压缩过的大数据Spark蘑菇云行动前置课程视频百度云分享链接

链接：http://pan.baidu.com/s/1cFqjQu

**SCALA专辑**

Scala深入浅出经典视频

链接：http://pan.baidu.com/s/1i4Gh3Xb 密码：25jc

DT大数据梦工厂大数据spark蘑菇云Scala语言全集（持续更新中）

[http://www.tudou.com/plcover/rd3LTMjBpZA](qq://txfile/)/

|  |  |
| --- | --- |
| 1 | Spark视频王家林第1课：大数据时代的“黄金”语言Scala |
| 2 | Spark视频王家林第2课：Scala零基础实战入门的第一堂课及如何成为Scala高手（修改版） |
| 3 | Spark视频王家林第3课：在IDE下开发第一个Scala程序纯傻瓜式彻底透彻解析 |
| 4 | Spark视频王家林第4课：零基础彻底实战Scala控制结构及Spark源码解析 |
| 5 | Spark视频王家林第5课：零基础实战Scala函数式编程及Spark源码解析 |
| 6 | Spark视频王家林第6课：零基础实战Scala集合操作 |
| 7 | Spark视频王家林第7课：零基础实战Scala面向对象编程及Spark源码解析 |
| 8 | Spark视频王家林第8课：零基础实战Scala最常用数据结构Map和Tuple及Spark源码鉴赏 |
| 9 | Spark视频王家林第9课：Scala类和对象彻底实战和Spark源码鉴赏 |
| 10 | Spark视频王家林第10课：Scala继承彻底实战和Spark源码鉴赏 |
| 11 | Spark视频王家林第11课：Scala面向接口彻底实战和Spark源码鉴赏 |
| 12 | Spark视频王家林第12课：Scala函数式编程进阶（匿名函数、高阶函数、函数类型推断、Currying）与Spark源码鉴赏 |
| 13 | Spark视频王家林第13课Scala模式匹配实战和Spark源码鉴赏 |
| 14 | 第14课Scala集合上的函数式编程实战及Spark源码鉴赏 |
| 15 | 第15课：Scala类型参数编程实战 |
| 16 | 第16课Scala implicits编程彻底实战 |
| 17 | Spark视频王家林第17课：Scala并发编程实战及Spark源码鉴赏 |
| 18 | Spark视频王家林第18课：scala偏函数、异常、lazy值编码实战及spark源码鉴赏 |
| 19 | Spark视频王家林第19课：Scala的包、继承覆写及Spark源码鉴赏 |
| 20 | Spark视频王家林第20课：scala提取器、注解深度实战详解及spark源码鉴赏 |
| 21 | Spark视频王家林第21课：scala文件和xml操作实战及spark源码鉴赏（上）关于文件序列和和操作详解实战 |
| 22 | Spark视频王家林第22课：scala集合和高级函数操作实战及spark源码鉴赏 |

《Spark零基础视频：从零起步到调通第一个Wordcount》<http://www.tudou.com/plcover/IB9YwzdU8f0/>

|  |  |
| --- | --- |
| 1 | Spark视频王家林第1集：大数据最火爆语言Scala光速入门 |
| 2 | Spark视频王家林第2集：Scala面向对象彻底精通及Spark源码阅读 |
| 3 | Spark视频王家林第3课：Scala函数式编程彻底精通及Spark源码阅读 |
| 4 | Spark视频王家林第4课：Scala模式匹配、类型系统彻底精通与Spark源码阅读 |
| 5 | Spark视频王家林第5课：彻底精通Scala隐式转换和并发编程及Spark源码阅读 |
| 6 | Spark视频王家林第6课：精通Spark集群搭建与测试 |
| 7 | Spark视频王家林第7课：实战解析Spark运行原理和RDD解密 |
| 8 | Spark视频王家林第8课：彻底实战详解使用IDE开发Spark程序 |
| 9 | Spark视频王家林第9课：彻底实战详解 IntelliJ IDEA下的Spark程序开发 |
| 10 | Spark视频王家林第10课：底实战详解使用Java开发Spark程序 |
| 11 | Spark视频王家林第11课：彻底解密WordCount运行原理 |
| 12 | Spark视频王家林第12课：HA下的Spark集群工作原理解密 |

**SPARK专辑**

大数据Spark内核core源码解密

[http://www.tudou.com/plcover/JdpoUtqxmNs](qq://txfile/)/

|  |  |
| --- | --- |
| 13 | Spark视频王家林第13课：Spark内核架构解密 |
| 14 | Spark视频王家林第14课：Spark RDD解密 |
| 15 | Spark视频王家林第15课：RDD创建内幕彻底解密 |
| 16 | Spark视频王家林第16课：RDD实战（RDD基本操作实战及Transformation流程图） |
| 17 | Spark视频王家林第17课：RDD案例（join、cogroup、reduceByKey、groupByKey等） |
| 18 | Spark视频王家林第18课：RDD持久化、广播、累加器 |
| 19 | Spark视频王家林第19课：Spark高级排序彻底解秘 |
| 20 | Spark视频王家林第20课：Top N彻底解秘 |
| 21 | Spark视频王家林第21课：从Spark架构中透视Job |
| 22 | Spark视频王家林第22课：RDD的依赖关系彻底解密 |
| 23 | Spark视频王家林第23课：从物理执行的角度透视Spark Job |
| 24 | Spark视频王家林第24课：Spark Hash Shuffle内幕彻底解密 |
| 25 | Spark视频王家林第25课：世界第一的Spark Sort-Based Shuffle内幕工作机制、案例实战、源码剖析、优缺点及改进方式等彻底解密 |
| 26 | Spark视频王家林第26课：Spark Runtime（Driver、Masster、Worker、Executor）内幕解密 |
| 27 | Spark视频王家林第27课：Spark on Yarn彻底解密 |
| 28 | Spark视频王家林第28课：Spark天堂之门解密 |
| 29 | Spark视频王家林第29课：Master HA彻底解密 |
| 30 | Spark视频王家林第30课：Master的注册机制和状态管理解密 |
| 31 | Spark视频王家林第31课：Spark资源调度分配内幕天机彻底解密：Driver在Cluster模式下的启动、两种不同的资源调度方式源码彻底解析、资源调度内幕总结 |
| 32 | Spark视频王家林第32课：Spark Worker原理和源码剖析解密：Worker工作流程图、Worker启动Driver源码解密、Worker启动Executor源码解密等.av |
| 33 | Spark视频王家林第33课：Spark Executor内幕彻底解密：Executor工作原理图、ExecutorBackend注册源码解密、Executor实例化内幕、Executor具体工作内幕 |
| 34 | Spark视频王家林第34课：Stage划分和Task最佳位置算法源码彻底解密 |
| 35 | Spark视频王家林第35课：打通Spark系统运行内幕机制循环流程 |
| 36 | Spark视频王家林第36课：TaskScheduler内幕天机解密：Spark shell案例运行日志详解、TaskScheduler和SchedulerBackend、FIFO与FAIR、Task运行时本地性算法详解等 |
| 37 | Spark视频王家林第37课：Task执行内幕与结果处理解密 |
| 38 | Spark视频王家林第38课：BlockManager架构原理、运行流程图和源码解密 |
| 39 | Spark视频王家林第39课：BlockManager解密进阶：BlockManager初始化和注册解密、BlockManagerMaster工作解密、BlockTransferService解密、本地数据读写解密、远程数据读写解密 |
| 40 | Spark视频王家林第40课：CacheManager彻底解密：CacheManager运行原理流程图和源码详解 |
| 41 | Spark视频王家林第41课：Checkpoint彻底解密：Checkpoint的运行原理和源码实现彻底详解 |
| 42 | Spark视频王家林第42课：Spark Broadcast内幕解密：Broadcast运行机制彻底解密、Broadcast源码解析、Broadcast最佳实践 |
| 43 | Spark视频王家林第43课：Spark 1.6 RPC内幕解密：运行机制、源码详解、Netty与Akka等 |

《大数据spark性能优化》

[http://www.tudou.com/plcover/c74-UW2DP7o/](qq://txfile/)

|  |  |
| --- | --- |
| 44 | Spark视频王家林第44课：真正的Spark高手是如何炼成的？！ |
| 45 | Spark视频王家林第45课：王家林谈Spark性能优化第一季！ |
| 46 | Spark视频王家林第46课：王家林谈Spark性能优化第二季！ |
| 47 | Spark视频王家林第47课：王家林谈Spark性能优化第三季！ |
| 48 | Spark视频王家林第48课：王家林谈Spark性能优化第四季！ |
| 49 | Spark视频王家林第49课：王家林谈Spark性能优化第五季！ |
| 50 | Spark视频王家林第50课：王家林谈Spark性能优化第六季！ |
| 51 | Spark视频王家林第51课：王家林谈Spark性能优化第七季之Spark 钨丝计划解密! |
| 52 | Spark视频王家林第52课：王家林谈Spark性能优化第八季之Spark Tungsten-sort Based Shuffle 内幕解密 |
| 53 | Spark视频王家林第53课：王家林谈Spark性能优化第九季之Spark Tungsten内存使用彻底解密 |
| 54 | Spark视频王家林第54课：王家林谈Spark性能优化第十季之全球独家揭秘Spark统一内存管理！ |

《 王家林大数据Spark SQL从零起步彻底精通彻底实战  》全集[http://www.tudou.com/plcover/nhBqWLH2Asc/](qq://txfile/)

|  |  |
| --- | --- |
| 55 | Spark视频王家林第55课：60分钟内从零起步驾驭Hive实战 |
| 56 | Spark视频王家林第56课：揭秘Spark SQL和DataFrame的本质 |
| 57 | Spark视频王家林第57课：Spark SQL on Hive配置及实战 |
| 58 | Spark视频王家林第58课：使用Java和Scala在IDE中开发DataFrame实战 |
| 59 | Spark视频王家林第59课：使用Java和Scala在IDE中实战RDD和DataFrame转换操作 |
| 60 | Spark视频王家林第60课：使用Java和Scala在IDE中实战RDD和DataFrame动态转换操作 |
| 61 | Spark视频王家林第61课：Spark SQL数据加载和保存内幕深度解密实战 |
| 62 | Spark视频王家林第62课：Spark SQL下的Parquet使用最佳实践和代码实战 |
| 63 | Spark视频王家林第63课：Spark SQL下Parquet内幕深度解密 |
| 64 | Spark视频王家林第64课：Spark SQL下Parquet的数据切分和压缩内幕详解 |
| 65 | Spark视频王家林第65课：Spark SQL下Parquet深入进阶 |
| 66 | Spark视频王家林第66课：Spark SQL下Parquet中PushDown的实现 |
| 67 | Spark视频王家林第67课：Spark SQL下采用Java和Scala实现Join的案例综合实战（巩固前面学习的Spark SQL知识） |
| 68 | Spark视频王家林第68课：Spark SQL通过JDBC操作Mysql |
| 69 | Spark视频王家林第69课：Spark SQL通过Hive数据源实战 |
| 70 | Spark视频王家林第70课：Spark SQL内置函数解密与实战 |
| 71 | Spark视频王家林第71课：Spark SQL窗口函数解密与实战 |
| 72 | Spark视频王家林第72课：Spark SQL UDF和UDAF解密与实战 |
| 73 | Spark视频王家林第73课：Spark SQL Thrift Server实战 |
| 74 | Spark视频王家林第74课：Hive on Spark大揭秘 完整版 |
| 75 | Spark视频王家林第75课：Spark SQL基于网站Log的综合案例实战 |
| 76 | Spark视频王家林第76课：Spark SQL实战用户日志的输入导入Hive及SQL计算PV实战 |
| 77 | Spark视频王家林第77课：Spark SQL基于网站Log的综合案例实战之Hive更大规模数据导入、Spark SQL对数据UV操作实战 |
| 78 | Spark视频王家林第78课：Spark SQL基于网站Log的综合案例用户用户跳出率和新用户注册比例 |
| 79 | Spark视频王家林第79课：Spark SQL基于网站Log的综合案例综合代码和实际运行测试 |
| 80 | Spark视频王家林第80课：Spark SQL网站搜索综合案例实战 |
| 81 | Spark视频王家林第81课：一节课贯通Spark SQL工作源码流程 |

王家林大数据Spark Streaming从零起步（21集）

[http://www.tudou.com/plcover/nRM-f151vp0/](qq://txfile/)

|  |  |
| --- | --- |
| 82 | Spark视频王家林第82课：Spark Streaming第一课：案例动手实战并在电光石火间理解其工作原理 |
| 83 | Spark视频王家林第83课：透彻讲解使用Scala和Java两种方式实战Spark Streaming开发 |
| 84 | Spark视频王家林第84课：图解StreamingContext、DStream、Receiver并结合源码分析 |
| 85 | Spark视频王家林第85课：基于HDFS的SparkStreaming案例实战和内幕源码解密 |
| 86 | Spark视频王家林第86课：SparkStreaming数据源Flume实际案例分享 |
| 87 | Spark视频王家林第87课：Flume推送数据到SparkStreaming案例实战和内幕源码解密 |
| 88 | Spark视频王家林第88课：SparkStreaming 从Flume Poll数据案例实战和内幕源码解密 |
| 89 | Spark视频王家林第89课：SparkStreaming on Kafka之Kafka解析和安装实战 |
| 90 | Spark视频王家林第90课：SparkStreaming基于Kafka Receiver案例实战和内幕源码解密 |
| 91 | Spark视频王家林第91课：SparkStreaming基于Kafka Direct案例实战和内幕源码解密 |
| 92 | Spark视频王家林第92课：SparkStreaming中Tanformations和状态管理解密 |
| 93 | Spark视频王家林第93课：SparkStreaming updateStateByKey 基本操作综合案例实战和内幕源码解密 |
| 94 | Spark视频王家林第94课：SparkStreaming 实现广告计费系统中在线黑名单过滤实战 |
| 95 | Spark视频王家林第95课： 通过Spark Streaming的window操作实战模拟新浪微博、百度、京东等热点搜索词案例实战 |
| 96 | Spark视频王家林第96课： 通过Spark Streaming的foreachRDD把处理后的数据写入外部存储系统中 |
| 97 | Spark视频王家林第97课： 使用Spark Streaming+Spark SQL实现在线动态计算出特定时间窗口下的不同种类商品中的热门商品排名 |
| 98 | Spark视频王家林第98课： 使用Spark Streaming实战对论坛网站动态行为的多维度分析（上） |
| 99 | Spark视频王家林第99课： 使用Spark Streaming实战对论坛网站动态行为的多维度分析（下）：完整案例实现、测试、Debug等 |
| 100 | Spark视频王家林第100课： 使用Spark Streaming+ Spark SQL + Kafka+FileSystem综合案例 |
| 101 | Spark视频王家林第101课： 使用Spark Streaming企业实际数据处理流水线完整声明周期 |
| 102 | Spark视频王家林第102课： 动手实战Spark Streaming自定义Receiver并进行调试和测试 |
| 103 | Spark视频王家林第103课： 动手实战联合使用Spark Streaming、Broadcast、Accumulator实现在线黑名单过滤和计数 |

**06. 大数据Spark电商广告点击综合案例**

大数据Spark电商广告点击综合案例

http://www.youku.com/playlist\_show/id\_27881496.html

|  |  |
| --- | --- |
| 104 | Spark视频王家林第104课： Spark Streaming电商广告点击综合案例需求分析和技术架构 |
| 105 | Spark视频王家林第105课： Spark Streaming电商广告点击综合案例在线点击统计实战 |
| 106 | Spark视频王家林第106课： Spark Streaming电商广告点击综合案例黑名单过滤实现 |
| 107 | Spark视频王家林第107课： Spark Streaming电商广告点击综合案例底层数据层的建模和编码实现（基于MySQL） |
| 108 | Spark视频王家林第108课： Spark Streaming电商广告点击综合案例动态黑名单过滤真正的实现代码 |
| 109 | Spark视频王家林第109课： Spark Streaming电商广告点击综合案例动态黑名单基于数据库MySQL的真正操作代码实战 |
| 110 | Spark视频王家林第110课： Spark Streaming电商广告点击综合案例通过updateStateByKey等实现广告点击流量的在线更新统计 |
| 111 | Spark视频王家林第111课： Spark Streaming电商广告点击综合案例在线实现每个Province点击排名Top5广告 |
| 112 | Spark视频王家林第112课： Spark Streaming电商广告点击综合案例实战实现广告点击Trend趋势计算实战 |
| 113 | Spark视频王家林第113课： Spark Streaming电商广告点击综合案例实战模拟点击数据的生成和数据表SQL建立 |
| 114 | Spark视频王家林第114课 ：SparkStreaming+Kafka+Spark SQL+TopN+Mysql电商广告点击综合案例实战视频 |

**07.大数据Spark Streaming大型Spark项目性能优化**

大型Spark项目性能优化系列(115-124)

http://www.tudou.com/plcover/nMCMdKHtj4I/?bid=03&pid=29&resourceId=341542230\_03\_0\_29

|  |  |
| --- | --- |
| 115 | Spark视频王家林第115课： 超大规模Spark性能优化本质思考 |
| 116 | Spark视频王家林第116课： Spark Streaming性能优化：如何在毫秒内处理处理大吞吐量的和数据波动比较大 的程序 |
| 117 | Spark视频王家林第117课： Spark Streaming性能优化：如何最大程度的确保Spark Cluster和Kafka链接的稳定性 |
| 118 | Spark视频王家林第118课： Spark Streaming性能优化：如何获得和持续使用足够的集群计算资源？ |
| 119 | Spark视频王家林第119课： Spark Streaming性能优化：如何在生产环境下应对流数据峰值巨变？ |
| 120 | Spark视频王家林第120课： Spark Streaming性能优化：如何在End-to-End生产环境下安全高效的把结果数据存入HBase中？ |
| 121 | Spark视频王家林第121课： Spark Streaming性能优化：通过摄像头图像处理案例来说明Spark流处理性能评估新方法及性能调优参数测试 |
| 122 | Spark视频王家林第122课： Spark Streaming性能优化：Spark Streaming处理分布式拒绝服务案例及性能优化 |
| 123 | Spark视频王家林第123课： Spark Streaming性能优化：通过Spark Streaming发现Botnet及性能优化 |
| 124 | Spark视频王家林第124课： Spark Streaming性能优化：通过Spark Streaming进行设备日志监控报警及性能优化思考 |

**08.大数据Spark Streaming源码疯狂解密系列**

Spark Streaming源码疯狂解密系列(125-134)

http://www.tudou.com/plcover/f0AK\_UVxwoQ/?bid=03&pid=29&resourceId=341542230\_03\_0\_29

|  |  |
| --- | --- |
| 125 | Spark视频王家林第125课： Spark Streaming反思和启示：一切皆是流式处理及Spark Streaming 架构和运行机制 |
| 126 | Spark视频王家林第126课： Spark Streaming源码经典解读系列之一：基于DStream的DStreamGraph源码内幕 |
| 127 | Spark视频王家林第127课： Spark Streaming源码经典解读系列之二：Spark Streaming生成RDD并执行Spark Job源码内幕解密 |
| 128 | Spark视频王家林第128课： Spark Streaming源码经典解读系列之三：JobScheduler工作内幕源码解密 |
| 129 | Spark视频王家林第129课： Spark Streaming源码经典解读系列之四：JobGenerator工作内幕源码解密 |
| 130 | Spark视频王家林第130课：  Spark Streaming源码经典解读系列之五：Receiver工作内幕源码解密 |
| 131 | Spark视频王家林第131课： Spark Streaming源码经典解读系列之六：ReceiverTracker工作内幕源码解密 |
| 132 | Spark视频王家林第132课： Spark Streaming源码经典解读系列之七：Executor容错工作内幕源码解密 |
| 133 | Spark视频王家林第133课： Spark Streaming源码经典解读系列之八：Driver容错工作内幕源码解密 |
| 134 | Spark视频王家林第134课： Spark Streaming“魔镜秘境”总结 |

**09 Spark面试宝典（数据倾斜、性能调优等）**

Spark面试宝典（数据倾斜、性能调优等）

[http://www.tudou.com/plcover/yBrSt2Vz8G8/](qq://txfile/)

|  |  |
| --- | --- |
| 135 | Spark视频王家林第135课： Spark面试经典系列之数据倾斜：数据倾斜之痛 |
| 136 | Spark视频王家林第136课： Spark面试经典系列之数据倾斜解决原理和方法总论 |
| 137 | Spark视频王家林第137课： Spark面试经典系列之数据倾斜解决之Map 端Reduce及问题思考 |
| 138 | Spark视频王家林第138课： Spark面试经典系列之数据倾斜解决之采样分而治之解决方案 |
| 139 | Spark视频王家林第139课： Spark面试经典系列之数据倾斜解决之对于两个RDD数据量都很大且倾斜的Key特别多如何解决？ |
| 140 | Spark视频王家林第140课： Spark面试经典系列之数据倾斜解决之并行度的深度使用 |
| 141 | Spark视频王家林第141课： Spark面试经典系列之数据倾斜解决方案的“银弹”是什么？ |
| 142 | Spark视频王家林第142课： Spark面试经典系列之Cache和Checkpoint |
| 143 | Spark视频王家林第143课： Spark面试经典系列之Reduce端OOM和shuffle file not found如何解决 |
| 144 | Spark视频王家林第144课： Spark面试经典系列之NULL值问题及序列化错误 |
| 145 | Spark视频王家林第145课： Spark面试经典系列之Yarn生产环境下资源不足问题和网络的经典问题详解 |
| 146 | Spark视频王家林第146课： Spark面试经典系列之Yarn Cluster生产环境下JVM的OOM和Stack Overflow问题及解决方案 |
| 147 | Spark视频王家林第147课： Spark面试经典系列之Shuffle的性能调优问题 |

DT大数据梦工厂王家林Spark源码大师之路之源码解析与阅读全集

http://www.tudou.com/plcover/hlFqt6k1xUc/

|  |  |
| --- | --- |
| 1 | Spark视频王家林Spark源码大师之路0001讲：Spark源码阅读环境及Spark-shell解密 |
| 2 | Spark视频王家林Spark源码大师之路0002讲：Spark-shell内幕解密 |
| 3 | Spark视频王家林Spark源码大师之路0003讲：Spark-shell REPL内幕解密 |
| 4 | Spark视频王家林Spark源码大师之路0004讲：史上最细致Spark集群启动脚本源码彻底解密 |
| 5 | Spark视频王家林Spark源码大师之路0005讲：Spark集群Master启动源码彻底解密 |
| 6 | Spark视频王家林Spark源码大师之路0006讲：Spark集群启动Worker源码彻底解密 |
| 7 | Spark视频王家林Spark源码大师之路0007讲：Spark集群HA分析和源码解析 |
| 8 | Spark视频王家林Spark源码大师之路0008讲：Spark下ZooKeeper内幕 |
| 9 | Spark视频王家林Spark源码大师之路0009讲：Spark下的ZooKeeper源码内幕 Final |
| 10 | Spark视频王家林Spark源码大师之路0010讲：SparkContext介绍及SparkEnv源码解析 |
| 11 | Spark视频王家林Spark源码大师之路0011讲：SparkEnv源码解析 |
| 12 | Spark视频王家林Spark源码大师之路0012讲：SparkUI与ListenerBus |
| 13 | Spark视频王家林Spark源码大师之路0013讲：ListenerBus源码实现内幕详解 |
| 14 | Spark视频王家林Spark源码大师之路0014讲：SparkUI源码实现内幕详解 |
| 15 | Spark视频王家林Spark源码大师之路0015讲：Spark下的Hadoop源码彻底解密 |
| 16 | Spark视频王家林Spark源码大师之路0016讲：MetricsSystem与ExecutorSource解密 |
| 17 | Spark视频王家林Spark源码大师之路0017讲：Spark UI的扩展定制 |
| 18 | Spark视频王家林Spark源码大师之路0018讲：TaskScheduler内幕彻底解密 |
| 19 | Spark视频王家林Spark源码大师之路0019讲：SchedulerBackend内幕彻底解密：启动的前世今生、初始化的过程、内部的AppClient、DriverEndpoint、ClientEndpoint等 |
| 20 | Spark视频王家林Spark源码大师之路0020讲：Spark ClientEndpoint内幕源码详解：创建过程、生命周期、工作机制、注册Application到集群全过程等 |
| 21 | Spark视频王家林Spark源码大师之路0021讲：Spark DriverEndpoint内幕源码详解 |
| 22 | Spark视频王家林Spark源码大师之路0022讲：Executor启动注册全流程彻底剖析：源码逐行解析CoarseGrainExecutorBackend在启动的时候会向driverUrl所代表的Endpoint进行注册，这个driverUrl代表的Endpoint到底实体是谁？ |
| 23 | Spark视频王家林Spark源码大师之路0023讲：Spark中的Pool彻底解析：FIFO与Fair彻底解密 OK |
| 24 | Spark视频王家林Spark源码大师之路0024讲：：Spark 1.6.1中RPC通信源码分析 |
| 25 | Spark视频王家林Spark源码大师之路0025讲：：Spark 1.6.1中延迟调度 |
| 26 | Spark视频王家林Spark源码大师之路0026讲：：Spark Job调度中的引擎机制和消息循环源码解析 |
| 27 | Spark视频王家林Spark源码大师之路0027讲：源码解密Spark中的算子Pipeline的合并和展开 |
| 28 | Spark视频王家林Spark源码大师之路0028讲：实战Eclipse调试Spark程序 |

**11 大数据JVM性能优化**

《DT大数据梦工厂王家林大数据JVM优化系列》

http://www.tudou.com/plcover/hzJp87qXtBA/

|  |  |
| --- | --- |
| 1 | 第1课：实战演示jvm内存四大类型问题：heap、stack、contant、directmemory等 |
| 2 | 第2课：全球详细图解jvm内存三大核心区域及其jvm内存案例实战剖析 |
| 3 | 第3课：图解jvm线程引擎和内存共享区的交互及此背景下的程序计数器内幕解密 |
| 4 | 第4课：10分钟内图解掌握全世界最易懂的jvm内存模型内幕 |
| 5 | 第5课：实战演示jvm三大性能调优参数：-xms -xmx -xss |
| 6 | 第6课：实战演示从oom推导出jvm Gc时候基于的内存结构：young Generation（eden、from、to）、old Generation、permanent Generation |
| 7 | 第7课：瞬间理解jvm Gc时的内存管理工作流程和频繁gc的原因 |
| 8 | 第8课：jvm的young Generation大小对gc性能的致命影响力内幕解密 |
| 9 | 第9课：jvm的gc时候核心参数详解：-xx：newratio、-xx：survivorratio、-xx：newsize、-xx：maxnewsize |
| 10 | 第10课：jvm的gc中关于对象的age问题以及jvm的内存分配策略彻底详解 |
| 11 | 第11课：jvm的gc日志young Generation下minor Gc每个字段彻底详解 |
| 12 | 第12课：jvm的gc日志full Gc每个字段彻底详解 |
| 13 | 第13课：java 8中的jvm的metaspace解密 |
| 14 | 第14课：java对象内存逃逸技术 |
| 15 | 第15课：GC时候的引用计数算法和根搜索算法 |
| 16 | 第16课：瞬间理解三种基本的GC算法基石 |
| 17 | 第17课：JVM垃圾回收分代收集算法 |
| 18 | 第18课：JVM垃圾回收器串行、并行、并发垃圾回收器概述 |
| 19 | 第19课：JVM中Serial收集器、ParNew收集器、Parallel收集器解析 |
| 20 | 第20课：JVM中CMS收集器解密 |
| 21 | 第21课：JVM中G1收集器解密 |
| 22 | 第22课：通过案例瞬间理解JVM中PSYoungGen、ParOldGen、MetaSpace |
| 23 | 第23课：使用MAT对Dump文件进行分析实战 |
| 24 | 第24课：使用MAT动态分析Spark应用程序初体验 |
| 25 | 第25课：基于MAT分析Shallow Heap和Retained Heap |
| 26 | 第26课：MAT中的Dominator Tree与Retained Set详解 |
| 27 | 第27课：MAT中的GC Root解析和具体类别分析 |
| 28 | 第28课：MAT中的内存泄露检测介绍 |

**12. 大数据Spark 机器学习系列**

DT大数据梦工厂王家林大数据Spark机器学习

http://www.tudou.com/plcover/zNefiPmqLW8/

|  |  |
| --- | --- |
| 1 | 30分钟彻底理解Spark核心API发展史：RDD、DataFrame、DataSet |
| 2 | 20分钟内解密spark第二代tungsten引擎测试数据和引擎实现内幕-2 |
| 3 | structured Streaming In Spark-3 |
| 4 | 深度解析structured Streaming-4 |
| 5 | 深入理解rdd、dataframe、dataset、structured Streaming-5 |
| 6 | Spark机器学习算法介绍-6 |
| 7 | Spark机器学习内幕剖析-7 |
| 8 | 王家林带您1分钟内彻底理解大数据机器学习-8 |

**13. DT大数据梦工厂周末班补充视频-Scala**

[DT大数据梦工厂大数据Spark周末](http://www.tudou.com/plcover/hI7hcy9GzSA/)

http://www.tudou.com/plcover/hI7hcy9GzSA/

|  |  |
| --- | --- |
| 1 | 用实例说明函数式编程到底是什么 |
| 2 | 动手编写和运行自己的第一个scala函数式编程的实例 |
| 3 | scala函数式编程中的泛型和匿名函数以及函数本质的思考 |
| 4 | scala函数的惊人的内幕价值：天然的可序列化解密与实战 |