

Data Engineer

Course info session

细说大数据工程师公开课

2017 Copyright @ Data Application Lab
2017-3-25

Agenda

- Data Application Lab 介绍
 - Data Engineer课程 4/1 开课. 简介
 - Jason介绍大数据行业, 海量数据处理基本流程
 - Case Study
 - Max介绍大数据构架系统
 - 近期Data Engineer学员, 分享拿到Offer的经历
 - 上期同学拿到offer的感想
 - 同时Q&A 提问, 多位名师在线解答疑惑.
 - 如何报名
-



Data Application Lab

Data Application Lab

Our graduates work for great companies

[24]7



BOSCH



Abbott
A Promise for Life

ACI
payment systems



Microsoft

Bloomberg



PlanGrid



amazon

Booking.com

DISCOVER

OPERA
MEDIWORKS

Qumulo

Quora



ebay

Fanatics

GoodData

PROMETHEUS
Therapeutics & Diagnostics
A Nestlé Health Science Company

SONY



Symantec

Google

HUAWEI



irdeto



TWO SIGMA

VISA

yelp



LinkedIn

RE/MAX
PACIFIC REALTY

MONSANTO

tableau

zest finance

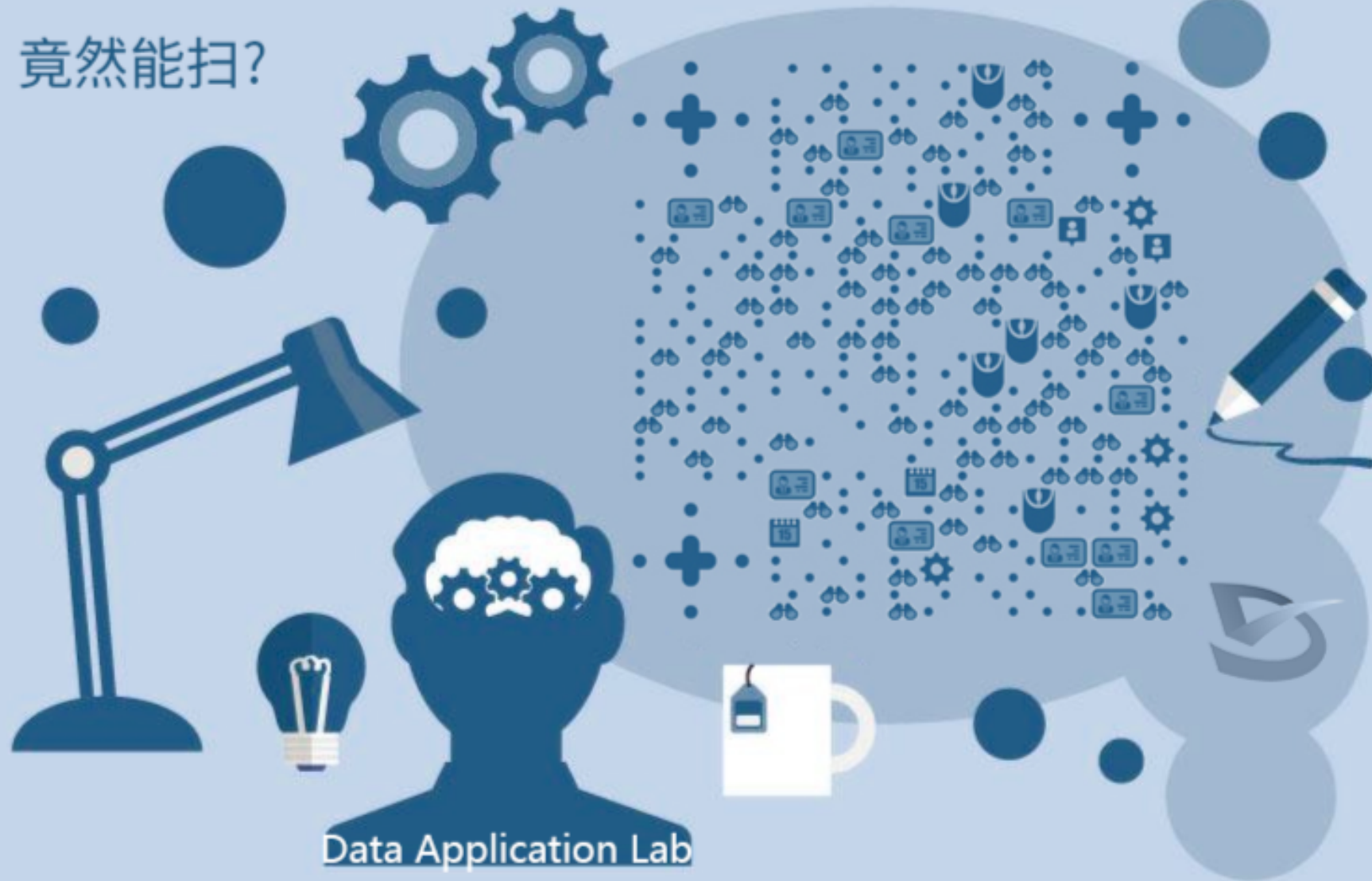
Zillow



FYI:

- <http://DataAppLab.com>
 - DataScienceAssociation
 - <http://DataScienceAssociations.org>
 - Facebook: <https://www.facebook.com/dataapplab/>
 - LinkedIn: <https://www.linkedin.com/company/7956968>
 - 微信: 大数据应用
-

竟然能扫?



Data Application Lab

北美最权威的-数据应用学院

DataAppLab.com



Data Application Lab

- [Home](#)
[About Us](#)
[Contact Us](#)
[FAQ](#)
[Privacy Policy](#)

[f](#)
[in](#)
[Events](#)
[Wechat](#)
[My account](#)
[Q](#)



Data Application Lab

大数据工程 | Data Application Lab

Education

Consulting

About us

Im(国际项目)

大数据工程师

HOME | COURSE



Application Procedure

FINISH APPLICATION

WAIT FOR CONTACT

PAYMENT BY INSTRUCTION

项目介绍:

一、紧跟行业前沿的课程设计

随着越来越多的公司将大数据技术应用到产品中,有集软件开发经验的工程师成为Software Engineer市场的宠儿。起独持的不可替代性确保了 一个稳定的高收入IT职业。我们收集和总结了各大TOP公司,包括FLAG在内的关于Big Data Engineer的招聘贴,从Entry level到 Management level的相关Job Description,制定了了最适合进入这个行业的全面培训求职课程。希望可以帮助具有工程专业、计算机专业背景的学生在职场中突破重围,找到一份有技能点增值的高薪职位。

关键词: Hadoop, Big Data, ETL, Data Pipeline, Machine Learning, Artificial Intelligence

 - 我们的讲师均来自北美顶尖科技企业,且具有3-8年大数据经验。他们凭借自己多年工作经验言传身教,特别是分布式系统调优,Java语言在分布式系统中的应用和最新实时数统系统的开发。
 - 授课内容乃当前大数据的最新技术与技巧, HDFS, MapReduce, Hive, Pig, Sqoop, Storm, Kafka, Spark等工具使用技巧和系统调优。都是学校里学不到, 网络课程不会讲, 而工作中又需要的重要技能; 每周8周授课, 每周6天2小时小班授课, 周日2小时实战课, 每周一个业界应用Case Study; 每周work load 10-15小时; 1月期渐速若干小时, 更多技能训练
 - 8周授课, 每周6天2小时小班授课, 周日2小时实战课, 每周一个中文业界应用Case Study; 每周work load 10-15小时;
 - 课后TA全程辅导, 让你在遇到问题第一时间得到帮助, 解决孤立无援的自学窘境
 - 算法辅导, 提升刷题能力
 - 同时熟悉业界最常用的大数据平台: Hortonworks (我们自己的实体服务器群), Cloudera (课上练习), Amazon Web Service (AWS, 项目实战)

Data Engineer 课程

•Lambda Architecture

- 随着课程的进行, 逐步了解和实践大数据行业最常用的Lambda构架
- 根据事件, 依据数据的时效分层次处理(参见下图)
- 实战Data Pipeline, 处理互联网Review平台中客户评价的数据

•Recommendation System

- 结合我们数据科学实际项目, 制作用户游戏推荐系统
- 实践和体验在Amazon云平台(AWS)中各项基础工具的使用方法(EC2, S3等)
- 从根基搭建基于Spark的大规模数据处理系统

•Financial Technology (FinTech) Online System

- 进阶演练Lambda构架,制作金融业务应用
- 练习业务接口, API, 实现数据整合
- 基于Storm和Kafka, 练习实时数据流处理
- 体会大数据系统在实际金融项目中的应用

•Healthcare 数据分析整理

- 基于时下流行话题, 展示和演练综合Data Analytics 技巧
 - 体会不同Domain Knowledge下的数据整理 (Integrating and normalizing diverse datasets)
 - 结合大数据平台下的分布式数据库构架, 操作复杂数据的存储和读取
-



Data Application Lab

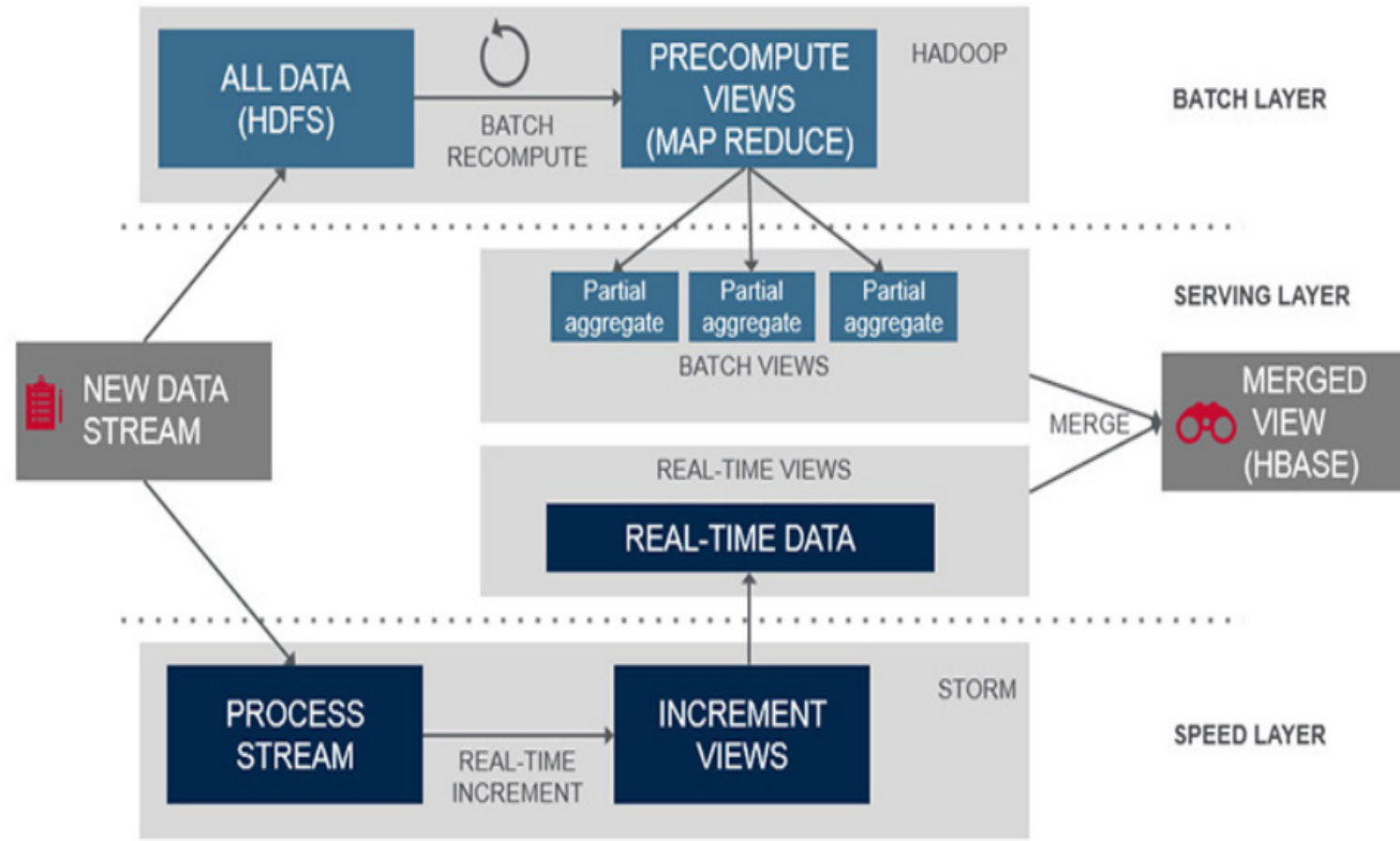
BIG DATA DATA SCIENCE

Data Engineer vs Software Engineer

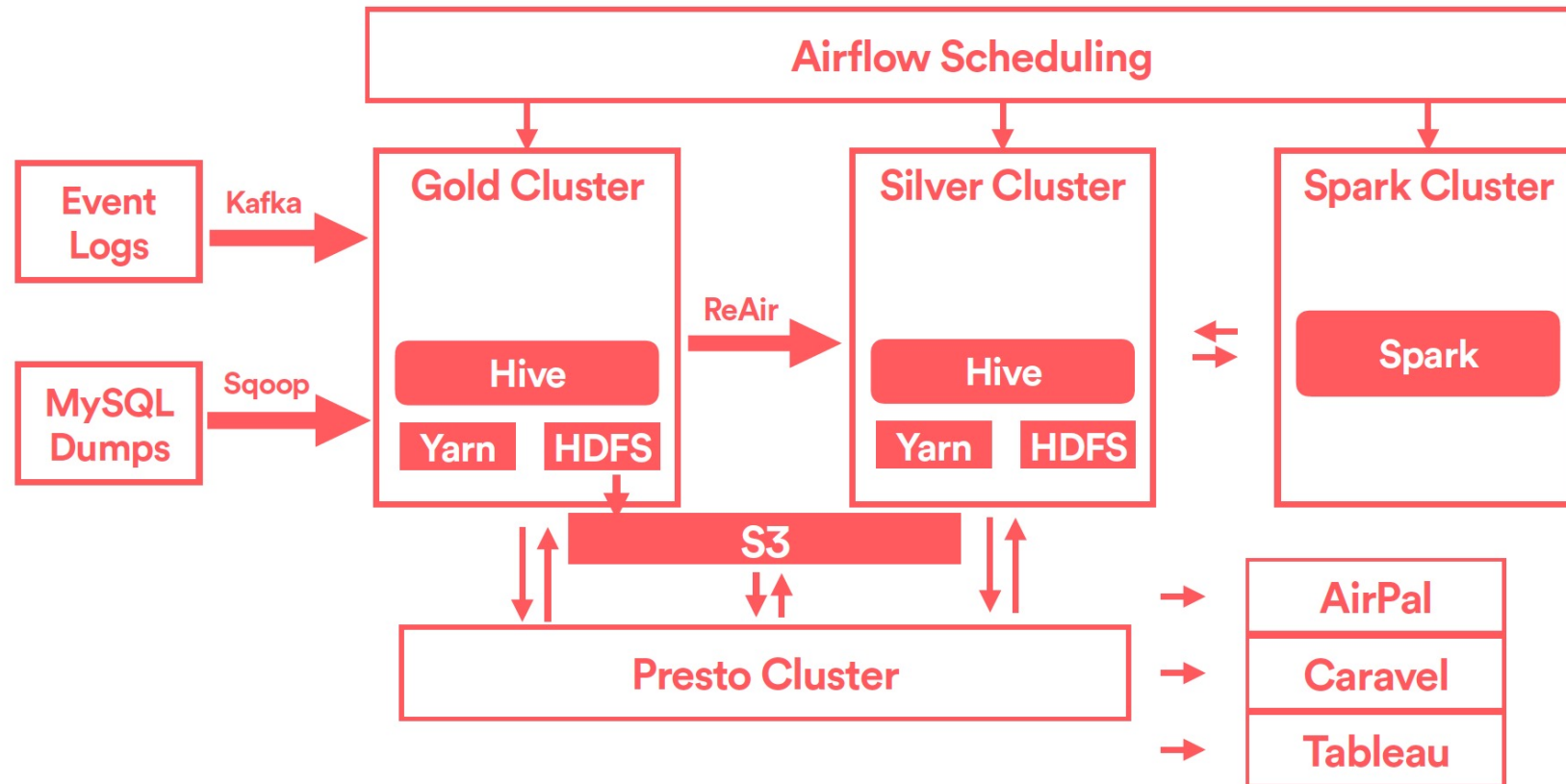
- Advanced data structures
 - Distributed computing
 - Concurrent programming
 - Knowledge of new & emerging tools: Hadoop, Spark, Kafka, Hive, etc.
 - Building ETL/data pipelines
- Frontend and backend development
 - Web apps
 - Mobile apps
 - Operating system development
 - Software design
-

Big Data System Architect

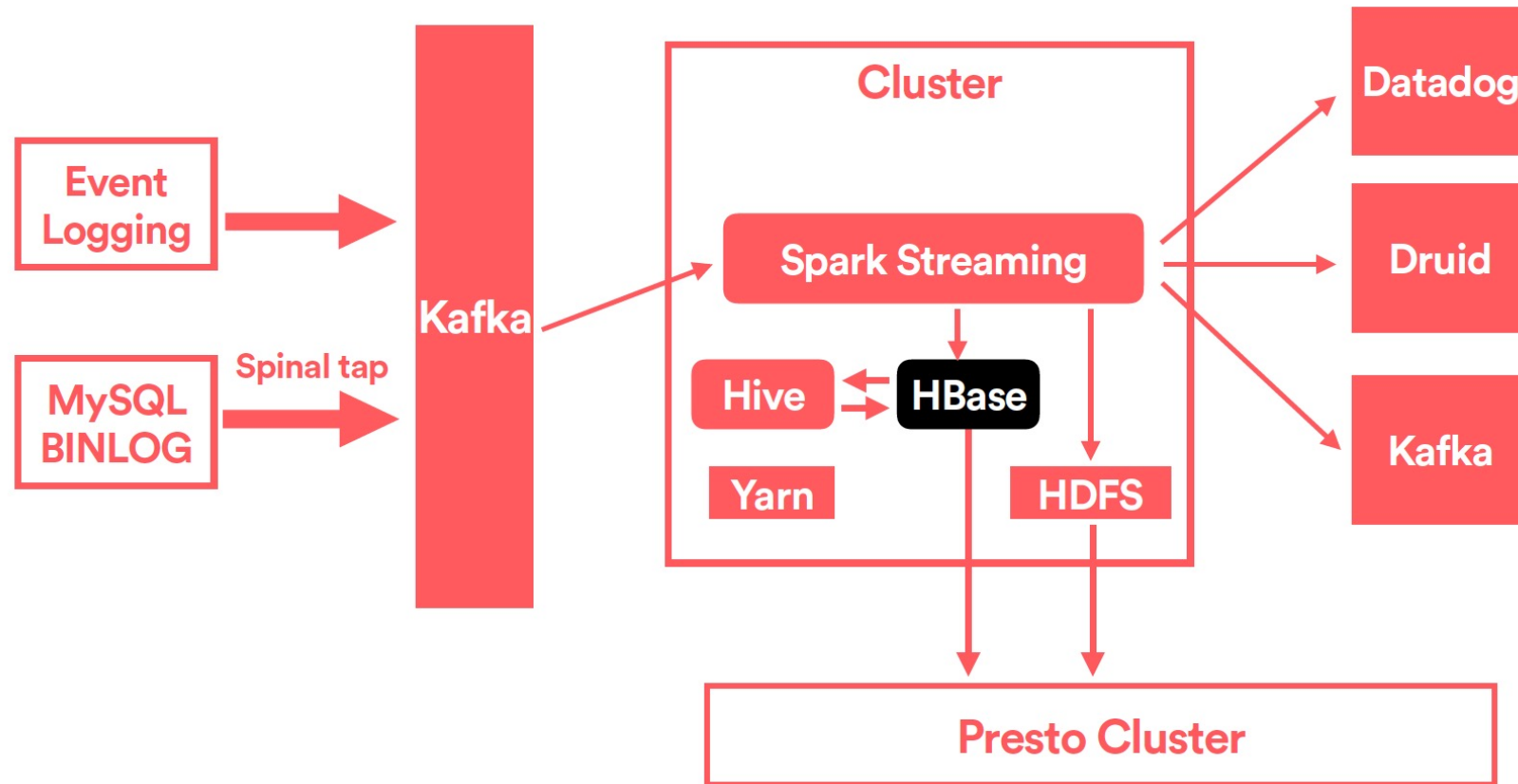
Lambda Architecture



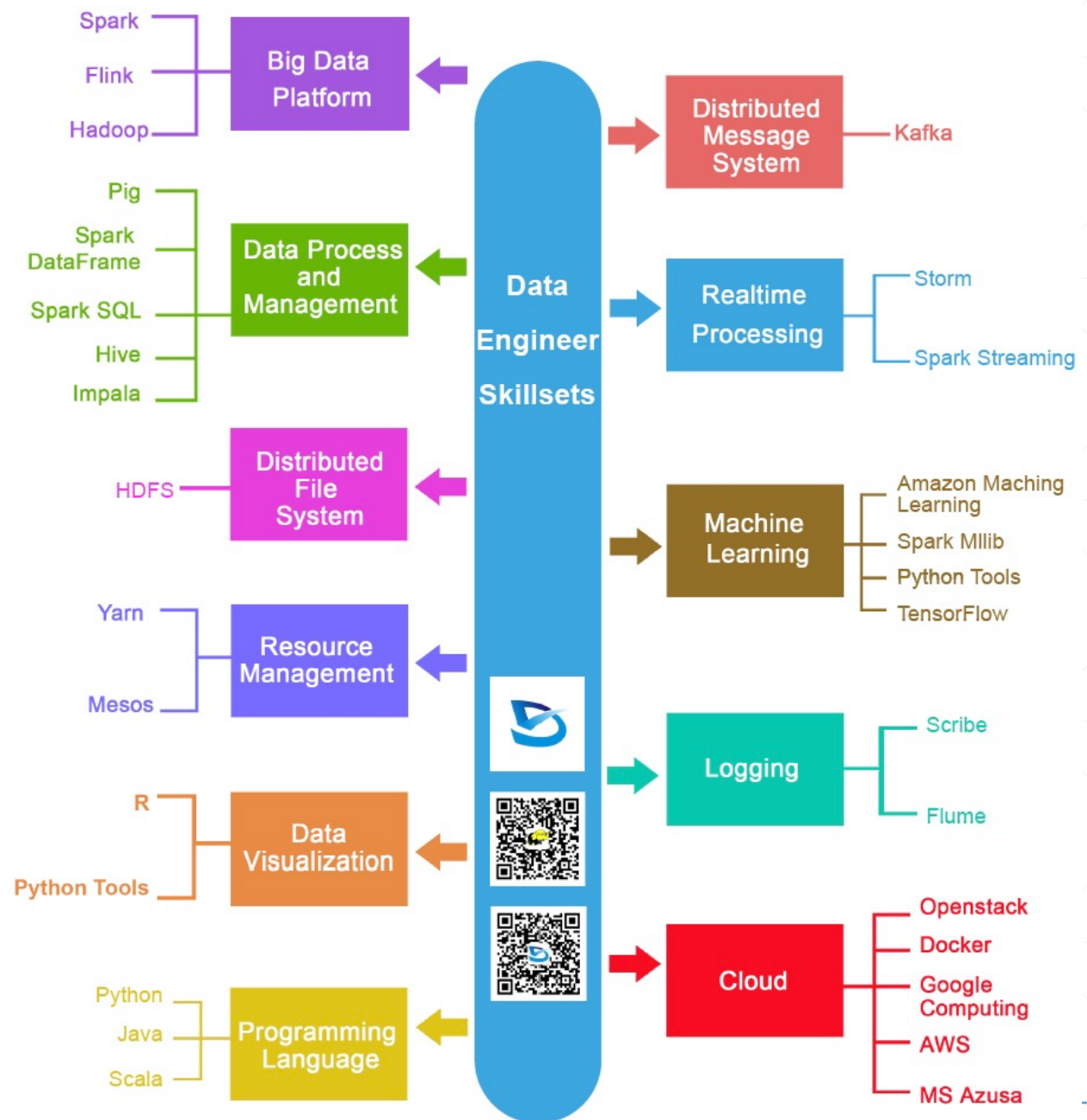
Batch process @Airbnb



Real-time process @Airbnb



Data Engineer



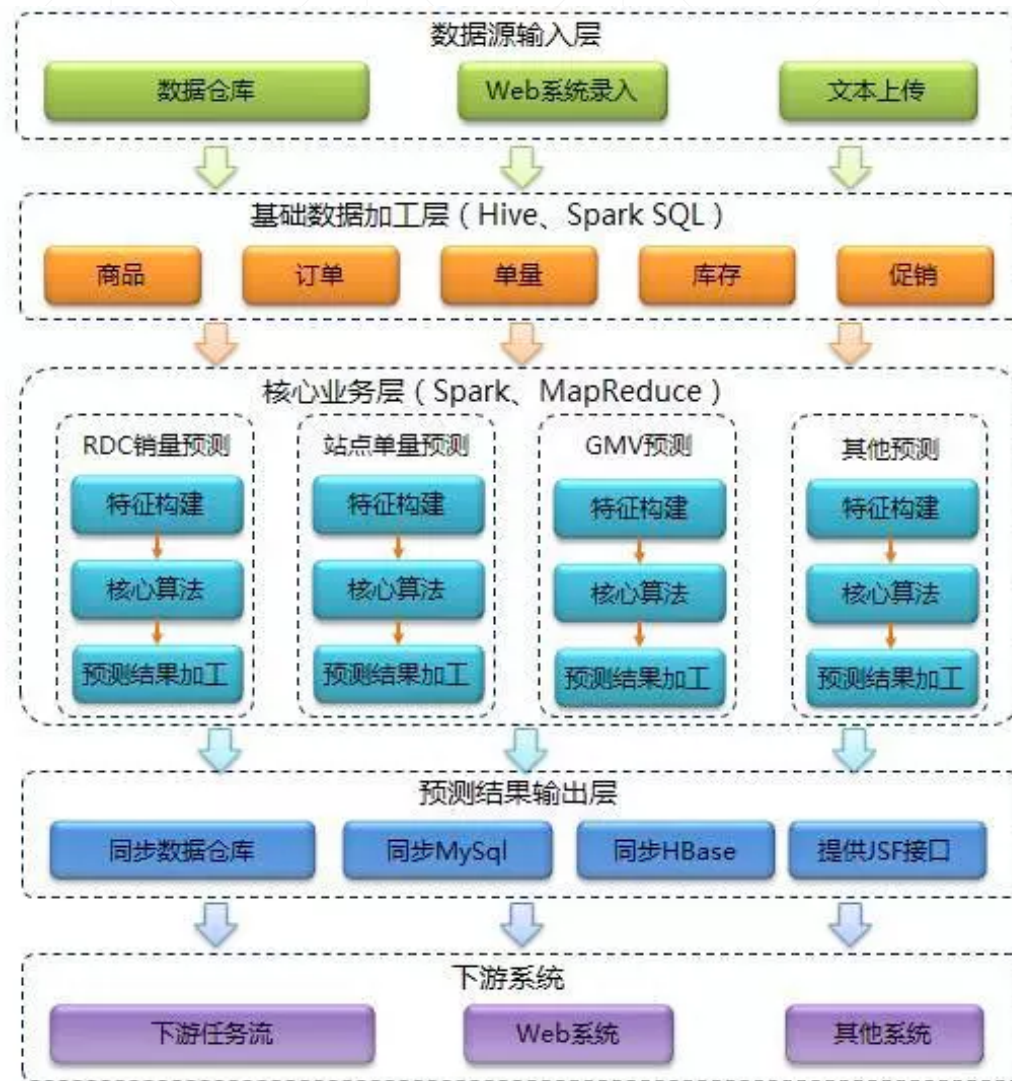
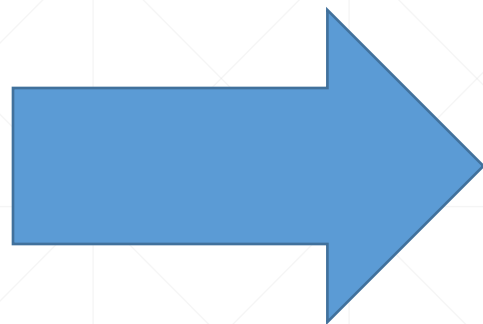
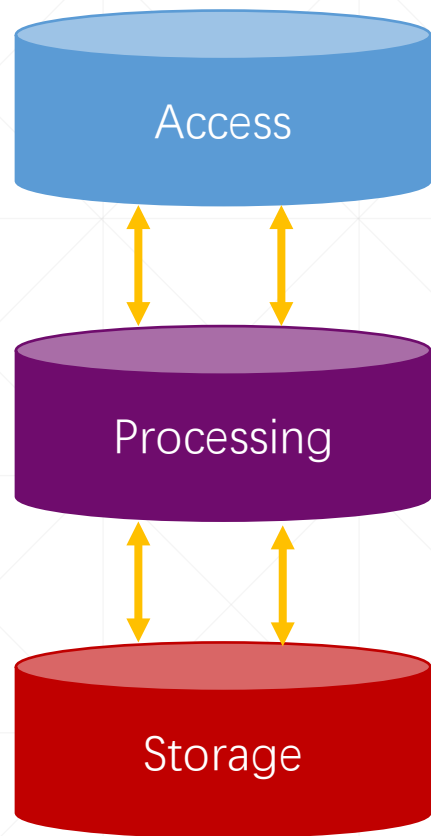


Max

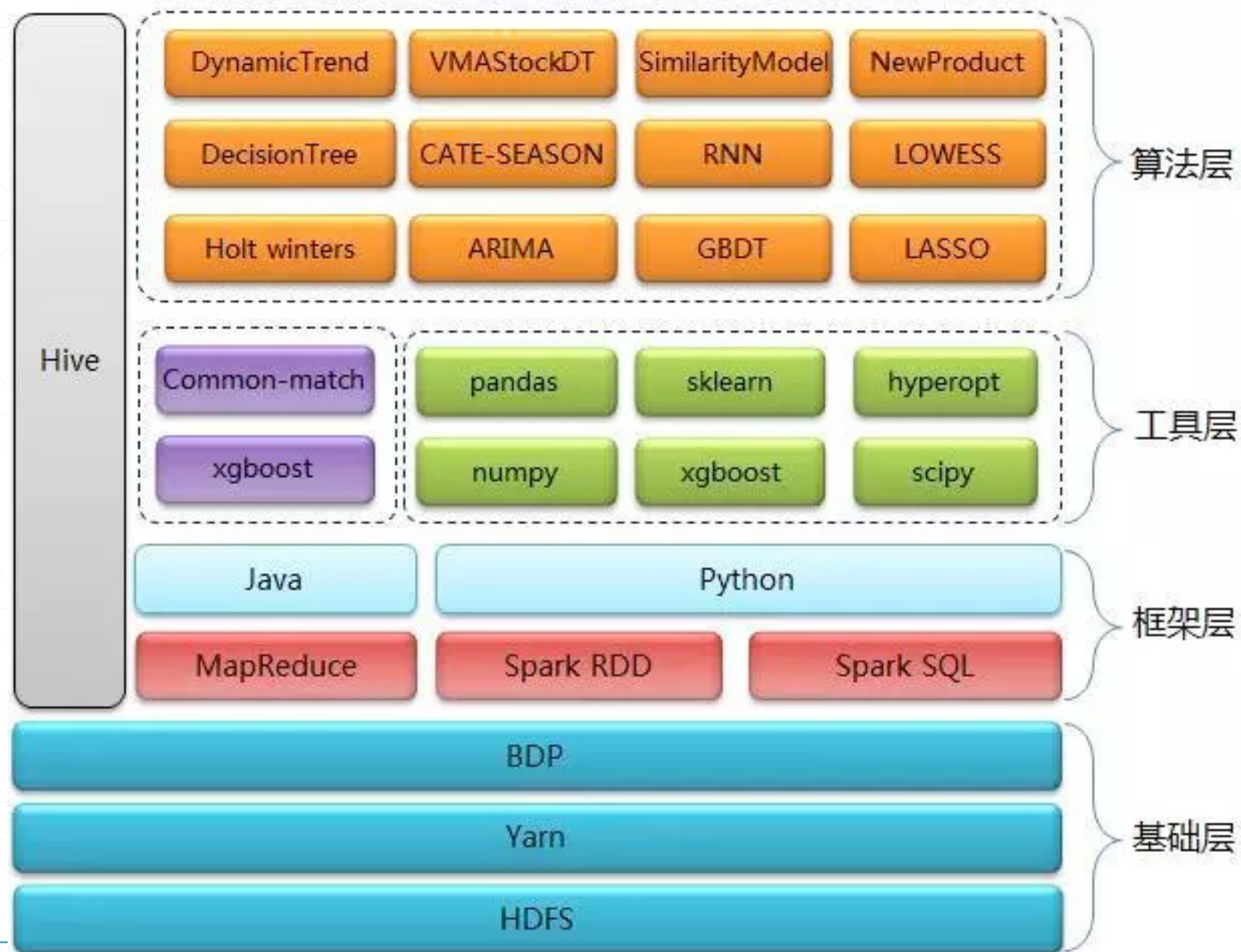
Case Study



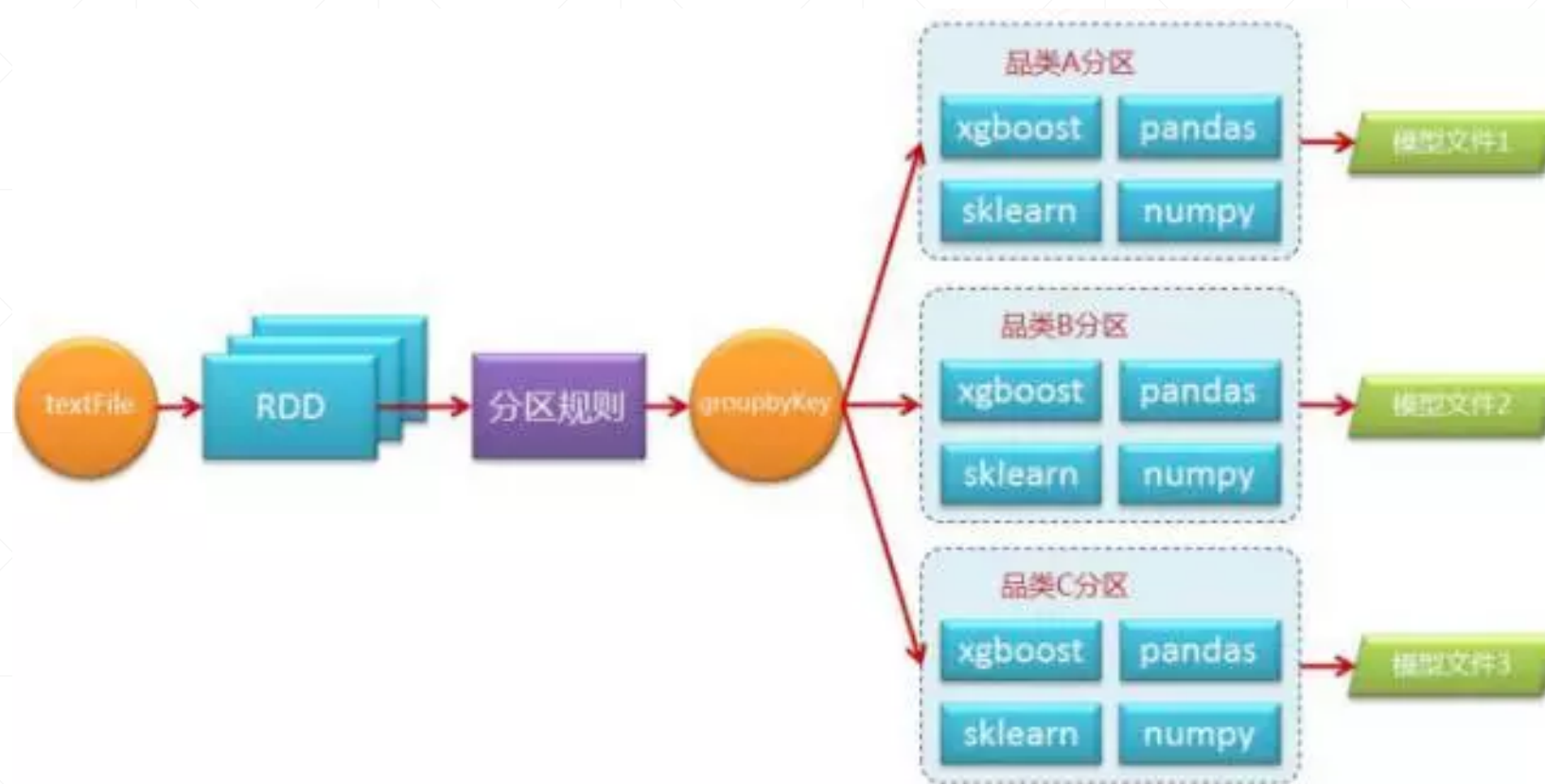
Architect



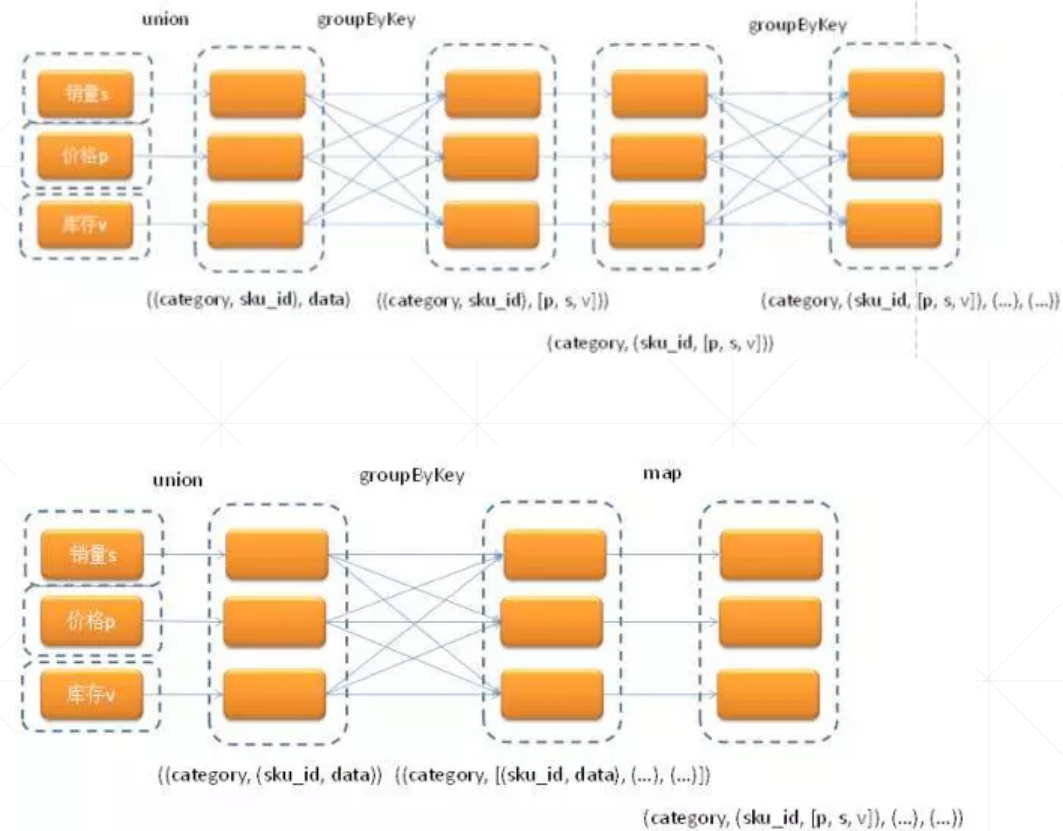
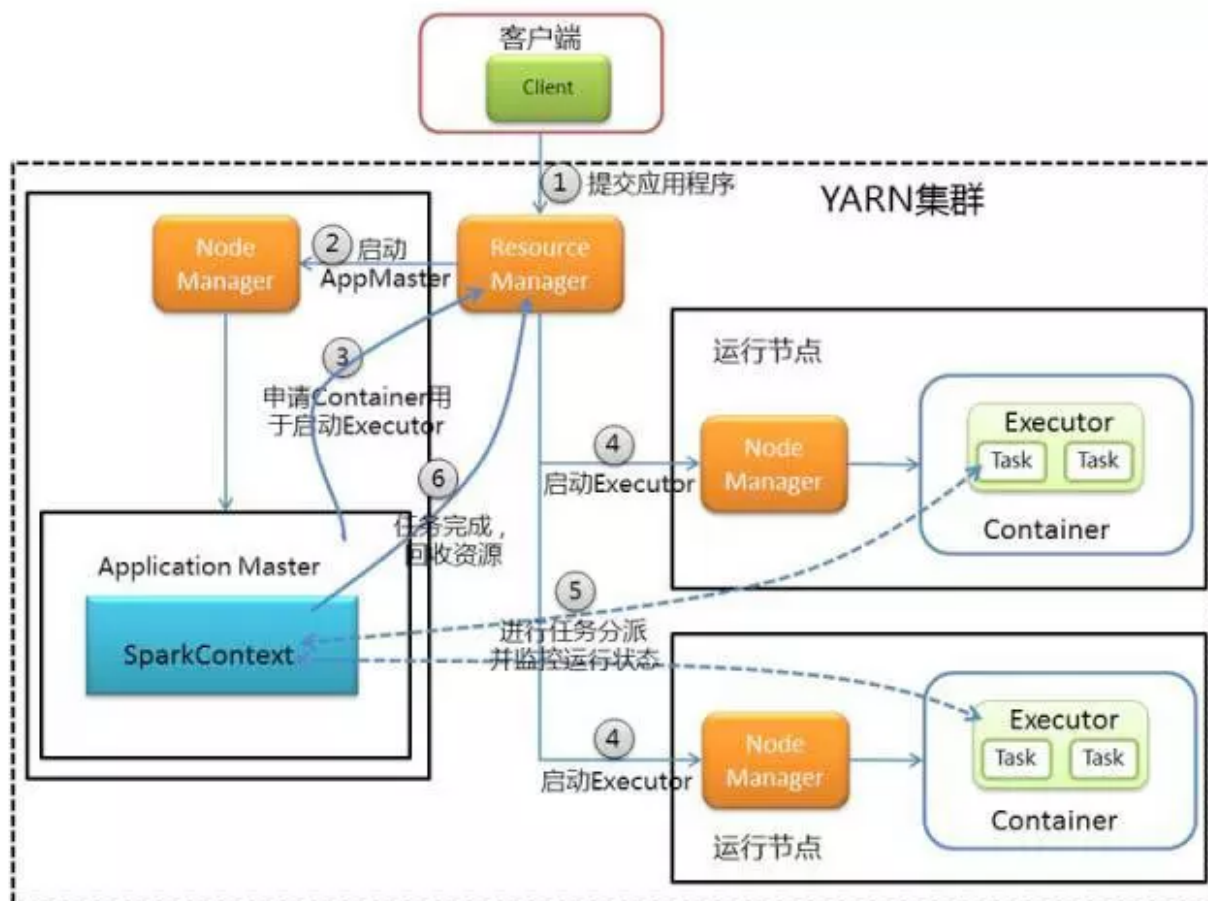
Tech Stack:



Parallel Processing Using Spark



Optimization:



参考:

- [Spark技术在京东智能供应链预测的应用](#)
 - [京东快物流背后的核心技术盘点](#)
-

求职拿Offer经历分享

Ben

Data Engineer 求职

1. 面试常见问题
2. 如何准备面试
3. 与SE的区别和联系
4. 需要注意的问题

Data Engineer 面试常见问题

1. Data 相关 projects
 2. 大数据基本知识 hadoop, spark, mr, hive, streaming
 3. Coding 能力
 4. Case study
-
-

Data Engineer 面试准备

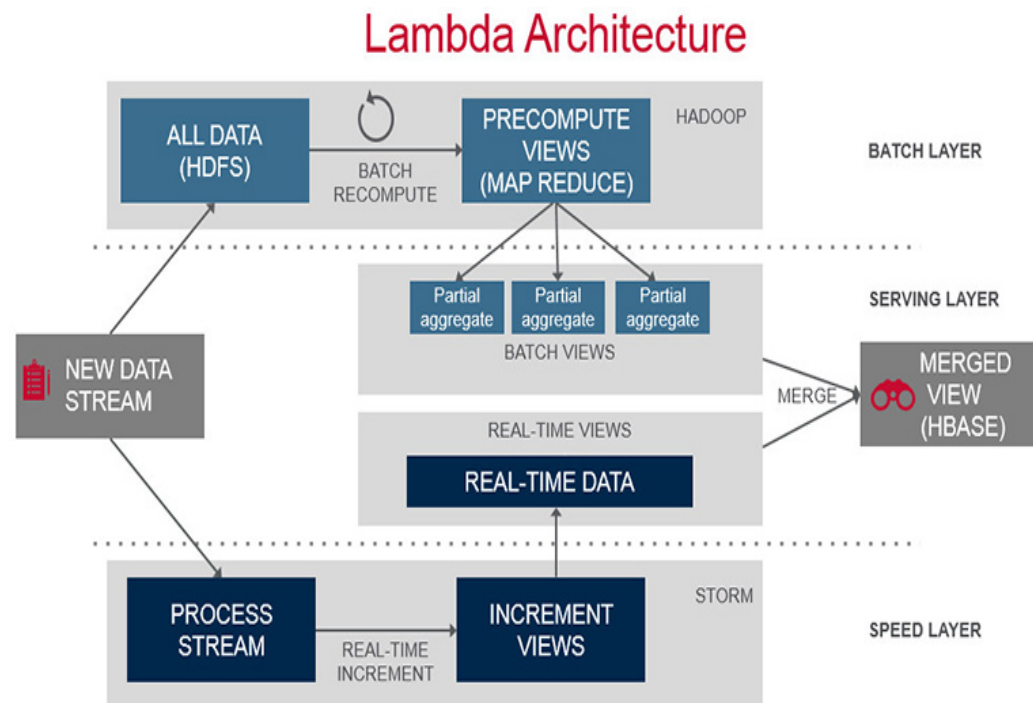
1. 熟悉项目每一个部分
 2. 课程内容, Hadoop: The Definitive Guide, 源码
 3. 刷题
 4. 积累经验
-
-

Data Engineer 与 Software Engineer 区别与联系

1. 知识点多但集中
2. Coding 要求相对低
4. 属于SE 一种

Data Engineer 注意事项

1. 同样的title, 要求完全不同
2. Lambda Architecture
3. 不断扩充知识



报名

- 今日报名, 享受立减优惠, 详情联络和咨询我们
- <https://www.DataAppLab.com/DE>
 - 填写报名表咨询
 - 电话咨询 +1 (800) 485 – 7918
 - Email咨询: DataEngineer@DataAppLab.com
 - 微信公众号 “大数据应用”
- 加微信群主私聊(微信号: data-engineer0)



Q&A?
