

题目描述

7-1 httpd日志分析 分数 10

全屏浏览题目

作者 翁恺 单位 浙江大学

Web servers typically maintain log files of client accesses to the web pages that they store. Given suitable tools, these logs enable web service managers to extract and analyze useful information such as:

- which are the most popular pages they provide;
- which ip took the most pages from the site;
- whether other sites appear to have broken links to this site’s pages;
- how much data is being delivered to clients;
- the busiest periods over the course of a day, or week, or month.

Such information might help managers to determine, for instance, whether they need to upgrade to more powerful server machines, or when the quietest periods are in order to schedule maintenance activities.

你的程序会读入一个日志文件，日志文件的每一行可能是以下的形式：

```
10.180.17.246 - - [24/Dec/2017:09:31:35 +0800] "GET /apple-touch-icon.png HTTP/1.1" 404 509 "-" "Safari/13604.4.7.1.3
CFNetwork/893.13.1 Darwin/17.3.0 (x86_64)"
10.180.17.246 - - [24/Dec/2017:09:31:44 +0800] "POST /login.php HTTP/1.1" 302 452 "http://fm.zju.edu.cn/" "Mozilla/5.0 (Macintosh;
Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
10.180.17.246 - - [24/Dec/2017:09:32:17 +0800] "GET /showCourse.php?id=57 HTTP/1.1" 200 1894 "http://fm.zju.edu.cn/index.php"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
```

程序要从中找出浏览器所访问的网址，统计所有的网址中，访问量最大的网址，输出日志中的访问量和网址。

当网址带有GET的请求时要去掉请求数据来统计。当存在多个网址的访问量相同时，按照字母顺序输出。

输入格式:

纯文本日志。

输出格式:

每行输出一个网址，首先是它的次数，然后是冒号，再跟上网址。

输入样例:

```
10.180.17.246 - - [24/Dec/2017:09:31:44 +0800] "POST /login.php HTTP/1.1" 302 452 "http://fm.zju.edu.cn/" "Mozilla/5.0 (Macintosh;
Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
10.180.17.246 - - [24/Dec/2017:09:31:44 +0800] "GET /index.php HTTP/1.1" 200 1400 "http://fm.zju.edu.cn/" "Mozilla/5.0 (Macintosh;
Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
10.180.17.246 - - [24/Dec/2017:09:32:17 +0800] "GET /showCourse.php?id=57 HTTP/1.1" 200 1894 "http://fm.zju.edu.cn/index.php"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
10.180.17.246 - - [24/Dec/2017:09:33:14 +0800] "GET /showCourse.php?id=126 HTTP/1.1" 200 3765 "http://fm.zju.edu.cn/showCourse.php?
id=57" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_2) AppleWebKit/604.4.7 (KHTML, like Gecko) Version/11.0.2 Safari/604.4.7"
```

输出样例:

2:/showCourse.php

代码长度限制	16 KB
时间限制	1200 ms
内存限制	64 MB

Java (javac)

?

🕒

⚙️

?

🔍

📄

1

//package org.example;

2