csc456 Reliable Systems
Project Requirements

Dr. Steven P. Crain

Due October 23, 2014 at 5 p.m.

In this assignment, you will gain experience designing a reliable system.

1 Overview

You will design a system that searches a set of files to find files that contain a provided sentence.

Example use case: a teacher regularly gives the same writing assignment, and wants to make sure students do not copy the papers from another student. The teacher makes a directory for each time she teaches the class, and places the student papers in the corresponding directory. As the teacher grades a paper, she picks a few distinctive words from the paper and searches for other papers containing the same words.

The system you design will be modified to simulate a relatively high failure rate for each component. The failure rate for each function call will be approximately $\lambda = 1000000/n^2$ per time the function is called, where n is the number of lines of code in the function.

2 Functional Requirements

1. The system will be provided with a configuration file containing the relative paths of up to 1000 directories, one directory per line.

2. The system must accept a line of input from the user containing a set of words to be searched, called the *query words*.

3. The system must report to the user the names of all files contained in any of the configured directories that contain all of the *query words* anywhere in the file (not necessarily together).

4. The system must keep track of when it last updated its model of what the files contain, called the *model age*.

5. The system must include the *model age* whenever it reports a list of files to the user.

3 Non-Functional Requirements

1. The system must not report a file name to the user unless it contained all of the query words when it was last read by the system.

2. The system must report a file name to the user if it contained all of the query words when it was last read by the system.

3. The system must maintain an accurate representation of file contents for any file that has not changed since the *model age*.

4. The system must have a *model age* no older than 1 minute when it begins to respond to a user query.

5. The system must finish responding to the user within 15 seconds of the user providing the *query words*.

6. The system must finish responding to the user within 5 seconds of the user providing the *query words* on average.

7. The system must support at least 1000 directories in the configuration files.

8. The system must support at least 100,000 files in the configured directories.

9. The system must support at least 1 GB of total file contents.

10. The system must support the US-ASCII character set for the files to be searched and user input.

11. The system must reject user input that it cannot securely process.

12. The system must support *queries* containing up to at least 80 characters.

13. The system must be designed with at least 4 separate components.

14. The system must achieve a reliability of 0.995 for an interval of 10 minutes.

15. The system must support at least 4 simultaneous users.

4 Analysis

Prepare an analysis of the reliability of your system to demonstrate that the design is expected to achieve the require reliability. You will need to estimate the reliability of each component by taking into account the number of functions, their complexity (lines of code) and how often they will be called in 10 minutes given the non-functional requirements. Once you have the reliability of each component, you can calculate the overall reliability. You may need to be creative to achieve the required reliability.

5 Submission

Please submit your answer on Moodle and bring a printed copy to class. You will be reviewing the designs of another group during class.

Your design should include, at a minimum:

1. (20 pts) An executive summary that describes in well-written paragraphs the problem you are solving, a high-level description of your solution and mentions the reliability that your solution achieves.

2. (20 pts) A diagram showing the overall architecture of your system.

3. (40 pts) Design detail for each component in your system. A minimum of 4 components are required.

4. (20 pts) An analysis of the reliability of your system, with explanation and calculations.