

## Programming Project, C++ Programming

The programming project is compulsory. You shall work in groups of two to four people, and it is up to you to find group partners.

You shall describe your results in a report and submit the report and your programs. Details on submission are in section 4.

Questions regarding the project can be directed to `sven.robertz@cs.lth.se`. Answers to questions of general interest will be posted in teams.

# A News System Implementation

## 1 Background

Usenet News is an old (started in 1986) system for sharing information over the web. Information is distributed in the form of “articles” that are posted to “newsgroups”. Articles are distributed to news servers, and a user can hook up to a news server and read articles and also post articles in newsgroups and create new newsgroups. The user runs a news client which uses the NNTP<sup>1</sup> protocol to communicate with the server.

In this project, you shall develop your own news system, consisting of a news server and a news client. Both programs shall be written in C++. The server is dedicated and does not share news with other servers. Instead of NNTP you will use the communication protocol that is defined in section 6.

## 2 System Requirements

### 2.1 General Requirements

The system consists of a server that handles a database containing newsgroups and articles and a client that accepts commands from the user and communicates with the server. Several clients may be connected to the server simultaneously.

The following tasks can be performed:

- List newsgroups.
- Create and delete newsgroups.
- List articles in a newsgroup.
- Read, write and delete articles in a newsgroup.

The system keeps track of the title and the author of each article. It cannot handle subject threads, follow-up articles or similar advanced concepts that are available in standard NNTP-based news implementations.

The communication between the server and the client follows the messaging protocol that is defined in section 6.

There are no security requirements on the system. For instance, any user can delete newsgroups and articles, even if he or she is not the creator.

### 2.2 Server Requirements

Detailed requirements on the server:

- There are two versions of the server: the first version uses a database stored in primary memory, the second a database stored on disk.
- The code shall be structured so that all code that is not directly related to the database should be agnostic of, and independent of, the database implementation. There should be no duplicated code in the implementation of the general server functionality.
- The in-memory version of the server starts from scratch each time it is invoked and builds the database in primary memory.
- The disk version of the database remembers the database between invocations. Changes to the database are immediately reflected on disk. See section 3.2.

---

<sup>1</sup> See <http://www.faqs.org/rfcs/rfc3977.html>.

- Each newsgroup has a unique name. A newsgroup also has a unique identification number, greater than zero. Identification numbers may not be reused.
- Each article has a title, an author and an article text. The article names need not be unique. An article also has an identification number, which must be unique in the newsgroup. Identification numbers may not be reused.
- Listings of newsgroups and articles are in chronological order, with the oldest item first.
- There are no limitations on the number of newsgroups or the number of articles in a newsgroup.
- There are no limitations on the length of newsgroup titles, article titles, author names or article texts.
- If a client misbehaves, for example by not following the messaging protocol, it is immediately disconnected by the server.
- The server tries to handle all errors. If it cannot recover from an error, it terminates with an informative error message.
- The disk database should be implemented using the standard library (either `<filesystem>` in C++-17, or the C standard library. You may not use an existing database library.

### 2.3 Client Requirements

Requirements on the client:

- The client reads commands from the keyboard, communicates with the server and presents the replies from the server on the terminal.
- The client is easy to use and contains informative help texts. No manual is necessary to use the client program.
- The client can handle articles with more than one line in the message body.
- The client tries to handle all errors. If it cannot recover from an error, it terminates with an informative error message.

## 3 Development Procedure

### 3.1 General Advice

The following development procedure is recommended:

- Start by writing the in-memory version of the server. Use the test clients described in section 7.1 during development.
- Then write the client, using your own server.
- Finally, write the disk version of the server. Your server design should be such that you only have to change the part of the server that deals with the database.

### 3.2 The Server

Advice for development of the server:

- Use the classes described in section 5.2 for the low-level communication between the server and the clients.
- You must use the messaging protocol described in section 6 for communication.

- It is a good idea to define a class `MessageHandler` to handle the communication on “low protocol level”. It should be possible to use this class also in the client. See the file `MessageHandler.java` in the java code for the test clients, for ideas on how to write such a class. That code is available at <https://gitlab.com/cs-lth-cpp/news-server-test-java>
- Write a class that functions as an interface to the database, so it is easy to switch between the in-memory version and the disk version of the server. The rest of the server shall be written against this abstract interface to the database.
- A production version of the server would most probably be multithreaded with one thread for each client. In this project, the server should be singlethreaded; the communication delays that are caused by this are acceptable.
- Separate the communication and the database parts of the server as far as possible. For instance, you should not have to modify the database in any way if the messaging protocol is changed.
- You may choose any structure for storing the database in the disk version of the database, as long as you implement the storage yourselves (you may not use a database library), and the implementation fulfills the requirements. One method is to dedicate a root directory to the database. A newsgroup is represented as a directory in the root directory, and the articles as regular files in the group directories. In C++-17, the `<filesystem>` header, which standardizes operations on the filesystem, was introduced. Previous to that, you had to rely on the C standard library functions. You can use either of these options: either study the `<filesystem>` header, or the system calls `mkdir` (chapter 2 of the man pages) and `remove`, `opendir`, `readdir`, and `closedir` (chapter 3C).

## 4 Submission of Results

You shall hand in a written report that describes your system, and also all files necessary to build your system. Don't hand in anything before everything (for example, the tests in `TestServer2`) works. *Please note that this only tests the correct operation, you need to ensure that your system also handles things like wrong user input gracefully.* The latest date for submission is in the course plan.

The report must be well structured, complete, easy to understand, well formulated, etc. It should (minimally) contain:

- A cover sheet with the title of the project, the course name, your names and e-mail addresses.
- A detailed description of your system design, both for the server and the clients. Preferably use UML diagrams to give an overview of the design. (It is not necessary that you list attributes and methods in these diagrams.) You must also describe the classes, at least as far as stating the responsibilities of each class.

Also give an overview of the dynamics of the server, i.e., trace an interaction between a client and the server from the point that the server receives a command until it sends the reply. UML sequence diagrams are good for this purpose.

- Conclusions: requirements that you fulfill, problems that you haven't succeeded in solving, etc. If you have found that the system ought to have more features you should elaborate on this. Also discuss known weaknesses in your design or implementation, and state if there is any code you want feedback on. Any suggestions for improvements to the project are also welcome.

When you are ready to submit your report and your programs, do the following:

1. Produce a *pdf* file of your report.
2. Create a directory *username* (i.e., the username of one of the project members<sup>2</sup>) with subdirectories *src* and *bin*. Collect all source files in the *src* directory and write a makefile. `make all` should create the executables (the two versions of the server and the client), `make install` should copy the executables to the *bin* directory. Give additional instructions, if necessary, in a *README* file in the *username* directory. `tar` and `gzip` the *username* directory. Remove all executables and object files from the directories before you do this. Also, please avoid any swedish or special characters in file names.
3. Submit the *.pdf* file and the *.tar.gz* file. Instructions for submission will be published under "Project" on the course web page. (Attach the files separately; do *not* place the *.pdf* file in the *.gz* file).

---

<sup>2</sup> It does not matter which name, the point of this is to have a unique directory name to avoid name collisions when marking and identify the group

## 5 Low-Level Communication

### 5.1 Communication Between Unix Computers

Unix computers communicate via numbered *ports*. Some ports are dedicated to special purposes, but all ports with numbers above 1024 are freely available for application programs. Two programs that wish to communicate must have agreed on which port to use. The server listens for traffic on the specified port, and clients may connect to this port on the server's computer.

Once a connection has been established, messages are exchanged via *sockets*. A program writes to a socket and another program reads from another socket.

### 5.2 Connection Classes

You shall use two classes in your system: *Connection*, which handles a connection to another program (in principle, a socket), and *Server*, which handles several simultaneous connections. Each client creates a *Connection* object for communication with the server. The server creates a *Server* object that keeps track of the clients with the help of a *Connection* object for each client.

Class specifications (only the public interface):

```
/* A ConnectionClosedException is thrown when a connection is closed */
struct ConnectionClosedException {};

/* A Connection object represents a connection (a socket) */
class Connection {
    friend class Server;

public:
    /* Establishes a connection to the computer 'host' via
       the port 'port' */
    Connection(const char* host, int port);

    /* Creates a Connection object which will be initialized
       by the server */
    Connection();

    /* Closes the connection */
    virtual ~Connection();

    /* Returns true if the connection has been established */
    bool isConnected() const;

    /* Writes a character */
    void write(unsigned char ch) const;

    /* Reads a character */
    unsigned char read() const;

    /* Connection cannot be copied */
    Connection(const Connection&) = delete;
    Connection& operator=(const Connection&) = delete;
    Connection(Connection&&);
    Connection& operator=(Connection&&) = delete;
};
```

---

```

/* A server listens to a port and handles multiple connections */
class Server {
public:
    /* Creates a server that listens to a port */
    explicit Server(int port);

    /* Removes all registered connections */
    virtual ~Server();

    /* Returns true if the server has been initialized correctly */
    bool isReady() const;

    /* Waits for activity on the port. Returns a previously registered
       connection object if an existing client wishes to communicate,
       nullptr if a new client wishes to communicate */
    std::shared_ptr<Connection> waitForActivity() const;

    /* Registers a new connection */
    void registerConnection(const std::shared_ptr<Connection>& conn);

    /* Deregisters a connection */
    void deregisterConnection(const std::shared_ptr<Connection>& conn);

    /* Servers cannot be copied */
    Server(const Server&) = delete;
    Server& operator=(const Server&) = delete;
    Server& operator=(Server&&) = delete;
    Server(Server&&);
};

```

### 5.3 A Communication Example

In the following programs, the client reads 32-bit integers from the terminal and sends them to the server. The server reads an integer and responds with a string, *Positive/Zero/Negative*, depending on the sign of the number. Notice that there are rudiments of a protocol here: integers are transmitted as four bytes with the high-order byte first and the end of a variable-length string is marked with a special character.

```

/* myclient.cc: sample client program */
#include "connection.h"
#include "connectionclosedexception.h"

#include <cstdlib>
#include <iostream>
#include <stdexcept>
#include <string>

using namespace std;

/*
 * Send an integer to the server as four bytes.
 */
void writeNumber(const Connection& conn, int value)
{
    conn.write((value >> 24) & 0xFF);
    conn.write((value >> 16) & 0xFF);
    conn.write((value >> 8) & 0xFF);
    conn.write(value & 0xFF);
}

```

```
/* Read a string from the server. */
string readString(const Connection& conn)
{
    string s;
    char ch;
    while ((ch = conn.read()) != '$') {
        s += ch;
    }
    return s;
}

/* Creates a client for the given args, if possible.
 * Otherwise exits with error code.
 */
Connection init(int argc, char* argv[])
{
    if (argc != 3) {
        cerr << "Usage: myclient host-name port-number" << endl;
        exit(1);
    }

    int port = -1;
    try {
        port = stoi(argv[2]);
    } catch (exception& e) {
        cerr << "Wrong port number. " << e.what() << endl;
        exit(2);
    }

    Connection conn(argv[1], port);
    if (!conn.isConnected()) {
        cerr << "Connection attempt failed" << endl;
        exit(3);
    }

    return conn;
}

int app(const Connection& conn)
{
    cout << "Type a number: ";
    int nbr;
    while (cin >> nbr) {
        try {
            cout << nbr << " is ...";
            writeNumber(conn, nbr);
            string reply = readString(conn);
            cout << " " << reply << endl;
            cout << "Type another number: ";
        } catch (ConnectionClosedException&) {
            cout << " no reply from server. Exiting." << endl;
            return 1;
        }
    }
    cout << "\nexiting.\n";
    return 0;
}

int main(int argc, char* argv[])
{
    Connection conn = init(argc, argv);
    return app(conn);
}
```



---

```

/* myserver.cc: sample server program */
#include "connection.h"
#include "connectionclosedexception.h"
#include "server.h"

#include <cstdlib>
#include <iostream>
#include <memory>
#include <stdexcept>
#include <string>

using namespace std;

/*
 * Read an integer from a client.
 */
int readNumber(const shared_ptr<Connection>& conn)
{
    unsigned char byte1 = conn->read();
    unsigned char byte2 = conn->read();
    unsigned char byte3 = conn->read();
    unsigned char byte4 = conn->read();
    return (byte1 << 24) | (byte2 << 16) | (byte3 << 8) | byte4;
}

/*
 * Send a string to a client.
 */
void writeString(const shared_ptr<Connection>& conn, const string& s)
{
    for (char c : s) {
        conn->write(c);
    }
    conn->write('$');
}

Server init(int argc, char* argv[])
{
    if (argc != 2) {
        cerr << "Usage: myserver port-number" << endl;
        exit(1);
    }

    int port = -1;
    try {
        port = stoi(argv[1]);
    } catch (exception& e) {
        cerr << "Wrong format for port number. " << e.what() << endl;
        exit(2);
    }

    Server server(port);
    if (!server.isReady()) {
        cerr << "Server initialization error." << endl;
        exit(3);
    }
    return server;
}

```

```
int main(int argc, char* argv[])
{
    auto server = init(argc, argv);

    while (true) {
        auto conn = server.waitForActivity();
        if (conn != nullptr) {
            try {
                int    nbr = readNumber(conn);
                string result;
                if (nbr > 0) {
                    result = "positive";
                } else if (nbr == 0) {
                    result = "zero";
                } else {
                    result = "negative";
                }
                writeString(conn, result);
            } catch (ConnectionClosedException&) {
                server.deregisterConnection(conn);
                cout << "Client closed connection" << endl;
            }
        } else {
            conn = make_shared<Connection>();
            server.registerConnection(conn);
            cout << "New client connects" << endl;
        }
    }
    return 0;
}
```

## 6 Message Protocol

A message consist of a number of characters (bytes). A message has the following format:

```
commandbyte parameters endbyte
```

A parameter is a string or a number, in one of the following formats:

```
string_p: PAR_STRING N char1 char2 ... charN // N is the number of characters
num_p: PAR_NUM N // N is the number
```

PAR\_STRING and PAR\_NUM are one-byte constants. A number N is transmitted as four bytes with the most significant byte first.

The list below shows the format of all messages. Messages from the server to a client contain a status indication: ANS\_ACK if the command was executed successfully, ANS\_NAK followed by an error code otherwise. The names of the error codes are hopefully self explanatory. The expression [x | y] means x or y, [x y] means x followed by y, and x\* means zero or more repetitions of the expression x.

1. List newsgroups. The reply contains the number of newsgroups followed by the identification numbers and titles of the groups.

```
COM_LIST_NG COM_END
ANS_LIST_NG num_p [num_p string_p]* ANS_END
```

2. Create a newsgroup. The title of the group is sent as a parameter.

```
COM_CREATE_NG string_p COM_END
ANS_CREATE_NG [ANS_ACK | ANS_NAK ERR_NG_ALREADY_EXISTS] ANS_END
```

3. Delete a newsgroup. The identification number of the group is sent as a parameter.

```
COM_DELETE_NG num_p COM_END
ANS_DELETE_NG [ANS_ACK | ANS_NAK ERR_NG_DOES_NOT_EXIST] ANS_END
```

4. List articles in a newsgroup. The identification number of the group is sent as a parameter. The reply contains the number of articles, followed by the identification numbers and titles of the articles.

```
COM_LIST_ART num_p COM_END
ANS_LIST_ART [ANS_ACK num_p [num_p string_p]* |
ANS_NAK ERR_NG_DOES_NOT_EXIST] ANS_END
```

5. Create an article. The identification number of the group is sent as a parameter, followed by the article title, author and text.

```
COM_CREATE_ART num_p string_p string_p string_p COM_END
ANS_CREATE_ART [ANS_ACK | ANS_NAK ERR_NG_DOES_NOT_EXIST] ANS_END
```

6. Delete an article. The group and article identification numbers are sent as parameters.

```
COM_DELETE_ART num_p num_p COM_END
ANS_DELETE_ART [ANS_ACK |
ANS_NAK [ERR_NG_DOES_NOT_EXIST | ERR_ART_DOES_NOT_EXIST]] ANS_END
```

7. Get an article. The group and article identification numbers are sent as parameters. The reply contains the title, author, and text of the article.

```
COM_GET_ART num_p num_p COM_END
ANS_GET_ART [ANS_ACK string_p string_p string_p |
             ANS_NAK [ERR_NG_DOES_NOT_EXIST | ERR_ART_DOES_NOT_EXIST]] ANS_END
```

All symbolic constants are defined in the scoped enum `Protocol` (see section 7.2).

## 7 Test Programs and Access to Communication Classes

### 7.1 Test Programs

Two test programs are available: `TestServer1` and `TestServer2`. Both programs are written in Java, and `.jar`-files may be downloaded from the course homepage (the files are also available in the directory `/usr/local/cs/cpp`). Execution:

```
java -jar TestServer1.jar hostname port
java -jar TestServer2.jar hostname port
```

The programs open a connection to the server running on `hostname` (which may be `localhost`) on the port `port`. `TestServer1` is intended to be used during development of the server. It contains buttons to execute the commands that the server should be able to handle.

`TestServer2` is a more complete program that systematically tests all commands, including erroneous commands (creating a newsgroup with a duplicate name, deleting a non-existing newsgroup, and so on).

### 7.2 Access to Communication Classes

The communication classes are defined in `connectionclosedexception.h`, `connection.h`, `server.h`, and `protocol.h`. These files, together with the corresponding `.cc` files, are in a file `clientserver.tar.gz`, which can be downloaded from the course homepage. There is also a makefile that compiles the files and creates a library `libclientserver.a`. The test programs from section 5.3 are in a subdirectory `test`. Check the `README` file for possible additional information on linking and other practical details.

The file `protocol.h` contains an enum class with the following outline:

```
enum class Protocol {
    /* Command codes, client -> server */
    COM_LIST_NG    = 1, // list newsgroups
    ...

    /* Answer codes, server -> client */
    ANS_LIST_NG    = 20, // answer list newsgroups
    ...

    /* Parameters */
    PAR_STRING     = 40, // string
    ...

    /* Error codes */
    ERR_NG_ALREADY_EXISTS = 50, // newsgroup already exists
    ...
};
```

The test programs rely on the same constant definitions, so the definitions must not be changed.