百度文学：

笔试：

1.有A、B两份1亿条的url，你有一台4G可用内存的PC，请设计一种算法，把这A、B两份数据集中重复的url挑选出来。

2.根据时间戳计算ARPU值

3.关于一些shell脚本的处理

4.谈谈你对BI的理解

... ...(给忘了，总共是7道题目)

面试：

1.自我介绍，讲述自己的经历，做过哪些项目等

2.做过哪些项目，用过哪些算法，原理是什么，有哪些关键点

3.现场提一些算法方面的问题，需要你去解决

    (1) 用Hadoop实现全排序，说一下思路和Key、Value的表述（考察对Hadoop的掌握）

    (2) 给你一棵任意的二叉树，求出树中的最长路径，并输出（考察对回归算法的理解）

4.和人力面谈，一些软技能方面的东西，最后谈薪资

    (1) 历次离职的原因解释

    (2) 个人兴趣爱好之类的问题

    (3) 谈薪资

百度LBS部门：