

Airbnb架构要点分享

谢丽 架构师联盟 2015-09-23

Airbnb成立于2008年8月，总部位于加利福尼亚州旧金山市。Airbnb是一个值得信赖的社区型市场，在这里人们可以通过网站、手机或平板电脑发布、发掘和预订世界各地的独特房源，其业务已经覆盖190个国家的40000座城市。目前，Airbnb已经完成了7轮融资，共募集了23亿美元的资金，估值达250亿美元。Mike Curtis是Airbnb工程部门的副总裁。近日，The Platform编辑Timothy Prickett Morgan就Airbnb的架构要点对他进行了采访报道。

Airbnb自成立之初就部署在AWS上，Curtis认为：

这很重要，我们的工程师可以尽可能地专注于让我们与众不同的东西上，而不用运行一个庞大的基础设施。

目前，Airbnb已经使用了大约5000个AWS EC2实例，其中大约1500个实例用于部署其应用程序中面向Web的部分，其余的3500个实例用于各种分析和机器学习算法。而且，随着Airbnb的发展，相比于业务处理所需的计算能力，分析和机器学习的计算能力需求增长更快。这是因为，随着Airbnb的用户越来越多，其挑战不是找到一个可供出租的地方，而是找到一个合适的地方，并且让房主和房客都满意。按照Curtis的说法：

我们在工程上所做的一切工作都是为了在人与人之间创建良好的匹配关系。每个旅行者或房主的情况都不相同。人们都有不同的偏好，希望获得的旅行体验就有所不同。因此，我们在工程上所做的大量工作都是关于，如何为了真实世界的线下体验将合适的人匹配在一起。这是我们所做的一切工作的组成部分，其中包括机器学习、搜索排序、欺诈检测.....

关键就是，根据房客和房主的偏好（个人资料及在Airbnb上的交易和搜索历史），使用搜索引擎和机器学习，找出最好的租房/出租选择。而且，选项不能太多，否则用户会很难决定。因此，Airbnb需要借助机器学习强化搜索，为用户（房客和房主）提供5到10个选项。这缩短了交易时间，降低了Airbnb系统的负载，而用户也更满意。

同许多公司一样，Airbnb开始时也使用了若干开源软件，但进行了大量的修改和扩展，以便创建一个机器学习强化的搜索引擎。其核心索引技术使用了Lucene，但只用了反向文本部分。而所有的排序和机器学习功能则是他们自己编写的。他们的第一个机器学习实验就将预订率提升了4%。Airbnb的动态定价也是由机器学习算法所支撑。据Curtis说，当房主按照Airbnb推荐的价格上下浮动不超过5%定价，其出租成功率提升了4倍。

为了使机器学习更容易嵌入应用程序，Airbnb创建了一个名为Aerosolve的工具（GitHub页面）。该工具有许多功能，但最重要的是让数据科学家可以了解机器学习算法（房屋推荐或定价建议）内部正在执行什么操作，然后运用直觉获得更好的结果。

驱动Airbnb站点和分析的核心数据平台是Hadoop，所有东西都存储在Hadoop分布式文件系统中。之前，他们使用Amazon Elastic MapReduce服务，并在AWS EC2计算实例上部署了自己的Hadoop，但去年，他们已经迁移到Cloudera的企业级Hadoop上。此外，Airbnb使用S3存储Web站点的图片和Hadoop集群的备份数据。

在HDFS文件之上，Airbnb使用由Facebook创建并开源的Hive和Presto创建了一个数据仓库。对于长时间运行的查询，他们使用MapReduce。Hive不支持子查询，使用MapReduce可以获得同子查询类似的结果。Curtis指出，Presto非常强大，它兼容SQL，使得数据查询非常容易，而且速度很快，虽然功能不如Hadoop MapReduce或Hive多，但在日常分析中非常有用。他们还创建了一个名为Airpal的工具（GitHub页面），用于设计SQL查询，并发送到数据仓库的Presto层。

Curtis告诉The Platform：

实际上，去年年底，我们做了一次很大的集群迁移，将所有的数据基础设施分成了两个单独的镜像集群：一个用于运行所有关键的业务任务.....另一个用于即时查询.....

这两个Hadoop集群使用Kafka保持同步。Airbnb创建有自己的工作流和ETL工具Airflow（GitHub页面）。该工具已经钩入HDFS、Hive、Presto、S3、MySQL和Postgres。此外，Airflow接收来自Airbnb站点的非结构化数据流，增加结构后转储到HDFS上的Hive表中。曾经很长一段时间，他们使用成百上千的CRON任务来做这项工作，那很容易出问题。而Airflow提供了一种简洁、可编程的系统，使他们能够查看、编辑、监控和调试ETL系统。

为了管理AWS上的集群，他们使用Chef进行配置管理，并自己开发了很多工具。虽然Mesos集群控制器已经被前Twitter和Airbnb工程师商业化，但他们现在使用Mesos并不多，因为他们觉得：

本质上讲，Mesos是一个抽象层。其设计初衷是让很多节点看起来像一个节点。我们发现，尤其是在Mesos上运行某些新东西时，这个抽象层实际上会隐藏一些东西，增加了我们调试的难度。

最后，Curtis指出，虽然他们无意运行自己的数据中心，但IT团队每隔六个月就会对计算和存储能力及成本进行分析，然后与创建并运营本地数据中心进行对比。他们发现，租用AWS的服务可以节省20%到30%的费用。而按Curtis的说法，实际节省的费用可能更多，而且，在数据中心投入过多的精力就无法专注于业务了。

阅读原文