-related calls, as well as the file descriptor-related calls for the file descriptors you create, manipulate, and close. You may find difference before and after those.

Important: even though the relative order of some of the system calls is not important for the correct functioning of `dpipe`, your assigned variant of `dpipe` will have a specific ordering your submission must reproduce to obtain full credit.

Make sure that your `dpipe-ABC` does not return to the prompt prematurely, or fails to return to the prompt after its children have exited.

Important: We will use the compiler switches `-O -Wall -Werror` when compiling your program. If you compile without these switches, and any warnings occur in your code, you run the risk of failing our autograding scripts.