**File Log Processing  
  
Problem Statement :**

implement a file log processor app in java that is capable of prepending line numbers for every line in a file present in the target directory as shown below:

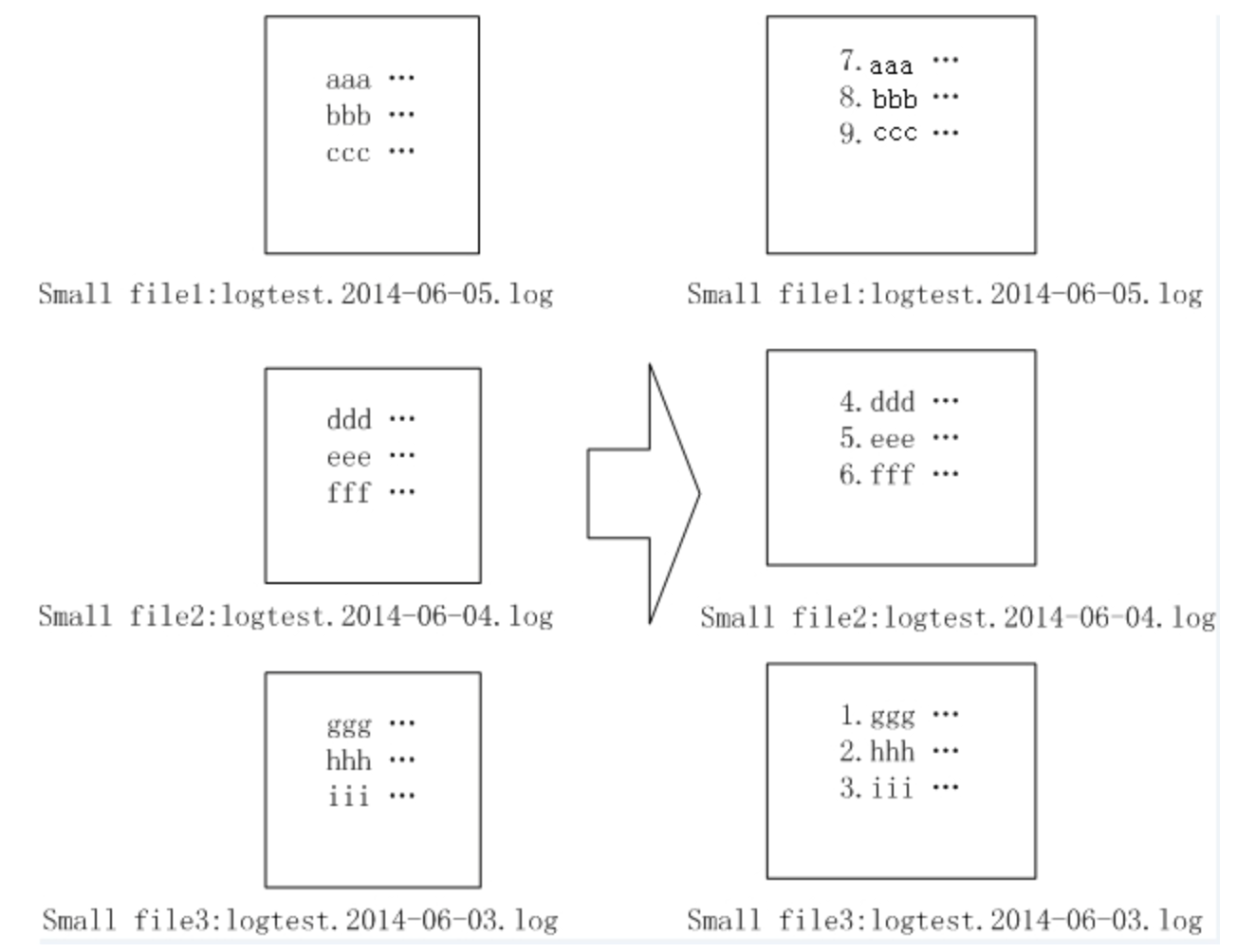


Figure : file log processor requirement

**Implementation:**

Below are certain assumptions that I had considered.

1. The total number of files are roughly 1000000.
2. The naming convention for the small files follows the format “logtest.yyyy-mm-dd.log”
3. The files that do not follow above naming convention are not processed.
4. The files are ordered according to the timestamp (oldest to latest).
5. The file name is considered to order the files according to timestamp not the last modified date. (as it was not clear in the problem what to use)

The problem is solved in two phases the first phase will generate a lookup list for all the files in the directory. This lookup list will have the file index and the starting line number for this file.

The second phase will utilize this lookup values to prepend the numbers for every line in a valid files.

**Design:**

1. Class Diagram

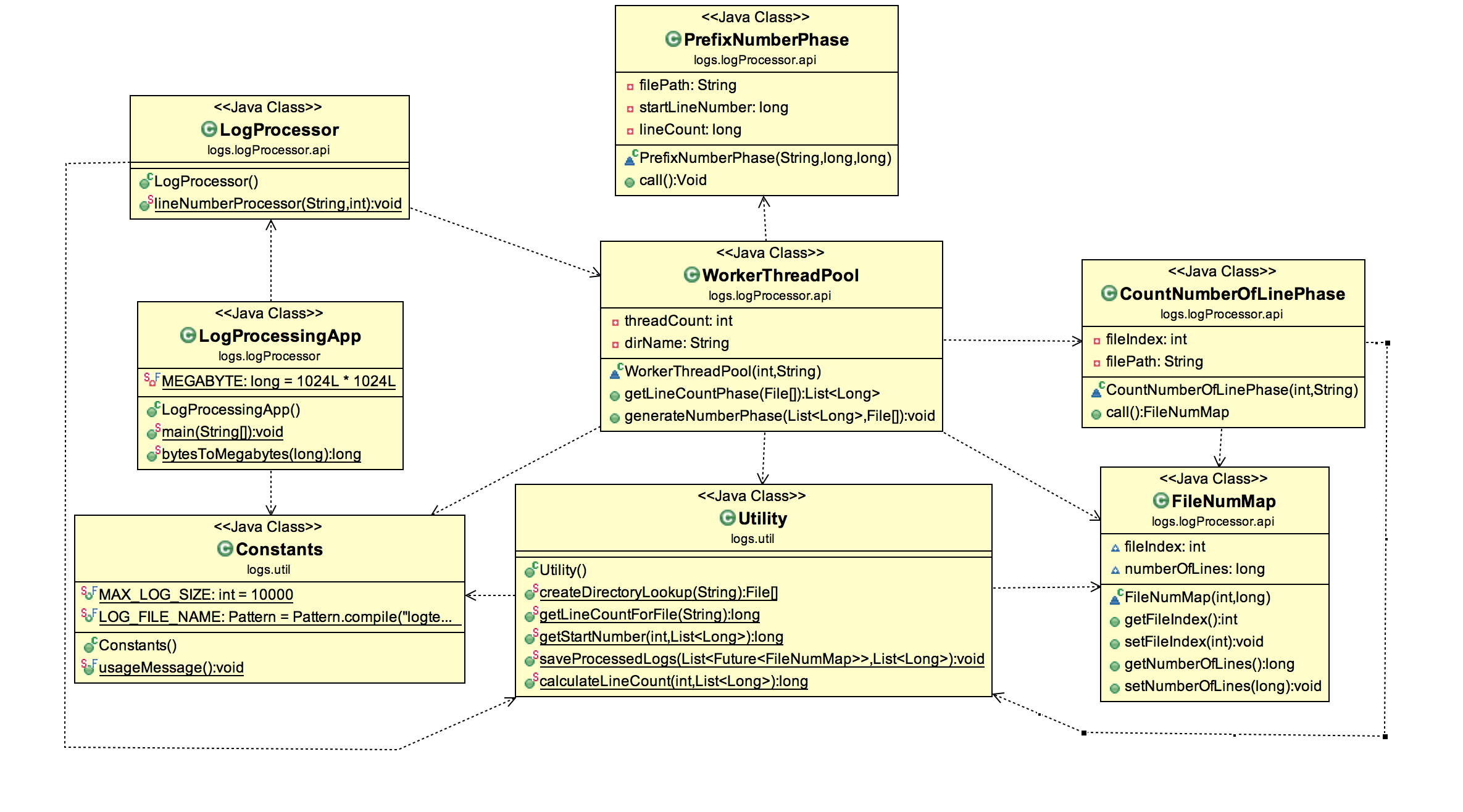
****

Figure Class Diagram

1. Approach/Algorithm
   1. The entry point for the app is LogprocessorApp, the main function gets two arguments as an input to initiate the process.
      1. Max number of threads that can be configured for processing
      2. The target directory path where all the log files are stored.
   2. Once the required information is received, task is delegated to Log Processor, which creates a worker threads (Thread Pool) which are responsible for performing the task.
   3. The following task are carried out by the pool of threads
      1. Calculating the total number of lines in the files, which helps in creating a list lookup of file index and the starting line number for each valid file.
      2. Starts prepending the numbers for every line in the file with the help of list lookup generated in first phase.
2. Application uses **MemoryMapped** concept to improve the write performance.

**Steps to build & run the project:**

The project requires jre 1.7 and above

1. Build the project using “mvn clean install” command and then execute the below command:

java -classpath < fileLogprocessor-0.0.1-SNAPSHOT.jar > logs.logProcessor.LogProcessingApp --max-threads=<number of threads> --dir=<directory path for the log files>  
  
eg : java -classpath fileLogprocessor-0.0.1-SNAPSHOT.jar logs.logProcessor.LogProcessingApp --max-thread=4 --dir=/Users/Jhon/logs/

1. Import the Project in eclipse and run the project as “ run configuration”, select the main class as LogProcessorApp and in the arguments tab provide below details

--max-threads=<number of threads> --dir=<directory path for the log files>

eg: --max-thread=4 --dir=/Users/Jhon/logs/