# 电信客服项目

## 背景说明：

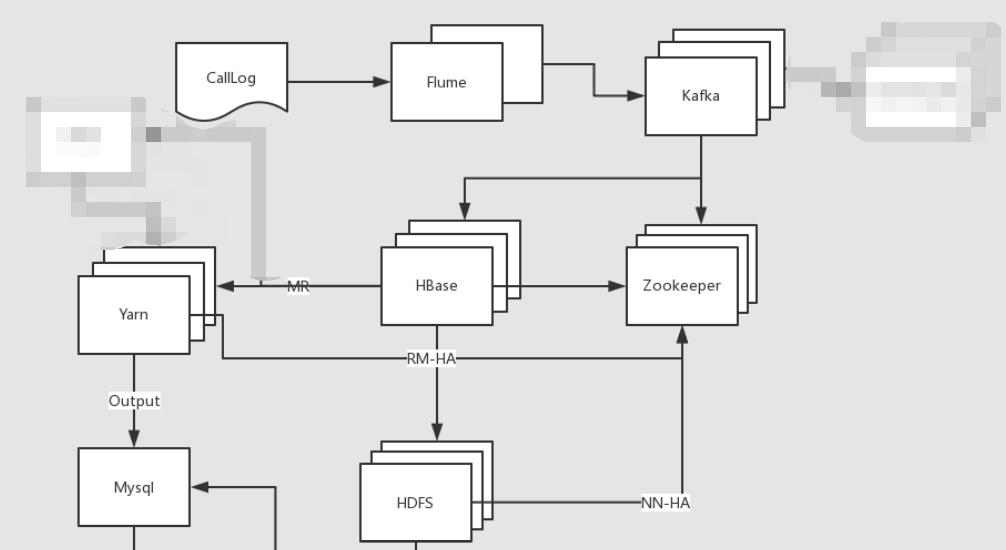
通信运营商每时每刻会产生大量的通信数据，例如通话记录，短信记录等信息。数据量如此巨大，需要定时定期的对已有数据进行离线的分析处理。例如，当日话单，**月度话单，季度话单，年度话单**，通话详情，通话记录等等。当前该项目需求是：

每人每天的通话次数和时长

每人每个月的通话次数和时长

每人每年的通话次数和时长

项目架构



模拟数据模块

随机主叫和被叫

随机电话建立的时间

随机通话的时长

将通话记录**写入文件**，

数据采集：

Flume：

Source：exec

Channel：memoryChannel

Sink：kafkaSink

Flume就是用来往HDFS里面放置数据的，所以flume对HDFS的支持非常好，而kafka向HDFS的支持没有flume好。

Kafka：

接收flume传过来的信息

数据存储

HBase做为消费者，消费kafka中的数据，将数据存储成结构化的数据，

预分区，防止数据热点。Redis存储的数据恰恰是热点数据（大家都使用比较多的，比如说一个APP的首页），作为数据库的前置。

数据分析：

从HBase里读取数据，使用MapReduce程序来分析；

并且需要自定义output将数据写入到MySQL

Web的展示：

SSM：springMVC+spring+Mybatis 为什么是这个叫法？这里表示了数据的流向的问题，SpringMVC作为服务器端的框架来接收页面的数据， 然后在将数据传递给 🡪service 🡪dao

Service和dao的组合和功能的扩展依赖于spring框架。最后和数据的交互交给Mybatis。

定义类的时候的package的标准：公司域名反写+项目名+模块+程序类型名。