

文章硬不硬核，你说了算，
土哥怒肝大数据学习路线一条龙！
(万字长文+资源分享)

本文作者：在 IT 中穿梭旅行

本文档来自公众号：3 分钟秒懂大数据

微信扫码关注，土哥拉你进大数据技术
交流群，获取更多大数据技术



备注：加技术群



前言

大家好，我是土哥。

随着读者数量的不断增加，从 8 月底到 10 月中旬，3 分钟秒懂大数据 公众号粉丝数已经突破 7200+。

用户概况

7216

用户总数

146

今日新增

• 统计今日00:00到当前数据 •

短短几个月，看着越来越多的读者通过我写的文章而关注到该公众号，我好开心，同时也诚惶诚恐，害怕无法持续输出高质量文章，让读者失望。但我知道，只有写出让读者心动的文章，才能得其心，土哥会好好加油，为大家持续输出精品文章的。

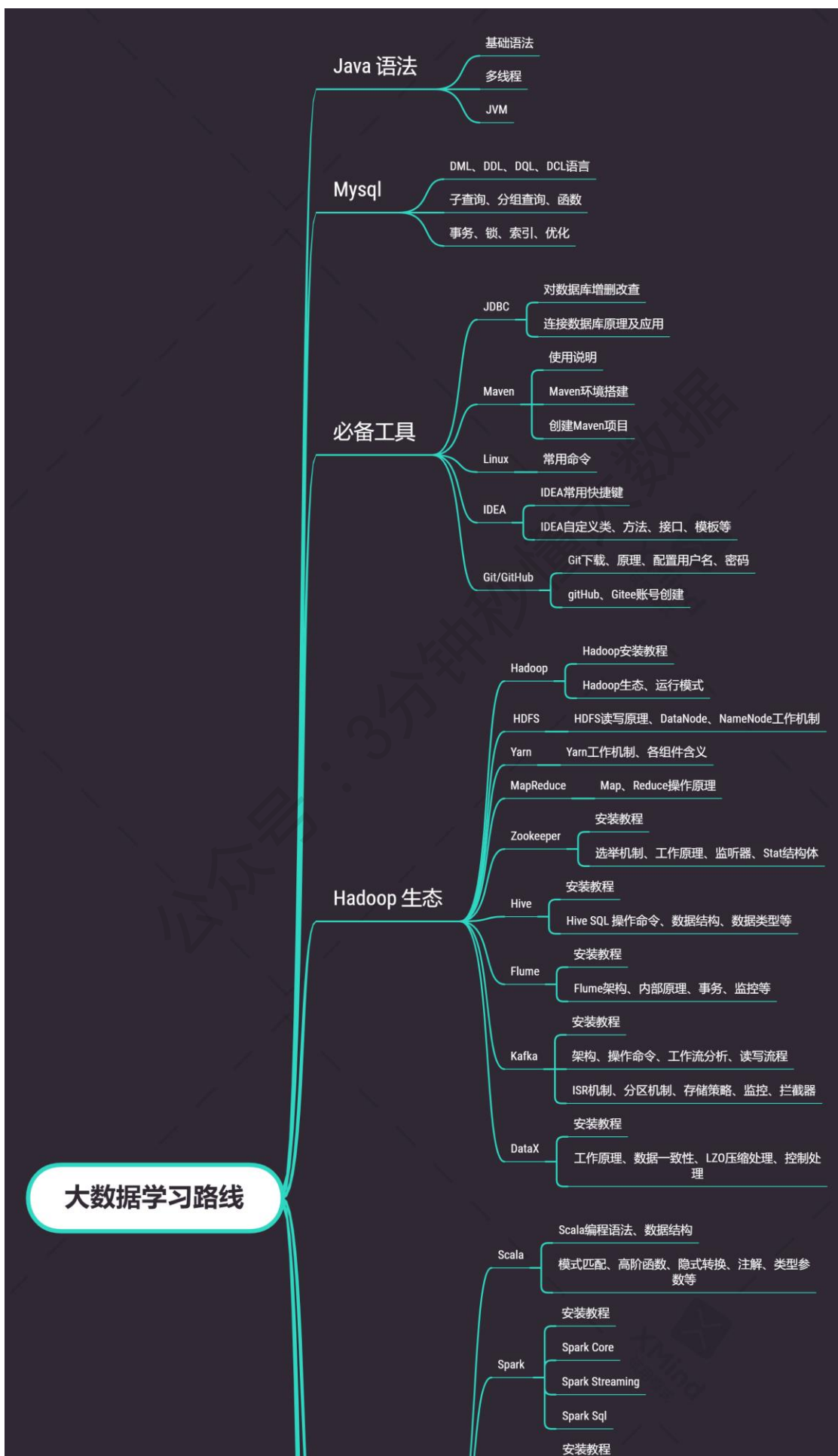
有读者问我，对于大数据小白或者跨行业者，应该如何学习大数据呢？或者说有什么推荐的书籍或者网站？

其实这个我很有心得啦，因为我研究生学的就是大数据与人工智能方向，所以了解很多关于这方面的书籍，并保留了一些学习网站。

哈哈，不要问我为什么会这么清楚，因为我就是这样采坑过来的，因为土哥很清楚，对于一个大数据小白或者跨行业者，如果自学大数据系列组件和教程，其实难度很大，光各种组件加起来有二三十个，是不是想想就头疼。

但是，跟着土哥走，让你学会大数据肯定没问题的。

接下来，我将根据自己的学习经验并结合一些书和视频，给大家全面分享，如何从一名小白摇身一变成为大数据开发工程师！学习网站、书籍以及软件安装教程已经全部帮大家整理好了。大纲目录如下：



一、Java

★ 官网链接：

<https://www.java.com/zh-CN/>

JAVA 语言是学习大数据的基础，因为好多大数据组件的源码都是由 JAVA 所编写的，如果你刚接触 JAVA、推荐先阅读《JAVA 核心技术》内容。



视频部分推荐：

★ 宋红康 - 全网最全 Java 零基础入门教程

<https://www.bilibili.com/video/BV1Kb411W75N>

（只看 Java 8 部分即可）

通过视频教程和书籍结合，安装 JDK1.8 或者 JDK11, 熟悉 JAVA 语法。

学完基础后，开始学习一些 JAVA 的高级特性，JAVA 虚拟机和并发编程是面试时的必考点，推荐看《深入理解 java 虚拟机》，重点学习内存、垃圾回收、类加载机制这几部分。

并发编程推荐《java 并发编程的艺术》，强烈推荐从头到尾好好阅读几遍，这是你冲击一二线大厂的本钱。

如果觉得安装软件麻烦，准备好一款在线、随时随地写代码的工具，不用在本地去安装和配置，是初学者的好帮手

★ 推荐网站 - 菜鸟工具

<https://c.runoob.com/compile/10>

准备一款记笔记的软件，学编程的过程中老师的思路、自己遇到的问题、解决问题的方法、心得感悟、遇到的好资料，都要记下来哦，**好记性不如烂笔头**，可以帮助你事半功倍。

★ 推荐软件 - Typora 工具

<https://www.typora.io/>

推荐使用 Markdown 语法编写，

★ 教程 - Markdown 教学

<https://www.bilibili.com/video/BV1no4y1k7YT/>

上述 4 本书结合视频教程，你相信，你很快可以成为合格的 JAVA 编程者。

二、Mysql

★ 官网链接：<https://www.mysql.com/>

为什么要把 Mysql 放到第二部分呢，因为它很重要，可以说非常重要，在大数据存储组件中，包含很多分布式数据库，都需要编写一定的 SQL 代码，而这些都是以 SQL 语法为基础。

同时在大数据计算组件，如 Flink SQL 和 Spark SQL 中，也需要用到 SQL 语法，所以，相信小伙伴们知道 SQL 的重要性了吧。

刚学习数据库时，强烈推荐下载解压版进行安装，安装教程如下，包含压缩包。

★ Mysql 安装教程：

https://blog.csdn.net/weixin_38201936/article/details/81605640

安装成功后开始学习 SQL 语法，推荐书籍《mysqlbi 必知必会》简单入门，配合下面的视频教程一块学习。



对于 Mysql 的视频资源，推荐以下内容：

★ 老杜 - mysql 入门基础 + 数据库实战

<https://www.bilibili.com/video/BV1Vy4y1z7EX>

★ 尚硅谷 - MySQL 基础教程

<https://www.bilibili.com/video/BV1xW411u7ax>（小姐姐讲课）

在线练习

★ SQL 自学网:<http://xuesql.cn/>

★ SQL 在线运行：

<https://www.bejson.com/runcode/sql/>

★ SQL - 菜鸟教程

<https://www.runoob.com/sql/sql-tutorial.html>

★ MySQL - 菜鸟教程

<https://www.runoob.com/mysql/mysql-tutorial.html>

学完简单语法之后，还不够，因为面试时这些只是最基础知识，你还需要掌握深层次知识点，例如 InnoDB 存储引擎，索引，锁，事务，性能调优等。

★ 推荐书籍《mysql 技术内幕 InnoDB 存储引擎》、《高性能 Mysql》等。

完成 JAVA 和 Mysql 的学习教程后，你已经基本具备转行学大数据内容了。

三、必备工具

3.1 JDBC

★JDBC Maven 下载链接:

<https://mvnrepository.com/artifact/org.xerial/sqlite-jdbc>

JDBC 是一个连接数据库的驱动，通过该驱动，可以对数据库进行增删改查操作。

★JDBC 教学视频:

<https://www.bilibili.com/video/av21399547>

3.2 Maven

★Maven 官网链接:

<http://maven.apache.org/what-is-maven.html>

★Maven 仓库链接:

<https://mvnrepository.com/>

Apache Maven 是一个软件项目管理和理解工具。基于项目对象模型 (POM) 的概念，Maven 可以从一条中央信息中管理项目的构建、报告和文档。

说简单点，Maven 就是一个仓库，通过 Maven 仓库，你可以下载你想要的 jar 包和依赖，你可以通过菜鸟教程学习 Maven 如何使用，其实很简单。也可以看 Maven 实战书籍。



★ Maven 菜鸟教程:

<https://www.runoob.com/maven/maven-tutorial.html>

★ Maven 安装教程:

<https://www.zhihu.com/question/20104186>

★ Maven 视频教程:

https://www.bilibili.com/video/BV1Fz4y167p5?from=search&seid=987708061267699949&spm_id_from=333.337.0.0

3.3 IDEA

★ IDEA 官网链接:

<https://www.jetbrains.com/idea/>

DEA 全称 IntelliJ IDEA，是 Java 编程语言开发的集成环境。目前在业界被公认为最好的 Java 开发工具，尤其在智能代码助手、代码自动提示、重构、JavaEE 支持、各类版本工具(git、svn 等)、JUnit、CVS 整合、代码分析、创新的 GUI 设计等方面的功能可以说是超常的。

在学习 Java 的时候就可以下载这个软件，傻瓜式安装，然后百度找一下破解教程就可以，如果不会使用的话，可以按照下面的教程简单学习下。

★ IDEA 使用视频教程:

https://www.bilibili.com/video/BV117411A7Ef?from=search&seid=15589514687856441260&spm_id_from=333.337.0.0

3.4 Git

★ GIT 官网链接:

<https://git-scm.com/>

Git 是一个开源的分布式版本控制系统，用于敏捷高效地处理任何或小或大的项目。

说简单点，Git 工具 可以将你写的代码 通过本地仓库 - push 到互联网的远程仓库上 (github / gitee) ,直接在官网下载一个 Git 安装包，然后傻瓜式安装。根据下面的 Git 教程，学一下简单的 Git 命令,比如简单的一些命令:

1. git clone
2. git pull
3. git add .
4. git commit -m 'xxx'
5. git push
6. git log

书籍推荐，可以看 Pro Git 中文书籍就足够了，可以根据下面的教学链接辅助学习。

★ Git 菜鸟教程:

<https://www.runoob.com/git/git-tutorial.html>

★ Pro Git 中文书籍在线教程

<http://git-scm.com/book/zh/v2>

★ Git 视频教程:

https://www.bilibili.com/video/BV1FE411P7B3?from=search&seid=4816308847064529471&spm_id_from=333.337.0.0

★ 通过 Git 往远程仓库新建项目教程:

https://blog.csdn.net/weixin_38201936/article/details/84836043

3.5 Linux

上面的必备工具都掌握后，就可以开始进入大数据学习阶段，大数据学习的第一课就是要掌握 **Linux** 教程。

我们知道，在企业中，所有的大数据组件、服务器等都在 **Linux** 环境进行安装、部署、运行，所以学习大数据组件的前提就是必须学会如何使用 **Linux** 命令

★ Linux 视频教程:

https://www.bilibili.com/video/BV1rq4y1s7yE?spm_id_from=333.788.b_765f64657363.2

直接从第 30 集开始入手就行。

★ Linux 推荐书籍:《鸟哥的 Linux 私房菜》

四、Hadoop 生态

恭喜各位小伙伴，前面的 **Linux** 教程完成后，你已经踏进大数据的门槛，接下来土哥会将大数据分为多个领域，给大家介绍一下，通过学习哪些组件，就可以成为一名大数据工程师!

1. Hadoop 开发工程师

职位描述:

1. Hadoop 技术栈的开发和管理，解决实际业务挑战，YARN, HDFS, MapReduce, Spark;
2. 承担千台-万台规模 Hadoop YARN 集群的管理工作，与业务一起解决性能优化、容量规划、预算审计等问题，保障集群高效稳定经济运行。

职位要求:

1. 熟悉 Hadoop 相关基础设施;

2. 大数据开发工程师-数据治理方向

职位描述：

- 1、负责数据中台数据治理工作，包括元数据管理、数据质量检查、成本优化等系统的设计、开发及应用，提升数据易用性、安全性及稳定性；
- 2、孵化并持续迭代优化数据治理产品，以系统化、智能化能力高效支撑数据业务快速发展。

职位要求：

1. 熟练使用 SQL、Python、Java 等工具进行大型数据分析及建模；
2. 熟练使用 Hadoop、Hive、Spark、Flink 等组件进行大数据场景数据开发；

3. 大数据算法工程师-数据挖掘方向

职位描述：

1. 为短视频和直播业务提供基础数据挖掘支持，包括用户画像挖掘、热点挖掘等；
2. 全面了解业务发展，提供基于机器学习、深度学习、文本挖掘、复杂网络等算法方案解决复杂的实际业务问题。

职位要求：

1. 有大数据集、分布式计算工具(Hadoop, Spark, Hive, Storm, Flink)等应用开发经验优先；
2. 有扎实的数据结构和算法功底，熟悉机器学习、自然语言处理、数据挖掘中一项或多项；

4. 大数据运维工程师

主要负责大数据相关系统/平台的维护，确保其稳定性，更多的是对大数据系统的维护。 职位描述：

1. 负责商业化 OLAP 引擎稳定性保障工作，提升服务 SLA 标准；
2. 参与设计并实现商业化 OLAP 引擎自动化运维平台，及其日常需求开发；

职位要求：

1. 熟悉 Java 编程，掌握至少一种脚本语言：python/shell 等；
2. 熟悉 Linux 操作系统的基础运维，有 ElasticSearch、Druid 等大数据引擎运维或使用经验者优先。

5. 大数据开发工程师-数仓方向

职责描述：

1. 负责离线和实时数仓构建；
2. 解决海量数据处理的调度、资源优化、数据质量、可用性、可靠性、监控等问题。

职位要求：

1. 具有海量实时、离线数仓建设落地经验；
2. 具备较强的编码能力，熟悉 SQL，Python，Hive，Spark，Kafka，Flink 中的多项，有至少 TB 以上级大数据处理经验，理解平台原理；

6. 大数据开发工程师-流计算方向

职责描述：

实时计算方向，负责推荐系统架构实时计算系统的设计和开发，保障系统稳定和高可用；为大规模推荐系统设计和实现合理的实时（流式计算）数据系统；

职位要求：

1. 对流式计算系统有深入的了解，在生产环境有 TB 级别 Flink 实时计算系统开发经验，深入掌握 Flink DataStream、FlinkSQL、Flink Checkpoint、Flink State 等模块，有 Flink 源码阅读经验优先；
2. 熟悉常见消息队列原理和应用调优，有 Kafka、Plusar、RocketMQ 等项目源码阅读经验优先；
3. 熟悉 Java、C++、Scala、Python 等编程语言；
4. 有数据湖开发经验，熟悉 Hudi、Iceberg、DeltaLake 等至少一项数据湖技术，有源码阅读经验优先；
5. 熟悉其他大数据系统经验者优先，YARN、K8S、Spark、SparkSQL、Kudu 等；
6. 有存储系统经验加分，HBase、Cassandra、RocksDB 等。

土哥将大数据分为 6 个方向：包含 Hadoop 方向、数据治理方向、数据挖掘方向、数据仓库方向、流计算方向、大数据运维方向。当然，还可以细分更多的领域。

在这么多方向中，不同的方向需要学习不同的组件，但是，万变不离其宗，大数据领域的组件其实可以从三部分进行概括，分别为：**存储组件**、**计算组件**、**调度组件**。

1.存储组件：HDFS、Yarn、Zookeeper、Hive、Flume、Kafka、Hbase、Redis、ElasticSearch、ClickHouse 等。

2.计算组件：MapReduce、Spark、Flink、Ray 等。

3.调度组件：DolphinScheduler、Oozie、AirFlow，Yarn 等。

将组件进行分类后，小伙伴们应该一目了然，然后根据不同的岗位，学习对应的组件即可，毕竟在工作岗位中，并不是要对所有组件都**熟悉、精通**，对自己工作中的组件要做到熟悉、精通，深入到源码。对工作之外的组件做到了解其用途、明白其使用方式即可。

小伙伴学习大数据组件时，必须要掌握的基本能力包含以下 3 部分：

1. 能成功部署组件，并运行。
2. 了解组件的架构设计、运行原理。
3. 可以编写简单的代码运行。

下面土哥根据大纲的路线图，给大家整理大数据领域的各组件的学习视频、资料、以及书籍，如下图所示：

Hadoop3. X组件兼容版本对应表格				
组件名称	版本型号	Window/Linux	是否互相兼容	Jar包依赖
Hadoop	hadoop-3.2.2.tar.gz	Linux	兼容	
Zookeeper	zookeeper-3.4.14.tar.gz	Linux	兼容	
Kafka	kafka_2.12-2.2.1.tgz	Linux	兼容	
scala	scala-2.11.12.zip	Linux	兼容	
Hbase	hbase-2.3.6-bin.tar.gz	Linux	兼容	
Hive	apache-hive-3.1.2-bin.tar.gz	Linux	兼容	
Flink	flink-1.13.2-bin-scala_2.11.tgz	Linux	兼容	flink-sql-connector-hive-3.1.2.11-1.13.2.jar
Spark	spark-3.1.2-bin-hadoop3.2.tgz	Linux	兼容	
Redis	redis-6.2.6.tar.gz	Linux	兼容	
Mysql	mysql-5.7.19-1.el7.x86_64.rpm	Linux	兼容	
JDK	jdk-8u202-linux-x64.tar.gz	Linux	兼容	
Flume	apache-flume-1.9.0-bin.tar.gz	Linux	兼容	

4.1 Hadoop

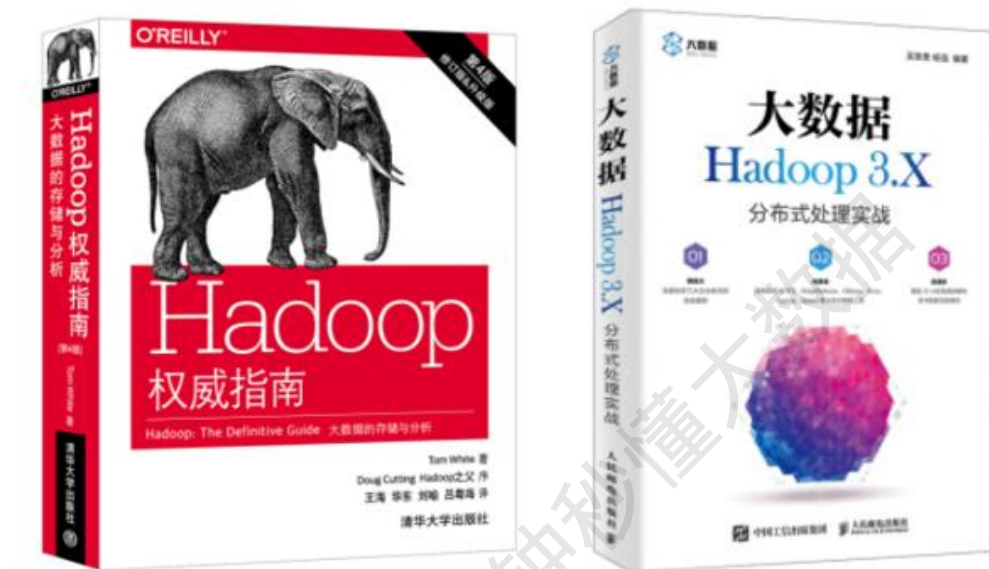
★Hadoop 官网链接：

<http://hadoop.apache.org/>



Hadoop 是 Apache 基金会所开发的分布式系统基础架构。目前 80% 的组件都要基于 Hadoop 组件来运行，用户可以在不了解分布式底层细节的情况下，开发分布式程序。Hadoop 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。

★ 推荐书籍：Hadoop 权威指南、Hadoop3.X 分布式处理实战



书中全面讲解了 Hadoop 的由来、设计架构，对 HDFS、MapReduce 进行了全面的讲解。

★ Hadoop 组件安装教程：

https://blog.csdn.net/weixin_38201936/article/details/85096496

★ Hadoop 视频教程：

https://www.bilibili.com/video/BV1rq4y1s7yE?spm_id_from=333.788.b_765f64657363.2

当 Hadoop 基本概念、原理都掌握后，可以看看其源码是如何设计的

★ Hadoop GitHub 源码链接：

<https://github.com/apache/hadoop>

4.2 Yarn

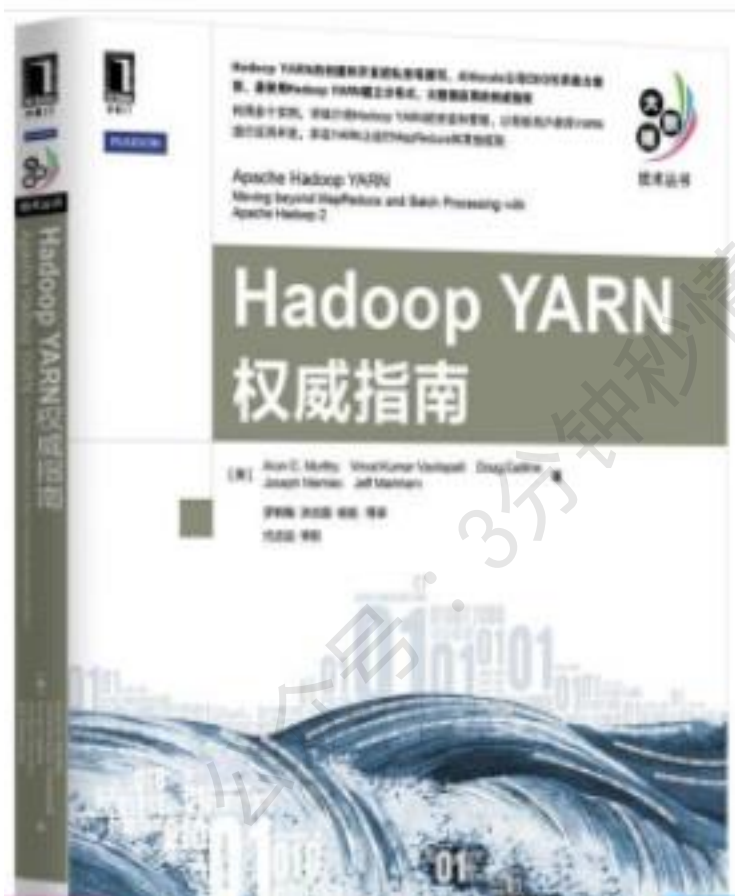
★ Yarn 官网链接：

<https://yarnpkg.com/>

Apache Hadoop YARN（Yet Another Resource Negotiator）是一种新的 Hadoop 资源管理器，它是一个通用资源管理系统，可为上层应用提供统一的资源管理和调度，它的引入为集群在利用率、资源统一管理和数据共享等方面带来了巨大好处。

目前计算组件如 Spark、Flink、MapReduce 都和 Yarn 资源管理进行关联，在集群部署时，都会采用高可用进行部署，将计算任务都提交到 Yarn 上进行管理。

★Yarn 书籍推荐：Hadoop Yarn 权威指南，这本书对 Yarn 资源管理进行了很详细的讲解，值得小伙伴们观看呦！



★Yarn 视频教程：5 分钟让你轻松学会 Yarn 的原理，包含 ResourceManager，NodeManager 等概念。

https://www.bilibili.com/video/BV1Pa4y1t7nf?from=search&seid=11247307393237909491&spm_id_from=333.337.0.0

★Yarn GitHub 源码链接：

4.3 Zookeeper

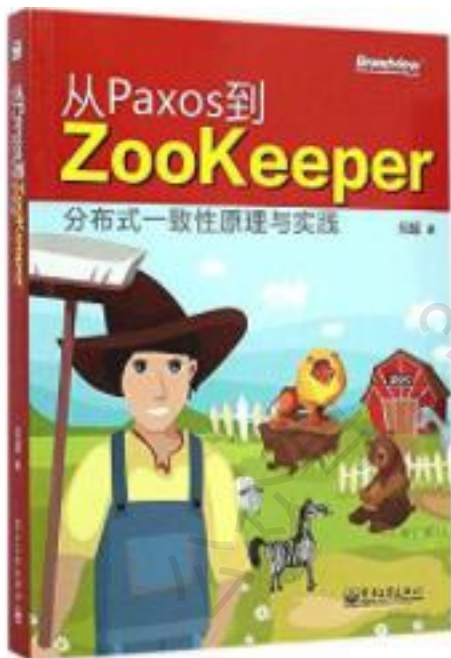
★ Zookeeper 官网链接:

<http://zookeeper.apache.org/>

ZooKeeper 是一个分布式的，开放源码的分布式应用程序协调服务，是 Google 的 Chubby 一个开源的实现，是 Hadoop 和 Hbase 的重要组件。它是一个为分布式应用提供一致性服务的软件，提供的功能包括：配置维护、域名服务、分布式同步、组服务等。

说的简单点，ZooKeeper 就是一个动物园管理员，他是拿来管大象(Hadoop)、蜜蜂(Hive)、小猪(Pig)、Spark、Flink、Kafka 等组件的管理员，从程序员的角度来讲可以理解为是一个整体监控系统。如果这些组件采用分布式部署时，其中主节点宕机后，这时候 Zookeeper 就会重新选举出 leader。这是它最大的作用所在。

★ zookeeper 书籍推荐:



★ Zookeeper 视频教程:

https://www.bilibili.com/video/BV1to4y1C7gw?from=search&seid=2755195091410266102&spm_id_from=333.337.0.0

★ Zookeeper 组件安装教程:

https://blog.csdn.net/weixin_38201936/article/details/88821559

★ Zookeeper GitHub 源码链接:

<https://github.com/apache/zookeeper>

4.4 Hive

★ Hive 官网链接：

<https://hive.apache.org/>

Hive 在大数据中扮演着比较重要的角色，不管是离线数仓还是实时数仓或者数据研发工程师，基本都会用到 Hive。

Hive 是一款建立在 Hadoop 之上的开源数据仓库系统，可以将存储在 Hadoop 文件中的结构化、半结构化数据文件映射为一张数据库表，基于表提供了一种类似 SQL 的查询模型，称为 Hive 查询语言（HQL），用于访问和分析存储在 Hadoop 文件中的大型数据集。

Hive 核心是将 HQL 转换为 MapReduce 程序，然后将程序提交到 Hadoop 群集执行。

学习 Hive 给大家推荐的书籍有：Hive 编程指南、Hive 性能调优实战，这两本书讲的都很全面，有丰富是实战案例。

★ Hive 集群安装教程链接：

https://blog.csdn.net/weixin_38201936/article/details/88598008

★ Hive 实战离线系列文章

<https://mp.weixin.qq.com/s/5YDO-Y4mfj5dMnQodG-7vQ>

★ Hive 实战实时数仓系列文章

https://mp.weixin.qq.com/s/1c-iPaGSlMEv_J_960SIYw

★ Hive GitHub 源码链接：

<https://github.com/apache/hive>

4.5 Kafka

★ Kafka 官网链接：

<http://kafka.apache.org/>

Kafka 是一种高吞吐量的分布式发布订阅消息系统，目前在后端、大数据领域中使用特别多，Kafka 包含生产者、消费者模型，通过生产者将数据实时存储在消息队列中，使用 Spark、Flink 等计算引擎从消费者读取数据进行计算。

Kafka 组件非常重要，一定要好好学习，这里土哥给大家推荐 KafKa 权威指南这本书，是初学者入门的不二选择。

★ Kafka 视频学习链接：

<http://kafka.apache.org/>

2021 年最新版 Kafka 视频教程，跟着视频可以快速学会 Kafka 的使用

对 Kafka 理论知识了解清楚后，建立着手部署 Kafka 集群，超级详细的安装教程请看这里：

https://blog.csdn.net/weixin_38201936/article/details/89226897

★ Kafka 操作命令：

https://blog.csdn.net/weixin_38201936/article/details/117595931

★ Kafka gitHub 链接：

<https://github.com/apache/kafka>

针对前面的内容和岗位介绍，选择合适的组件进行学习，Hadoop 岗位已经可以胜任。

五、Spark 生态

在大数据领域，计算组件有多个，从离线计算框架 MapReduce 到流式计算框架 Storm、Spark、Flink 等。在 Spark 领域，小伙伴们需要学习 Scala 语言，因为 Spark 的底层源码是用 Scala 所编写。

★ Scala 官网链接：

<https://www.scala-lang.org/>

★ Scala 菜鸟教程链接：

<https://www.runoob.com/scala/scala-tutorial.html>

Scala 语言要重点掌握编程语法、Scala 的数据结构、模式匹配、高阶函数、隐仕转换、注解、类型参数等。

掌握 Scala 之后，就可以着手学习 Spark 计算框架。

★ Spark 官网

<http://spark.apache.org/>

由于现在 Spark3.0 版本已经出来了，建议刚入行大数据的小伙伴直接学习 Spark3.0，Spark 于 2009 年诞生于加州大学伯克利分校实验室，2010 年被开源，在大数据领域，短短 10 年 Spark 已经成为世界顶尖的分布式快速通用计算框架。

Spark 生态圈包含 Spark Core，Spark SQL，Spark Streaming，Spark MLib，GraphX 等等，小伙伴们刚学时，建议重点学习 Spark Core,Spark SQL，Spark Streaming 等内容，多看官网的官方文档。

Spark 书籍推荐：图解 Spark 核心技术与案例实战

★ Spark3.0 入门教学视频

https://www.bilibili.com/video/BV1Xz4y1m7cv?from=search&seid=17655987947940698399&spm_id_from=333.337.0.0

★ Spark 入门安装视频

https://blog.csdn.net/weixin_38201936/article/details/85233263

★ Spark GitHub 链接：

<https://github.com/apache/spark>

六、Flink 生态

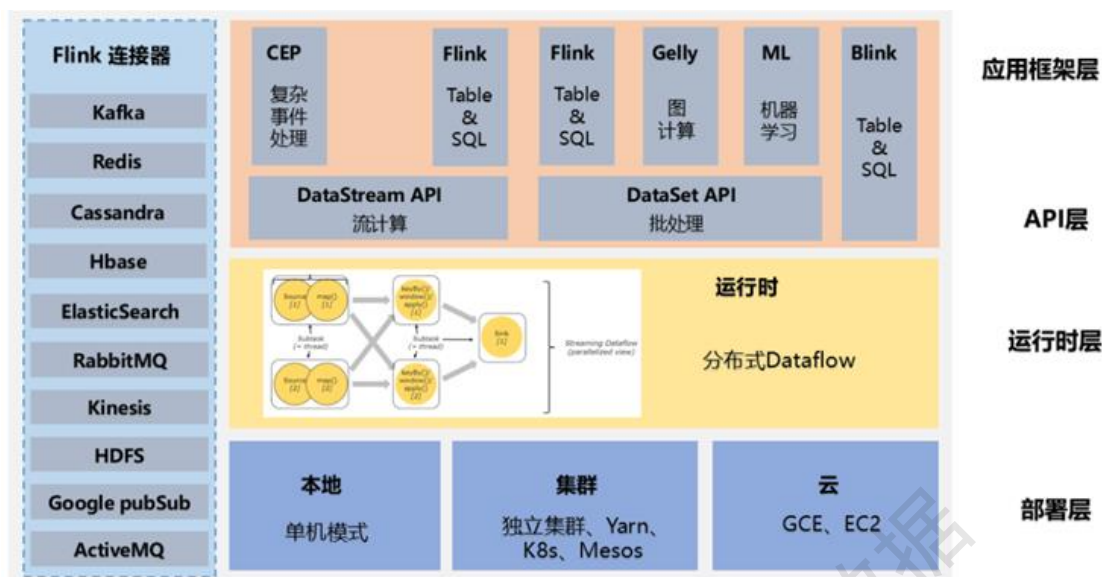
Flink 组件是大数据领域必学的实时流计算组件。

Flink 是一个以流为核心的高可用、高性能的分布式计算引擎。具备流批一体，高吞吐、低延迟，容错能力，大规模复杂计算等特点，在数据流上提供数据分发、通信等功能。在大数据领域扮演着非常重要的角色。

★ Flink 官网

<https://flink.apache.org/>

Flink 生态同样包含模块，如 Flink DataStream，Flink DataSet, Flink Table\SQL，Flink ML, Gelly 等，具体如下图：



出入门学习 Flink,土哥强烈推荐一本书籍：**Flink 核心原理与实践**，这本书从十六个章节全面介绍 Flink 的架构设计、运行原理、任务提交机制、四大基石、内存模型、SQL 提交机制、监控指标等。

★ Flink 推荐教学视频：

https://www.bilibili.com/video/BV1qy4y1q728?from=search&seid=3522561233646130305&spm_id_from=333.337.0.0

集群安装方面，直接看土哥写的 入门安装教程、非常详细，图文并茂，傻瓜式安装即可。

★ Flink 入门安装教程：[安装教程](#)

安装完成后，可以阅读我的 Flink 系列文章，一共发表了 23 篇 Flink 文章，从基础知识到源码分析

★ 秒懂 Flink 系列：[秒懂 Flink 系列](#)

掌握一定的基础知识后，可以通过提问方式巩固知识点，土哥写了 Flink 面试大全，全文 6 万字，110 知识点，160 张原理图，看完之后，轻松应对大厂的 Flink 面试。

★ 秒懂 Flink 面试大全：[6 万字 Flink 面经总结](#)

基础掌握扎实后，可以深入了解 Flink 的源码，如 Flink 应用提交、Flink SQL 应用提交、Flink 内存优化、指标监控等。直接在 GitHub 将源码下载下来。

Flink gitHub 源码链接：

<https://github.com/apache/flink>

七、总结

以上就是土哥花了几天大夜、结合个人经验、并且参考了大量网上的文章和评论总结而成的学习路线，真的是非常不容易。

文章全面介绍了大数据的学习路线，详细阐述如何从一名小白晋升为大数据开发工程师，土哥希望大家可以认真阅读、根据自己想要投递的岗位选择合适的组件进行学习。

如果大家觉得满意请务必三连支持下，路线文档和所有学习资料已经帮大家整理好，在我的公众号【3 分钟秒懂大数据】后台回复【大数据学习路线】就能获取。

非常欢迎大家加我**个人微信**
拉你进**大技术**交流群，有问题我们在群内一起讨论



长按上方扫码二维码，加我微信，拉你进群