CSE 344 HOMEWORK #4 Arife Yurtseven 210104004294

Introduction:

The fundamental working principle of the program is built upon a manager thread that scans files and directories in a source directory and adds them to a buffer structure, while worker threads retrieve file information from this buffer to perform the copying process. The main components and the sequence of operations of the program are as follows:

Manager Thread: Scans files and directories in the source directory, adding file information to be copied into the buffer structure.

Worker Threads: Use the file information retrieved from the buffer to copy the files.

Cleanup and Reporting: Once all processes are completed, the program releases resources and prints operation statistics on the screen.

Pseudocode

Global Constants and Structures

```
Define constants:
PATH_MAX_LENGTH = 1024
HASH_SIZE = 65536

Define structures:
file_info_t:
src_path: string[PATH_MAX_LENGTH]
dest_path: string[PATH_MAX_LENGTH]

buffer_t:
buffer: array of file_info_t[BUFFER_CAPACITY]
start: size_t
end: size_t
done: int
mutex: pthread_mutex_t
not_empty: pthread_cond_t
not_full: pthread_cond_t
```

Global Variables

```
file_buffer: buffer_t
workers: array of pthread_t
num_workers: size_t
src_dir: string
dest_dir: string
copied_files_hash: array of unsigned long[HASH_SIZE]

regular_file_count: int
fifo_file_count: int
directory_count: int
total_bytes_copied: size_t
start_time: timeval
end_time: timeval
buffer_size: size_t
```

Helper Functions

```
35 ∨ Define function calculate path hash(path: string) -> unsigned long:
         Initialize hash to 5381
         For each character in path:
             hash = (hash << 5) + hash + character # hash * 33 + c
         Return hash % HASH SIZE
41 ∨ Define function add path to hash(path: string) -> bool:
         Calculate hash index = hash path(path)
         If copied files hash[hash index] == 0:
             Set copied files hash[hash index] = 1
             Return true
         Return false
48 ∨ Define function initialize copied hash():
         For each index from 0 to HASH SIZE-1:
             Set copied_files_hash[index] = 0
52 v Define function signal handler(signal: int):
         If signal == SIGINT:
             Print "SIGINT received. Exiting..."
             Call cleanup()
             Exit program
```

Manager and Worker Threads

```
60 ∨ Define function managerThreadStart(arg: void*):
         Call copy_directory(src_dir, dest_dir)
         Lock file buffer.mutex
         Set file_buffer.done = 1
         Broadcast file buffer.not empty
         Unlock file buffer.mutex
         Return NULL
68 ➤ Define function workerThreadStart(arg: void*):
         While true:
             Lock file buffer.mutex
             While file buffer.start == file buffer.end and file buffer.done == 0:
                 Wait on file buffer.not empty with file buffer.mutex
                 Unlock file buffer.mutex
                 Break loop
             Dequeue file from file buffer
             Signal file buffer.not full
             Unlock file buffer.mutex
             Print "Copying file from src_path to dest_path"
             Call copy_file(file.src_path, file.dest_path)
         Return NULL
```

File and Directory Copying

```
Define function copy_file(src_path: string, dest_path: string):

Open src_path for reading

If failed, print error and return

Get file status of src_fd

Open dest_path for writing, create if necessary, truncate if exists

If failed, print error, close src_fd and return

while reading bytes from src_fd:

Write bytes to dest_fd

Increment total_bytes_copied by bytes_read

Close src_fd and dest_fd

If file is regular:

Increment regular_file_count

Belse if file is FIFO:

Increment fifo_file_count

Define function copy_directory(src_dir: string, dest_dir: string):

Open src_dir

If failed, print error and return

For each entry in src_dir:

If entry is "." or "..", continue

Construct src_path and dest_path

If entry is directory:

Create directory dest_path

Increment directory_count

Recursively call copy_directory(src_path, dest_path)

flse:

Lock file_buffer.mutex

while buffer is full:

Wait on file_buffer.not_full with file_buffer.mutex

Enqueue file into file_buffer.end + 1) % buffer_size

Signal file_buffer.mutex

Unlock file_buffer.mutex

Close src_dir

Close src_dir

Close src_dir
```

Print and Cleanup

```
123 V Define function printOutputs():

124 Calculate elapsed time in minutes, seconds, and milliseconds

125 V Print statistics:

126 Number of workers

127 Buffer size

128 Number of regular files

129 Number of FIFO files

130 Number of directories

131 Total bytes copied

132 Total time taken

133

134 V Define function cleanup():

135 Destroy file_buffer.mutex

136 Destroy file_buffer.not_empty

137 Destroy file_buffer.not_full

138 Free workers array
```

Main Function

calculate_hash_path:

This function generates a unique hash value for a given file path. Each character in the file path is processed, thus influencing the hash value. As a result, this hash value can be used as a unique identifier for the file

add_path_to_hash:

After calculating the hash value of a given file path, this value is stored in a hash table. If this file path has not been added to the hash table before, the function successfully indicates that a new entry has been made.

initialize_copied_hash:

At the start of the program, all hash values are reset, clearing the hash table. This ensures that old data is cleaned out before the copying process begins.

manager ThreadStart:

Functions as the manager thread. This thread initiates the file and directory copying process from the specified source directory to the destination directory. Once all copying tasks are completed, it wakes up all worker threads to signal the end of the process.

workerThreadStart:

Worker threads take the files designated by the manager thread to be copied and transfer them to the target location. If there are no more files to be copied and all tasks are completed, the thread terminates itself.

copy_file:

This function opens a specific source file, reads its content, and writes this content to the target file. Error checking is performed during the file copying process, and a counter for successful copies is updated.

copy_directory:

Scans all subdirectories and files in the source directory. If a sub-item is a directory, the same structure is created in the target directory, and the copying process continues recursively for this new directory. For files, they are added to the queue to be processed by worker threads.

printOutputs:

After the completion of the copying process, various statistics are printed on the screen. These statistics include the number of files copied, the total bytes copied, and the total duration of the process.

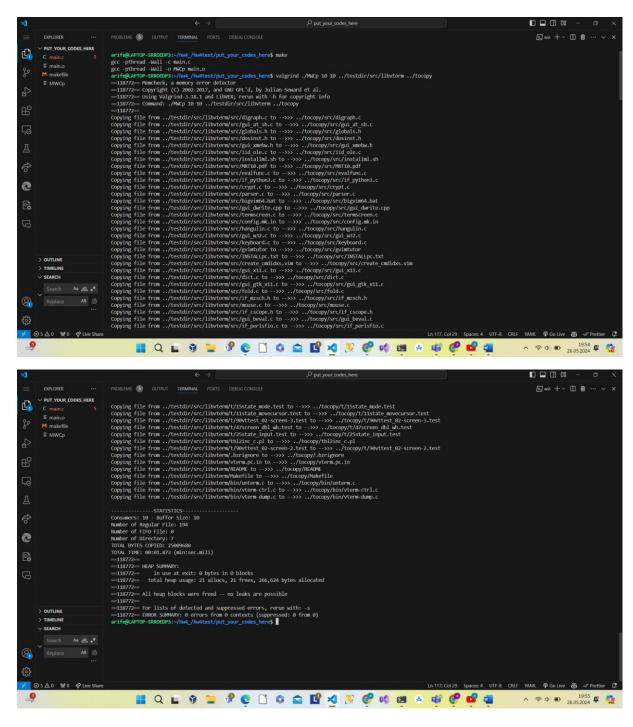
signal_handler:

Ensures the program terminates safely when it encounters an external interrupt signal (e.g., being stopped with CTRL+C).

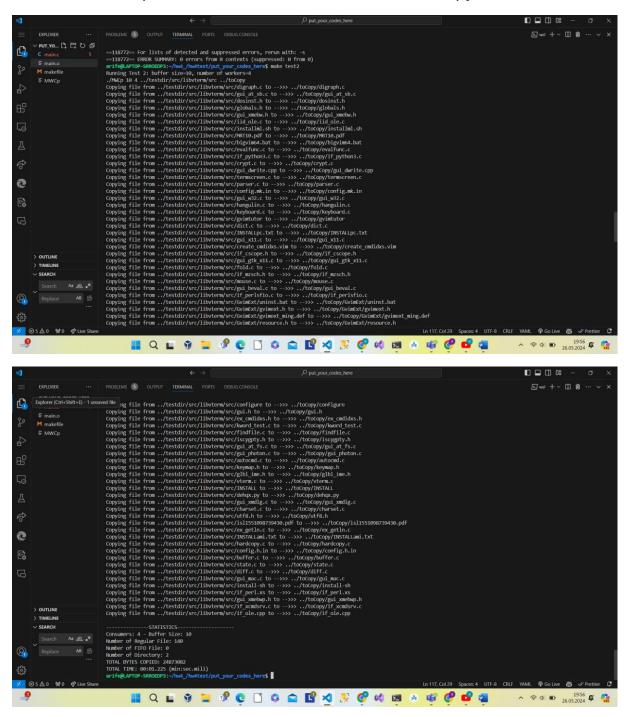
cleanup:

At the end of the program, all used resources are released. Mutexes are destroyed and memory allocated for thread arrays is reclaimed.

Test1 // also works as make test1



Test2 // ./MWCp 10 4 ../testdir/src/libvterm/src ../toCopy



Test3// ./MWCp 10 10 ../testdir ../toCopy

