# 4. 海量日志数据，提取出某日访问百度次数最多的那个 IP。

方案 1：首先是这一天，并且是访问百度的日志中的 IP 取出来，逐个写入到 一个大文件中。注意到 IP 是 32 位的，最多有 个 IP。同样可以采用映射的 方法，比如模 1000，把整个大文件映射为 1000 个小文件，再找出每个小文

中出现频率最大的 IP（可以采用 hash\_map 进行频率统计，然后再找出频率 最大的几个）及相应的频率。然后再在这 1000 个最大的 IP 中，找出那个频 率最大的 IP，即为所求。