**大数据（HadoopNoSQL等）的技术与实践 课程大纲（课时共2天）**

| **主题** | **授课内容** |
| --- | --- |
| 大数据时代关系数据库的挑战与应对 | 1. 现代数据管理技术综述2. 关系数据库技术的核心特征3.主流关系数据库的挑战4. 改进型关系数据库 |
| 大数据技术综述 | 1. 大数据概念澄清2. 大数据技术家族3. NoSQL技术综述3.1最早的NoSQL---BDB3.2 Hadoop之Hbase与Facebook之Cassandra3.3 MongoDB与CouchDB3.4Memcached与Redis3.5图形数据库Neo4j4. MapReduce5.关系数据库联邦6.海量分布式文件系统   7大数据技术理论基础       CAP，BASE，ACID |
| Hadoop实用教程 | 1. Hadoop技术概论1.1 Hadoop体系架构总论      1.2 HDFS-工作原理与架构      1.3 平民化的分布式计算MapReduce      1.4 MapReduce工作原理与架构      1.5 Hadoop数据仓库-Hive      1.6 Hadoop NoSQL数据库-HBase      1.7 工作流调度-Ooize      1.8 分布式协调系统--Zookeeper2. Hadoop部署2.1 Hadoop版本介绍与选择      2.2 Hadoop部署实践    2.3 Hadoop安装文件构成与配置体系    2.4 机器硬件建议配置    2.5 系统环境配置    2.6 基本参数配置与说明    2.7 进程分布规划与启动3. 分布式文件系统HDFS实用教程3.1 HDFS操作3.2 HDFS编程—文件读写3.3 HDFS数据压缩3.4 HDFS技术要点4.  MapReduce实用教程4.1 MapReduce原理与架构4.2 MapReduce编程方法4.3 MapReduce实用技术要点4.4 MapReduce排序与关联4.5 MapReduce工作流4.6 MapReduce调优5. MapReduce2.0-YARN5.1 YARN的原理5.2 YARN设计架构5.3 YARN工作流程5.4 YARN与MapReduce1.0比较6. MapReduce实例讲解6.1普通实例6.2 高级实例       6.3 MapReduce高级数据分析（时间允许时）7. Hadoop数据仓库Hive7.1 Hive编程      7.2 Hive环境部署与搭建      7.3 Hive工作机制7.4   Hive语法与实践8. 其它ZooKeeper，Sqoop，Chukwa，Avro…… |
| Hadoop技术分析 | 1.       Hadoop MapReduce技术解析6.1关于效率6.2关于扩展性6.3关于可靠性与可用性6.4关于与关系数据库6.5关于适用的数据类型6.6关于数据存储与管理2.       Hadoop与关系数据库2.1   MapReduce与关系数据库2.2   Hive与MPP关系数据库 |
| NoSQL实用教程 | 1.NoSQL理论基础---CAP与BASE深入分析2.NoSQL实用教程   2.1 HBase实用教程     2.1.1 HBase原理     2.1.2 HBase实用安装部署要点     2.1.3 HBase数据模型     2.1.4 HBase索引与关联的实现     2.1.5 HBase使用     2.1.6HBase性能调优     2.1.7 HBase高级设计教程---如何真正用好HBase     2.1.8 HBase与关系数据库结合3. NoSQL设计实例   3.1 HBase实现全属性查询   3.2 HBase实现时间序列数据管理   3.3 HBase与MapReduce结合示例4. Facebook Cassandra介绍5.MongoDB介绍6.图数据库Neo4J介绍 |
| NoSQL技术分析 | 1. NoSQL技术手段总结1.1 水平分割1.2 数据副本与读写一致性1.3   In-Memory架构1.4 MVCC1.5列存储1.6 COW2. NoSQL技术解析2.1   关于水平扩展性2.2关于模式自由3. NoSQL与关系数据库   3.1 理论原则分析   3.2 逻辑模型分析   3.3 物理模型分析   3.4 索引、事务与关联   3.5 使用场景定位   3.6 企业数据体系定位 |
| Spark教程 |    1. Spark组成与体系架构   2. Spark原理   3. Spark与Hadoop   4. Scala简介   5. Spark技术流程 |
| 超越Hadoop | 1.       Hadoop技术体系的不足与尴尬2.       新技术介绍3.       互联网技术体系介绍4.       数据管理技术发展趋势分析 |
| 大数据技术实践分享 | 1.       海量数据处理架构设计2.       大数据驱动与企业业务/运营3.       实践中的企业大数据分析技术流程3.1采集---各种方法的比较3.2存储---原始数据与业务数据提取3.3模型---Web分析指标体系3.4分析---大数据分析方法3.5 行动---个性化推荐4. 大数据与企业交易5. 大数据与企业交互5. 自已设计大数据技术体系 |
| 大数据与企业新一代数据体系建设 | 1.传统的以关系数据库为主的企业数据架构2.大数据时代的新一代企业数据逻辑架构2.1   数据分类2.2   数据分布2.3   数据流转2.4   数据集成2.5   数据交换2.6   数据分析2.7   应用展示3. 新一代企业数据技术架构3.1逻辑架构3.2技术方法3.3物理平台4.       新一代企业数据架构中Hadoop/NoSQL与关系数据库的相互配合5.       典型场景示例 |