CSE 344 System Programming 4th Assignment Report



Author

Mert Emir ŞEKER 200104004085

Date

26.05.2024

Table Of Contents

1.	Makefile	3
2.	Makefile Commands	3
3.	How to run the code?	3
4.	Code Explanation	4
	1 Buffer.c and Buffer.h	
4.2	2 Manager.c and Manager.h	6
4.3	3 Worker.c and Worker.h	8
4.4	4 Main.c	10
5.	Example Outputs	13
	1 Test 1 (in pdf)	
5.2	2 Test 2 (in pdf)	13
	3 Test 3 (in pdf)	
5.4	4 ./MWCp 10 10/testdir/tocopy	14
	valgrind ./MWCp 1024 10/testdir/tocopy	
5.6	6 ./MWCp 1024 10/testdir/tocopy	15
	7 valgrind ./MWCp 10 4/testdir/src/libvterm/src/newfile	
	3 valgrind ./MWCp 10 10/testdir/toCopy	
5.9	yalgrind ./MWCp 10 10/testdir/tocopy (with Ctrl-C signal)	17

1. Makefile:

```
M makefile
M makefile
  1
       all: compile
  2
  3
       compile:
           gcc -o MWCp 200104004085_main.c buffer.c manager.c worker.c -lpthread -lrt
  4
  5
  6
  7
      clean:
  8
           rm -f *.o
           rm -rf ../tocopy/*
  9
 10
           rm MWCp
           clear
 11
 12
```

2. Makefile Commands:

All: Compiles the code.

Compile: Compiles the code.

Clean: Clears terminal.

3. How to run the code?

To run the program, you can write **make all** or just **make** in terminal and then **./MWCp #buffer_size #ofworkers ../source_file_directory ../destination_file_directory**.

4. Code Explanation and pseudocodes:

4.1. Buffer.c and Buffer.h: A thread-safe buffer mechanism is implemented in this C code (buffer.c) and its accompanying header (buffer.h) in order to synchronize file descriptor operations between a management thread and many worker threads in a directory copying tool. The buffer, which is described in the buffer_str structure, leverages POSIX threading primitives for correct synchronization and stores file descriptors representing files that need to be duplicated. The buffer_init function allocates memory for file descriptors, initializes mutex and condition variables, specifies the size of the buffer, and initializes a number of indices and the done flag. By deleting synchronization primitives, freeing up memory, and shutting file descriptors, the buffer_destroy method tidys up. The management thread uses the buffer_add method to add file descriptors to the buffer, guaranteeing mutual exclusion and alerting worker threads when data becomes accessible. On the other hand, worker threads utilize buffer_remove to obtain file descriptors, guaranteeing mutual exclusion and alerting the manager when space becomes available. Buffer_set_done broadcasts to wake up any waiting threads and sets the done flag. The buffer.h header file includes function prototypes for buffer operations and signal handling, external variables for thread management and statistics tracking, and the essential data structures (buffer_str and file_descriptor). This method guarantees effective thread operation without idle waiting and appropriate resource cleaning afterward.

Pseudocode:

// Initialize the buffer

function buffer_init(buffer, size) set buffer size set buffer count to 0 set buffer done flag to 0 set buffer input index to 0 set buffer output index to 0 allocate memory for file descriptors if memory allocation fails print error message and exit initialize mutex initialize not full condition variable initialize not empty condition variable // Destroy the buffer and free resources function buffer_destroy(buffer) lock mutex for each file descriptor in the buffer close source file descriptor close destination file descriptor unlock mutex free file descriptor memory destroy mutex

destroy not full condition variable
destroy not empty condition variable
// Add a file descriptor to the buffer
function buffer_add(buffer, fd)
lock mutex
while buffer is full and not done
wait for not full condition
if done
unlock mutex and return
add file descriptor to buffer at input index
update input index
increment buffer count
signal not empty condition
unlock mutex

// Remove a file descriptor from the buffer function buffer_remove(buffer)
lock mutex
while buffer is empty and not done
wait for not empty condition
if buffer is empty and done
unlock mutex and return invalid file descriptor get file descriptor from buffer at output index
update output index
decrement buffer count
signal not full condition
unlock mutex
return file descriptor

// Set the done flag of the buffer function buffer_set_done(buffer) lock mutex set done flag broadcast not empty condition broadcast not full condition unlock mutex

4.2. Manager.c and Manager.h: The directory copying utility's manager thread functionality is implemented in this C file (manager.c) and the matching header (manager.h). In order for worker threads to handle file descriptors, the management thread must traverse the source directory, confirm that destination directories exist, and add file descriptors to the buffer. The create directory if not exists function determines if a directory already exists and, if not, creates it. By extracting the directory path and using create_directory_if_not_exists, the function ensure_parent_directories_exists verifies that all parent directories for a given file path exist. Copy_file is a function that copies files from one location to another, checks if parent directories are present, opens source and destination files, creates FIFOs if necessary, adds file descriptors to the buffer, and updates counters for normal and FIFO files. In order to process a directory, the function process_directory calls copy_file for files and process directory recursively for subdirectories. It also verifies that the destination directory exists, opens the source directory, reads its entries, and processes each entry according to its type (directory, FIFO, or regular file). Initializing the required variables, determining if the source path is a directory, FIFO, or ordinary file, and processing it appropriately are all done by the manager function. It also constructs the root destination directory for a directory, runs copy_file for files, and sets the buffer's done flag when it's finished.The manager.h header file defines the function prototypes for buffer operations and signal handling, external variables for managing threads and tracking statistics, and the manager_args structure for passing arguments to the manager function. These definitions guarantee effective thread operation and appropriate resource cleanup after completion.

Pseudocode:

if file is FIFO

create FIFO file if creation fails print error

```
function create_directory_if_not_exists(dir_path)
 if directory does not exist
    create directory
    if creation fails
     print error and exit
// Function to ensure the parent directories exist for the given file path
function ensure_parent_directories_exist(file_path)
 duplicate file path
 find the last slash in the path
 if last slash exists
    terminate string at the last slash
    create directory if it does not exist
 free duplicated string
// Function to copy a file from source to destination
function copy_file(src_path, dst_path, buffer, mode)
 ensure parent directories exist
```

// Function to create a destination directory if it does not exist

```
increment FIFO file counter
 else // Regular file
    open source file
    if opening fails
     print error
     return
    open destination file
    if opening fails
     print error
     close source file descriptor
     return
    create file descriptor structure
    add file descriptor to buffer
    increment regular file counter
// Function to process a directory
function process_directory(src_dir, dst_dir, buffer)
 ensure destination directory exists
 open source directory
 if opening fails
    print error
    set buffer as done
    return
 while reading directory entries and not done
    if entry is . or ..
     continue
    construct source path
    construct destination path
   get file status
    if getting status fails
     print error
     continue
    if entry is a directory
     recursively process directory
     increment directory counter
    else if entry is a FIFO
     copy FIFO file
    else if entry is a regular file
     copy regular file
 close source directory
// Manager function to handle copying process
function manager(arg)
 cast argument to manager_args pointer
 get buffer from arguments
 get source and destination paths from arguments
 get status of source path
```

return

```
if getting status fails
  print error
  set buffer as done
  return
if source path is a directory
  create root destination directory
  ensure root destination directory exists
  process source directory
else if source path is a FIFO
  construct destination file path
  copy FIFO file
else if source path is a regular file
  construct destination file path
  copy regular file
else
  print error for unsupported file type
  set buffer as done
  return
set buffer as done
return
```

5.3 Worker.c and Worker.h: The directory copying utility's worker thread functionality is implemented in this C code (worker.c) and the matching header (worker.h). Every worker thread updates the amount bytes transferred, transfers data from the source file to the destination file, and recovers file descriptors from the buffer. The worker thread begins by allocating memory for a buffer to read data from and converting the parameter to a buffer_str pointer. After that, the worker function goes into a loop where it keeps checking to see if the global done flag is set. If not, buffer_remove is used to extract a file descriptor pair (source and destination) from the buffer. The loop ends if an incorrect file descriptor which indicates the termination condition—is retrieved. The function writes data in chunks to the destination file while it is in the loop, reading data from the source file into the buffer. It retries writing if a signal (EINTR) interrupts the write operation. When reading or writing errors occur, the file descriptors are closed and error warnings are printed. It produces a message indicating the completion of the copy operation and closes the source and destination file descriptors after copying. Thread-safe access to the shared variable is ensured by the worker thread locking a mutex to update the global total bytes copied (total bytes copied) upon leaving the loop. Ultimately, it returns and releases the buffer RAM that was allotted. The worker function prototype is defined in the header file worker.h, which also contains the headers required for file operations, memory management, and error handling.

Pseudocode:

```
// Worker thread function
function worker(arg)
  cast arg to buffer_str pointer
  allocate memory for buffer
 if memory allocation fails
    print error and return
 initialize total bytes copied by this thread to 0
 while not done
    get file descriptors from buffer
    if file descriptors are invalid
      break
    while read data from source file into buffer
      if read fails and is not interrupted by signal
        print error
        close file descriptors
       free buffer
        return
      write data from buffer to destination file
      while data remains to be written
        if write fails and is not interrupted by signal
         print error
         close file descriptors
         free buffer
         return
        update write pointer and bytes remaining
      update total bytes copied by this thread
```

if read fails and is not interrupted by signal print error

close file descriptors
print copy completion message

lock mutex
update global total bytes copied
unlock mutex

free buffer
return

5.4 Main.c: The main program in this C file initializes the buffer, creates the worker and manager threads, and handles signal interruptions in order to coordinate the directory copying function. To establish the buffer size, number of worker threads, source and destination directories, and other program parameters, the software first checks and parses the command-line options. It first configures a SIGINT signal handler to gracefully manage interruptions, and then it uses gettimeofday to calculate the operation's execution time. The application launches the manager thread, passes the source and destination directories, and initializes the buffer with the given size. Additionally, the buffer needed to process file descriptors is created for each of the designated number of worker threads. The done flag is set by the handle_signal function, which is designed to detect the SIGINT signal and signal the buffer to wake up any waiting threads. This signals the end of the copying operation. The manager and worker threads work simultaneously, with the workers carrying out the actual copying and the management adding file descriptors to the buffer. The program waits for each worker thread to accomplish its task once the management thread ends. After that, the buffer is destroyed, memory that has been allotted is released, and the overall execution time is calculated. Statistics regarding the operation are printed, including the number of files transferred, directories handled, total bytes copied, and the total time.

Pseudocode:

```
// Function to print usage message
function print_usage(prog_name)
  print "Usage: prog_name <buffer_size> <num_workers> <src_dir> <dst_dir>"
// Signal handler for SIGINT
function handle_signal(sig)
 if sig is SIGINT
    print "Ctrl-C signal received. Copying stopped."
    set interrupted and done flags to 1
    set buffer as done
// Function to print statistics
function print_statistics(num_workers, buffer_size, elapsed)
 calculate elapsed time in minutes, seconds, and milliseconds
 print statistics including number of files, directories, total bytes copied, and total time
// Main function
function main(argc, argv)
 if argc is not 5
   print_usage(argv[0])
    exit with failure
 parse buffer_size, num_workers, src_dir, dst_dir from argv
 if buffer_size or num_workers is less than or equal to 0
    print_usage(argv[0])
    exit with failure
 print welcome message
 set up signal handling for SIGINT
 create destination directory if it does not exist
 start measuring time
 initialize buffer with buffer_size
 print "Files are copying..."
 create and start manager thread with src_dir, dst_dir, and buffer
 allocate memory for worker threads
 if allocation fails
```

print error and exit with failure

create and start worker threads with buffer

wait for manager thread to finish

wait for all worker threads to finish

destroy buffer and free memory

stop measuring time and calculate elapsed time

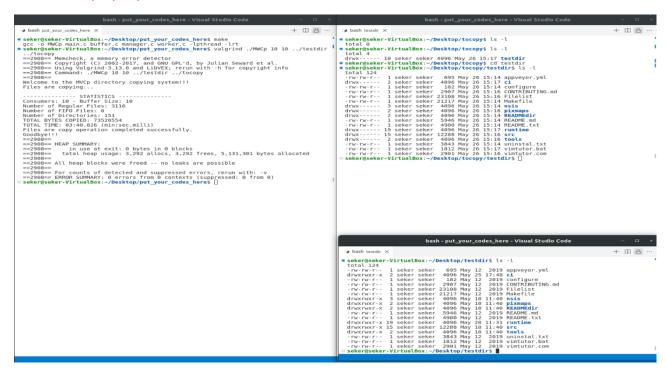
print statistics with num_workers, buffer_size, and elapsed time

print "Files copy operation completed successfully. Goodbye!"

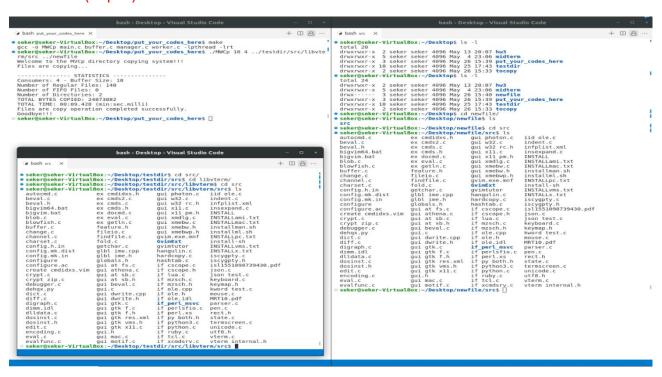
return 0

5. Example Outputs:

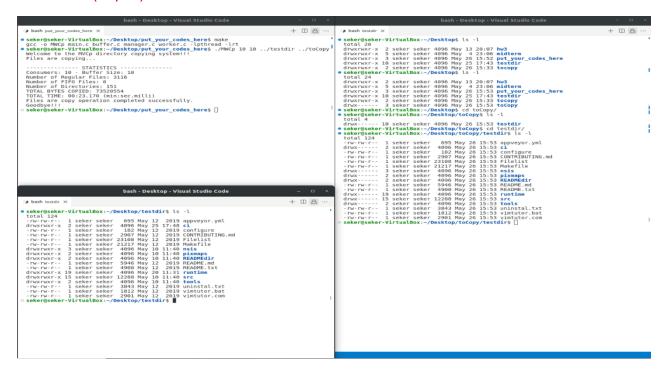
5.1. Test 1 (in pdf):



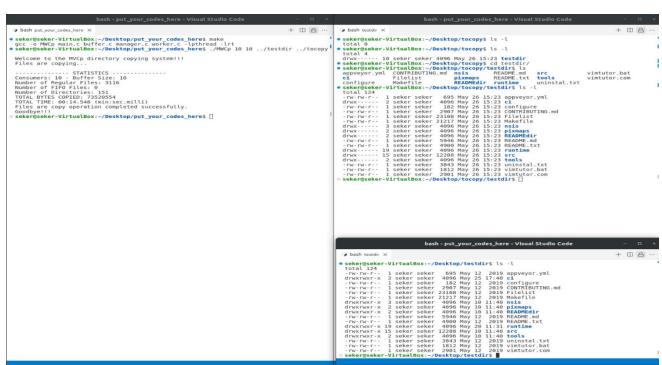
5.2. Test 2 (in pdf):



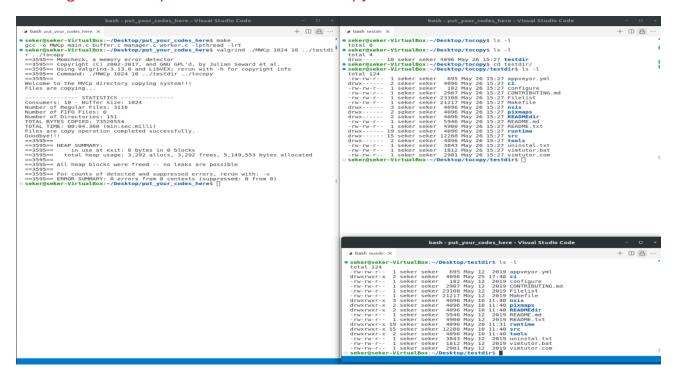
5.3. Test 3 (in pdf):



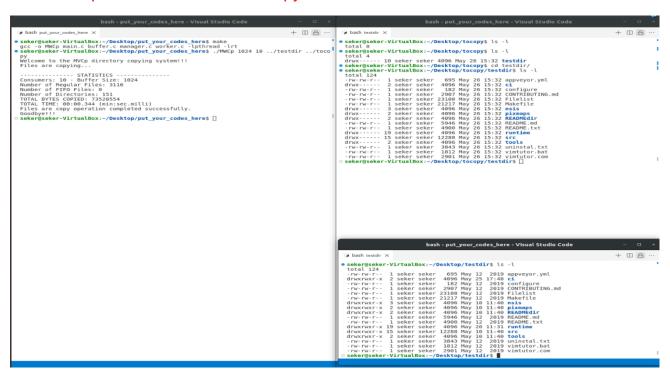
5.4. ./MWCp 10 10 ../testdir ../tocopy:



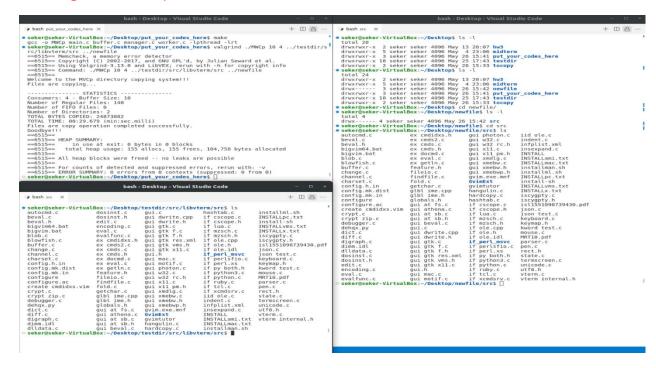
5.5. valgrind ./MWCp 1024 10 ../testdir ../tocopy:



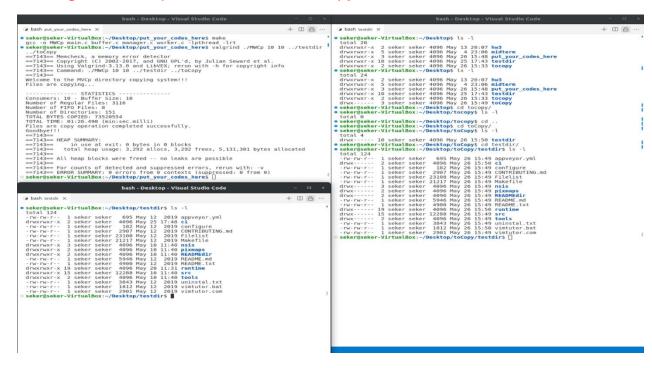
5.6. ./MWCp 1024 10 ../testdir ../tocopy:



5.7. valgrind ./MWCp 10 4 ../testdir/src/libvterm/src ../newfile:



5.8. valgrind ./MWCp 10 10 ../testdir ../toCopy:



5.9. valgrind ./MWCp 10 10 ../testdir ../tocopy (with Ctrl-C signal):

