上海金邱

Hadoop和spark有什么区别？

Hadoop是一个由Apache基金会所开发的分布式系统基础架构。 Hadoop实现了一个分布式文件系统HDFS。HDFS有高容错性的特点，并且设计用来部署在低廉的硬件上；而且它提供高吞吐量来访问应用程序的数据，适合那些有着超大数据集的应用程序。Hadoop的框架最核心的设计就是：HDFS和MapReduce。HDFS为海量的数据提供了存储，而MapReduce则为海量的数据提供了计算

Spark 是专为大规模数据处理而设计的快速通用的计算引擎。Spark拥有Hadoop MapReduce所具有的优点，Spark在Job中间输出结果可以保存在内存中，从而不再需要读写HDFS，因此Spark性能以及运算速度高于MapReduce。

汉克时代 一轮面试 公司地点：珠海

银行面试 时间长度：25分钟

面试问题：

1. 数据仓库的分层
2. 数据建模
3. 维度表和事实表
4. 项目介绍
5. 存储过程
6. 项目流程管理

数据建模不了解的不好

公司工作内容是银行的指标加工 后台的指标加工 传统的数仓

Kettle用过吗？

你们同步数据的周期是什么？

你们调度工具用的是什么？

大概什么时候过来？

为什么从上家走？

你对银行业务了解吗？