东莞理工学院计算机科学与技术学院

实验报告

课程名称：云计算与大数据应用开发 学期：2024年春季

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验名称 | 大数据处理 | | | 实验序号 | 4 |
| 姓 名 | \*\*\* | 学 号 | \*\*\* | 班 级 | \*\*\* |
| 实验地点 | \*\*\* | 实验日期 | \*\*\* | 指导老师 | 丁烨 |
| 1. 实验目标    1. 了解MapReduce机制；    2. 熟悉MapReduce工具及相关代码；    3. 使用MapReduce工具，部署MapReduce相关代码。 2. 实验条件    1. 硬件条件：PC兼容机或Mac；    2. 软件条件：浏览器或终端。 3. 实验内容    1. 安装Apache Spark或 Apache Hadoop，或开启PaaS MapReduce服务；    2. 准备一份（或多份）长文本数据；    3. 撰写MapReduce代码；    4. 使用MapReduce对长文本数据进行词频统计（Word Count）。 4. 实验作业及分析    1. 实验过程   \*\*\* 请将详细实验过程的截图和相关说明填写在此处 \*\*\*   * 1. 实验结果   \*\*\* 请将实验结果的截图和相关说明填写在此处 \*\*\*   1. 实验总结   \*\*\* 请撰写一段200字左右的实验总结 \*\*\* | | | | | |