

Summary of Design for Log Query tool

Khushboo Tekchandani

Task

To design a simulator that generates a log file for monitoring CPU usage for 1000 servers, each having two CPUs.

To design a command line tool which queries the log files to provide the CPU usage for a specific CPU between a given time.

Simulator

The directory structure for the logs is designed in a way to be able to uniquely log information for each CPU on a given date. To do this, the logs are classified according to the 'Year' and then according to 'Month' and then 'Day'. So, the simulator takes a date as input in the form of YYYY-MM-DD and creates a directory with the name 'YYYY'. Another directory with name 'MM' is created inside this directory. And inside the month directory, a text file with name 'DD.txt' is created to log the CPU usage.

Query Tool

The query tool allows the user to query the log monitoring system. It supports two commands

1. QUERY <serverIP> <CPUID> <start date> <start time> <end date> <end time>
2. EXIT

The query tool initializes to create a map of the entire information from the log files. The tool would require few minutes to do this, however, it would enable the tool to provide a quick response to user queries.

The said index consists of a hashmap and a linkedlist of tuples (or pairs) associated with each key in the hashmap. The key is essentially a combination of the IP address of the server and the CPU ID of the CPU, which uniquely identifies each CPU. And the value is a linked list of tuples of timestamp and CPU usage which comprises of the complete information about each CPU.

Data Structures and Classes

1. struct date: Stores date components in int format
2. struct utime: Stores human readable time in int format
3. class initialMapping: Used for storing the mapping that the tool creates upon creation
4. class query: Used for storing complete information about the query including, server IP, CPU ID, start time, start date, end date and end time.

Timeline

Three days were required to complete this task including system design, implementation and documentation. Following is the division of work during the timeline:

Day1: System design, simulator implementation and backbone for query tool

Day2: Query tool implementation

Day3: Query tool implementation, testing and documentation