WPI

CS 3516 - Computer Networks

Project 1 (100 points)
Assigned: Tuesday, October 25, 2022
Checkpoint: Tuesday, November 1, 2022 at 11:59pm
Due: Tuesday, November 8, 2022 at 11:59pm

# Project 1: Quick Response Codes

Quick Response (QR) codes have become more popular given the rise of smart phones. A QR code can encode a URL or other information in a two-dimensional barcode. With a QR-code reader, a smart phone can use its built-in camera to take a photo of the QR-code and load the URL. This allows a user to access a long URL without having to type it and reducing the risk of typos.

Figure 1: Example QR code.

Mobile devices typically translate the QR code from an image file to the text output locally. In this project, we will have the mobile device offload this functionality to a server that will translate the QR code and provide the client with a URL to retrieve. This is mainly for educational purposes to motivate the use of a client-server architecture. However, this approach shows how computationally intensive processes can be outsourced from a mobile device to save battery power.

In this project, we will create a QR code server for public use. Additionally, we will create a simple client to send the images and fetch the URLs. The goal is to become familiar with socket programming while learning more about how QR codes are used. Your program must be written in C/C++ and must work on a standard Ubuntu 20.04 Linux virtual machine (which is what the graders will be using).

## Expected QR code server functionality

Your server must run on the VM on a specified port. The operator of this server would invoke it as: `./QRServer [option1, ..., optionN]`. Below is a list of possible options. If an option is omitted, the default value should be used.

- PORT [port_number]
- RATE [number_requests] [number_seconds]
- MAX_USERS [number_of_users]
- TIME_OUT [number_of_seconds]

We now specify the details of these options:

- **Listening port (PORT):** The server should listen on the specified TCP port to receive QR codes. Possible values: 2000 - 3000. Default setting: 2012.
- **Rate limiting (RATE):** Without reasonable safeguards, public servers could be attacked. To prevent this, the server operator should be able to specify the number of QR codes, specified by "number requests", that should be allowed in a "number seconds" timeframe. Requests issued in excess of this rate should be discarded, and the occurrence of this condition logged at the server. The client should **not** be disconnected for violating rate limiting. Instead, they should receive an error message indicating why their command was not processed. Default: 3 requests per user per 60 seconds.

- **Maximum number of concurrent users (MAX_USERS):** Only a specified number of users should be able to connect to the QR code server at a given time. Above this threshold, connecting users should receive an error message indicating that the server is busy before terminating the connection. Such refused connections should be logged at the server. When a user disconnects, a slot should be opened for another connection. Default: 3 users.
- **Time out connections (TIME_OUT):** The server must monitor the connection for inactivity. If more than the specified number of seconds elapse since the last client interaction, the server must 1) report to the user that a time-out occurred using the return code and server message format specified, 2) close the TCP connection, and 3) log the event at the server. Default: 80 seconds.

## Administrative Log

The administrative log reports user activity and system behavior for the server operator. All entries should be prefixed with the time and client IP address in **YYYY-MM-DD HH:mm:ss xxx.xxx.xxx.xxx** format, where 1) "YYYY-MM-DD" is the year, month, and day, 2) "HH:mm:ss" is the time of day in 24 hour format (HH can take values from 00-23), and 3) "xxx.xxx.xxx.xxx" represents the IP address of the client in dotted decimal format. Valid QR code interactions and invalid actions from clients should be logged. The log file must also contain information about server start-ups, connections, and disconnections from users, events as described by the RATE/MAX_USERS instructions, and any other information regarding user behavior or system events.

## Expected Client and Server Interaction

Clients should connect to the QR code server using TCP. The server should provide separate and concurrent sessions for each client. This can be accomplished using a new process for each user (via fork) or via threads. Upon establishing a connection, the client is allowed to perform zero or more interactions.

In an interaction with the client, the client begins by sending the size of the QR code image file, followed by a byte array of the indicated size containing the image file contents. The server then replies with the URL that the QR code represents. The details of these messages follows:

- **Client Message:** [x] [y] (where x is a 32 bit unsigned integer representing the file size and y is a binary byte array of the size indicated that contains the file contents)
- **Server Message:** [x] [y] [z] (where x is a 32 bit unsigned integer representing the return code, y is a 32 bit unsigned integer representing the URL character array's length, and z is the character array of the size indicated that **contains**[1] the URL)

The server return codes are used to allow the server to reply to the client to provide additional feedback about the situation. Below are the valid return codes:

- 0 - Success. The URL is being returned as specified below.
- 1 - Failure. Something went wrong and no URL is being returned. The character array length is set to 0 and no character array is transmitted. This condition can be valid if the image uploaded does not represent a valid QR code or the client violates our network security requirements.
- 2 - Timeout. The connection is being closed. A human-readable text message is created and supplied as the character array. The character array length is set to the length of this human readable message.
- 3 - Rate Limit Exceeded. An error message about the rate limit being exceed is set in the character array with the size set to the character array.

[1]The decoding library may include other output in addition to the URL. That can be output as well; students do NOT need to parse the output to just return the URL.

### Network Security

In a networking environment, you cannot assume that your clients will be acting in good faith. These attackers may be attempting to compromise your server through buffer overflows or denial of service attacks. Your server must protect itself:

1. The server must set a reasonable upper-bound for the image size it is willing to process. In the protocol specified above, the client could specify and upload a file of size ($2^{32} - 1 = 4{,}294{,}967{,}295$ bytes or roughly 4GB). This upload could be used to reduce the server's available bandwidth while reserving a user connection slot, preventing other clients from connecting. Instead, you must set a lower limit to requests, returning a failure code for requests specifying a larger file size.
2. The server must not process any image file bytes that exceed the size specified by the client. The server must process only the the portion of the file specified by the client, discarding all other bytes that arrive before the server's response.
3. The server must prohibit clients from blocking other clients from connecting by using the timeouts specified above.

In each of these cases, we must resolve client violations within the above protocol. Do not disconnect a client simply for violating the file size upload limits or sending extraneous bytes.

### Concurrency

Each connecting client must have a separate process or thread to handle its interactions. No possible execution paths of these concurrent threads or processes should lead to an error. Specifically, ensure that you protect all functions and library calls explicitly unless they are noted to be thread safe. As an example, when updating the administrative log file, make sure that events corresponding to multiple clients are recorded correctly. All memory should be freed when your server program exits. Specifically, there should not be any memory leaks or zombie threads or processes. Do not rely on the OS to clean up terminated threads or processes.

## Resources and Restrictions

You are required to use either C or C++ for this assignment. You may use the C++ STL and TR1. You may use the native Linux/POSIX socket system calls, but you may **not** use any other socket libraries. No credit will be given to solutions that do not adhere to these requirements, so contact the instructor or the TAs for any clarification before using outside resources. These restrictions are being made so you become familiar with the details of lower-level socket programming.

While QR codes make a good motivational example of socket usage, understanding how they are created or parsed is not the goal for this project. Accordingly, students should use the **ZXing** library for generating and decoding QR codes. To simplify invocation, simply download the supplied jar files at `https://cerebro.cs.wpi.edu/cs3516/project1_jar.tar`, which will act as a decoder for QR codes. Before using this library, get a recent version of Java (e.g., `sudo apt install default-jre` on Ubuntu 20.04). Then, run: `java -cp javase.jar:core.jar com.google.zxing.client.j2se.CommandLineRunner [image_file]` in the directory where the files are decompressed. That will execute the decoder. In your server program, you will invoke this utility using system calls such as **system** or **exec**. An example of the output from that decoder follows:

```
file:/home/staneja/cs3516/project1/project1_qr_code.png (format: QR_CODE, type: URI):
Raw result:
http://web.cs.wpi.edu/~staneja/cs3516/
Parsed result:
```

```
http://web.cs.wpi.edu/~staneja/cs3516/
Found 4 result points.
  Point 0: (35.0,215.0)
  Point 1: (35.0,35.0)
  Point 2: (215.0,35.0)
  Point 3: (197.0,197.0)
```

For testing purposes, you may use the above QR code which decodes to the URL for the course Web page. You can generate additional examples by using the Web application at `https://www.qr-code-generator.com/` and providing a URL as input.

As always, you are encouraged to avail yourself to Internet resources and Linux manual pages when completing the assignment. Socket tutorials, such as the optional course textbook, will be helpful in understanding socket programming. However, you **must ensure that you write your own code** and **explicitly mention any resources you use** including the textbook and discussions with individuals outside of your group. In case of any questions about making use of a specific Web resource for the project, seek clarification in advance via the forum.

The following system calls will likely be useful in completing the assignment: **listen**, **accept**, **bind**, **recv**, **send**, **setsockopt**, **socket**, **close**, **gethostbyname**, **getprotobyname**, **htons**, **ntohs**, **fork**, **exec**, **dup2**, **waitpid**, **system**. Before getting started, you should consult the manual pages for each of these system calls. Additionally, you would need functions to perform thread or process synchronization.

### Road Map

Like all programming projects, you should build your program incrementally. Below are a reasonable set of milestones in completing the project:

1. Create a single threaded server that accepts text from a client in a file, displays the text and disconnects the client.
2. Add support in the server for decoding QR codes.
3. Add support for accepting binary transmissions (this should be a straight-forward change from Step 1).
4. Support concurrent clients.
5. Add error checking and logging functionality.
6. Add security features.

## Points of Clarification

- The command line arguments should be readable in any order. You are free to use libraries like getopt if you would like. However, you should not have just a string of numbers that have to be in a set order. Since getopt does not let you specify two parameters to an option, you may instead use two options RATE_MSGS and RATE_TIME to specify that the limit is RATE_MSGS every RATE_TIME seconds.
- The "all memory should be freed when your server program exits" requirement is largely to catch when you fork off server processes. You should do proper garbage collecting rather than relying on the kernel to clean up for you. This should be true for your main process as well, but that's mostly just to encourage good coding practices.
- You need to create the client portion of this protocol as well and submit it for grading. While the specifications focused on the server portion, the client should provide what the server expects.

- "Error checking" in the project rubric refers to proper handling of clients that send files that are too large for the server or that send files that are not valid QR codes (e.g., the Java library cannot find a URL).

## Checkpoint Contributions

Students must submit work that demonstrates substantial progress towards completing the project on the checkpoint date. Substantial progress is judged at the discretion of the grader to allow students flexibility in prioritizing their efforts. However, as an example, any assignment in which the first two roadmap items are completed will be considered as making substantial progress. **Projects that fail to submit a checkpoint demonstrating significant progress will incur a 10% penalty during final project grading.**

## Deliverables and Grading

When submitting your project, please include the following:

- All of the files containing the code for all parts of the assignment.
- One file called `Makefile` that can be used by the `make` command for building the executables. It should support the "`make clean`" command and the "`make all`" command.
- A document called `README.txt` explaining your project and anything that you feel the teaching staff should know when grading the project. Only plaintext write-ups are accepted.
- The .jar files and their tested image files.

Please compress all the files together as a single .zip archive for submission. As with all projects, please **only use standard zip files** for compression; **.rar, .7z, and other custom file formats will not be accepted**.

The project programming is only a portion of the project. Students should use the following checklist in turning in their projects to avoid forgetting any deliverables:

1. Sign up for a project partner or have one assigned (URL: `https://ia.wpi.edu/cs3516/request_teammate.php`),
2. Submit the project code and documentation via InstructAssist (URL: `https://ia.wpi.edu/cs3516/files.php`),
3. Complete your Partner Evaluation (URL: `https://ia.wpi.edu/cs3516/evals.php`), and

A grading rubric will be provided to give you a guide for how the project will be graded. No points can be earned for a task that has a prerequisite unless that prerequisite is working well enough to support the dependent task. Students will receive a scanned markup of this rubric as part of their project grading feedback. Students in teams are expected to contribute equally; unequal contributions may yield different grades for the team members. Students who work in teams but do not submit a Partner Evaluation by the project deadline will incur up to a 5% penalty on the project grade.