

# 问题及修改记录

## 1. 与文档格式相关

序号：1
来源：B 组、C 组
描述：文档没有目录
状态：接受
修改记录：在第三页正文前添加目录
具体信息： <div><div>▪ 目录</div><div><div>1 范围.....5</div><div><div>1.1 标识 .....5</div><div>1.2 系统概述.....5</div><div>1.3 文档概述.....5</div><div>1.4 目的 .....6</div><div><div>1.4.1 文档目的.....6</div><div>1.4.2 测试目的.....6</div></div><div>1.5 引用文档.....6</div></div><div>2 测试方法概述.....6</div><div>3 测试准备 .....8</div><div><div>3.1 软硬件环境的准备配置情况.....8</div><div>3.2 测试数据.....8</div><div>3.3 测试策略.....8</div><div>3.4 测试环境与软件运行环境的偏差.....8</div></div><div>4 测试用例.....9</div><div><div>4.1 功能性需求测试.....9</div><div>4.2 非功能性需求测试.....10</div></div><div>5 测试用例追踪.....13</div></div></div>

序号：2
来源：C 组
描述：参考资料没有需求规格说明书且格式不够规范。
状态：接受
修改记录：在“1.5 引用文档”一节添加需求规格说明书，并对格式加以更正。
具体信息： <div> <p>▪1.5 引用文档</p> <p>[1]许鹏,《Apache Spark 源码剖析》[M], 电子工业出版社, 2015-3</p> <p>[2]夏俊鸾. 程浩. 邵赛赛,《Spark 大数据处理技术》[M], 电子工业出版社, 2014. 12</p> <p>[3]《A-需求规格说明书-基于分布式 Apache Spark 的数据处理研究_v4.0 修订版》</p> </div>

序号:3
来源：C 组
描述：图表名应该比正文字体小一号
状态：接受
修改记录：更正“2 测试方法概述”中涉及的图的命名文字为宋体 5 号；更正“4 测试用例”中涉及的表格的命名文字为宋体 5 号。
具体信息：无

## 2、与文档内容描述相关

序号：1
来源：C 组
描述：测试目的那一部分现在写的是测试文档的目的，而应该写测试的目的
状态：部分调整，介绍编写测试文档的目的是有必要的，可以阐述我们编写该文档的初衷，并以此目标来规范文档编写过程。
修改记录：将“1.4 目的”中的内容改为“1.4.1 文档目的”和“1.4.2 测试目的”两部分内容。
<div>具体信息：</div> <div><div>▪1.4 目的↵</div><div><div>▪1.4.1 文档目的↵</div><div>结合《需求规格说明书》中涉及到的需求和目前 Apache Spark 系统的认识，对该系统的相关需求进行了初步测试计划的设计工作，其涵盖了测试准备、测试策略、测试用例设计、测试用例追踪设计等方面内容。测试人员根据此文档可高效完成对系统的测试工作。↵</div></div><div><div>▪1.4.2 测试目的↵</div><div>通过建立测试用例对系统进行测试，通过对产生的测试结果和系统的需求规格说明书进行比较，可以帮助我们判断系统是否满足了需求规格说明书中产生的功能性需求和非功能性需求，进而可以帮助我们发现系统中的问题并有机会对其加以改正。↵</div></div></div>

3. 与测试用例相关

序号：1									
来源：B 组、C 组									
描述：功能需求测试用例较少									
状态：接受									
修改记录：在“4.1 节功能性需求测试”一节添加两个功能性测试用例									
具体信息：增加 Spark 执行 SQL 数据处理的测试用例和 Spark Stream 流计算测试用例									
(2)测试需求——Spark 执行 SQL 数据处理									
测试用例名称	Spark 执行 SQL 数据处理	测试用例标识	TC002	测试需求项	功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	2.3 节 Sql 分析计算
简要描述	本测试 Spark 系统上运行 SQL 命令，针对不同格式的数据（如 JSON 数据集、数据库、Hive、RDD）执行 ETL 操作，然后进行增、删、改、查操作分析。								
前提和约束	Spark 集群配置完成，数据集准备妥当								
测试方法	黑盒测试								
测试过程描述									
序号	测试步骤					预期结果	评价准则		测试结论
1	使用 Spark Context 对象创建 SQLContext 对象： val sqlContext = new org.apache.spark.sql.SQLContext(sc)						实际结果与预期结果一致		无
2	使用导入语句的方式，隐式地将 RDD 转化成 DataFrame import sqlContext.implicits._						实际结果与预期结果一致		无
3	使用要进行分析的数据集文件创建一个与数据集结构相同的对象的 DataFrame						实际结果与预期结果一致		无
4	使用 DataFrame 将创建的 DataFrame 注册为一个表						实际结果与预期结果一致		无
5	添加增、删、改、查等操作（select、show、count）的实现函数						实际结果与预期结果一致		无

6	提交到集群执行，并监控执行状态					显示数据集相应的内容	实际结果与预期结果一致	无	
7	通过编程的方式指定数据集的模式，重新进行测试					显示数据集相应的内容	实际结果与预期结果一致	无	
8	对其他格式的数据集进行测试					显示数据集相应的内容	实际结果与预期结果一致	无	
备注									
测试人员：				监测人员：			测试时间：		
(3) 测试需求——Spark Stream 流计算测试									
测试用例名称	Spark Stream 流计算测试	测试用例标识	TC003	测试需求项	功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	2.3 节 Stream 流计算
简要描述	本测试是使用 Spark Stream 对实时数据进行高通量的流式处理，对输入的数据源（例如以 socket 连接作为数据源）进行某种操作，分析不同 batch interval 情况下的效率。								
前提和约束	Spark 集群配置完成，数据集准备妥当								
测试方法	黑盒测试								
测试过程描述									
序号	测试步骤					预期结果	评价准则	测试结论	
1	设置工作线程数目						实际结果与预期结果一致	无	
2	创建 StreamingContext 对象：val ssc = new StreamingContext(conf, Seconds(1))，设置 batch interval（例如 1s）指定处理数据的时间间隔——输入数据会按照 batch interval 分成一段一段的数据						实际结果与预期结果一致	无	
3	输入需要分析的数据源 InputDStream（可以是 Kafka、Flume、HDFS/S3、Kinesis 和 Twitter 等数据源）						实际结果与预期结果一致	无	
4	然后统计每一个 batch 中的 word 数目——Reduce					中间结果可以输出的外部设备	实际结果与预期结果一致	无	
5	统计（Map）并打印统计结果（可以进行					输出统计的结	实际结果与	无	

	时间统计)	果(和时间记录)	预期结果一致	
6	更改 batch interval 进行测试比较	显示数据集相应的内容	实际结果与预期结果一致	无
7	更改数据源进行测试比较			
备注				
测试人员:	监测人员:		测试时间:	

序号: 2
来源: C 组
描述: 测试准备的部分, 建议简要说明一下测试数据的具体信息, 以及为什么选择这些数据。
状态: 接受
修改记录: 在“3.2 测试数据”一节的添加对 Movielens 数据集及选择该数据集的理由的描述。
具体信息: 选择 Movielens 作为测试数据的原因主要体现在以下 4 个方面: <ul style="list-style-type: none"> <li>a. 数据量大</li> <li>b. 数据来源相对权威, 在一定程度上会减少测试结果的偏差</li> <li>c. 可验证机器学习的准确率</li> <li>d. 有实际应用价值</li> </ul>

序号: 3				
来源: B 组				
描述: “软硬件环境的准备配置情况” 相关表格信息应该补充完整				
状态: 接受				
修改记录: 对“3.1 软硬件环境的准备配置情况”中 Linux 平台下相关信息补充完整。				
具体信息: <div> <div>3.1 软硬件环境的准备配置情况</div> <table> <tr> <td>平台: Linux 平台</td></tr> <tr> <td>节点数: 2</td></tr> <tr> <td>cpu: i7-6700u i7-5960x</td></tr> <tr> <td>内存分配: 每台 16G</td></tr> </table> </div>	平台: Linux 平台	节点数: 2	cpu: i7-6700u i7-5960x	内存分配: 每台 16G
平台: Linux 平台				
节点数: 2				
cpu: i7-6700u i7-5960x				
内存分配: 每台 16G				

序号： 4

来源: B 组、C 组

描述: 需求和测试用例之间的追踪关系不明确 || 测试用例的追踪可以更丰富一些 || 建议与需求文档相对应

状态：适当调整

修改记录:

- 在“4 测试用例”中对表格 1-1, 表格 2-1~表格 2-4 中添加被测对象在需求规格说明书中对应的模块说明;
- 在“5 测试用例追踪”中添加测试需求项、测试用例项和需求规格说明书三者的追踪关系列表。

具体信息:

表格 1-1 测试需求——Spark 机器学习结果测试

测试用例名称	Spark 机器学习结果测试	测试用例标识	TC001	测试需求项	功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	第 2 章 2.3 节 MLib 机器学习
--------	----------------	--------	-------	-------	-------	--------	------	-----------------	-----------------------------

### 表格 1-2 测试需求——Spark 执行 SQL 数据处理

测试用例名称	Spark 执行 SQL 数据处理	测试用例标识	TC002	测试需求项	功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	第 2 章 2.3 节 Sql 分析计算
--------	-------------------	--------	-------	-------	-------	--------	------	-----------------	----------------------

表格 1-3 测试需求——Spark Stream 流计算测试

测试用例名称	Spark Stream 流计算测试	测试用例标识	TC003	测试需求项	功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	第 2 章 2.3 节 Stream 流计算
--------	--------------------	--------	-------	-------	-------	--------	------	-----------------	------------------------------

## 4.1 非功能性需求测试

表 2-1 测试需求——Spark 鲁棒性测试

测试用例名称	Spark 鲁棒性测试	测试用例标识	TC011	测试需求项	非功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说明书）	第 3 章 3.1 节 鲁棒性
--------	-------------	--------	-------	-------	--------	--------	------	-----------------	-----------------

表 2-2 测试需求——Spark 容错性测试

测试用例名称	Spark容错性测试	测试用例标识	TC012	测试需求项	非功能性测试	测试需求标识	TR01	测试用例追踪（需求规格说	第 3 章 3.2 节 容错性
--------	------------	--------	-------	-------	--------	--------	------	--------------	-----------------------

								说明书)	
表格 2-3 测试需求——Spark 安全性测试									
测试用例名称	Spark 安全性测试	测试用例标识	TC013	测试需求项	非功能性测试	测试需求标识	TR01	测试用例追踪 (需求规格说明书)	第 3 章 3.3 节 安全性
表格 2-4 测试需求——Spark 效率测试									
测试用例名称	Spark 效率测试	测试用例标识	TC014	测试需求项	非功能性测试	测试需求标识	TR01	测试用例追踪 (需求规格说明书)	第 3 章 3.4 节 效率

## ·5 测试用例追踪

测试需求项	测试需求项标识	需求规格说明对应模块	测试用例	测试用例标识
功能性需求	TR01	第 2 章 2.3 节 <u>MLib</u> 机器学习	Spark 机器学习结果测试	TC001
		第 2 章 2.3 节 <u>Sql</u> 分析计算	Spark 执行 SQL 数据处理	TC002
		第 2 章 2.3 节 Stream 流计算	Spark Stream 流计算测试	TC003
非功能性需求	TR11	第 3 章 3.1 节鲁棒性	Spark 鲁棒性测试	TC011
		第 3 章 3.2 节容错性	Spark 容错性测试	TC012
		第 3 章 3.3 节安全性	Spark 安全性测试	TC013
		第 3 章 3.4 节效率	Spark 效率测试	TC014

序号：5
来源：C 组
描述：鲁棒性测试和容错性测试的测试方法叫“性能测试”，是否合理？
状态：回应，性能测试是在功能正确的基础上考虑功能的响应时间、资源占用等。通过性能测试，会得到供参考的性能数据，可以验证系统是否达到用户提出的性能指标，同时发现系统中存在的性能瓶颈，起到优化系统的目的。
修改记录：无
具体信息：无



序号：6
来源：C 组
描述：负载测试与压力测试没有找到对应的测试用例
状态：回应，通过负载测试，确定在各种工作负载下系统的性能，目标是测试当负载逐渐增加时，系统各项性能指标的变化情况。压力测试是通过确定一个系统的瓶颈或者不能接收的性能点，来获得系统能提供的最大服务级别的测试。可以粗略的认为负载测试和压力测试是一种实现测试的方式，如在容错性测试或运行速度测试时，对测试环境不断加压以获取系统性能便是一种负载测试或压力测试的思想。
修改记录：无
具体信息：无

序号：7
来源：C 组
描述：测试完成的标准当中。结合某个测试阶段中单位时间查出错误的数量的曲线信息，确定是否进入下一测试阶段。”表意模糊，望详悉。
状态：修改并回应，这一完成标准属于理想测试状态下，测试周期较长的情况。通过记录测试阶段 A 在固定时间间隔下发现问题的数量来判断测试阶段 A 已没有发现新问题的潜力，此时可以选择进入下一测试阶段。
修改记录：在“3.3 测试策略”中对该判断标准进行阐述，并标明属于理想测试状态下的参考依据，适合于测试周期较长的情况。
<p>具体信息：</p> <p><b>•3.3 测试策略</b></p> <p>1、在测试设计上，可从正常情况、非正常情况、边界情况、非法情况、极端情况(强度测试)、性能测试几个角度进行。</p> <p>2、对于测试完成的标准，可以从以下两方面衡量：</p> <p>第一、执行了所有测试情况，但没有发现错误，则停止测试。</p> <p>第二、下述标准属于理想测试状态下，适合于测试周期较长的情况。具体来说：这一完成标准通过记录测试阶段 A 在固定时间间隔下发现问题的数量绘制曲线信息，来判断测试阶段 A 已没有发现新问题的潜力，此时可以选择进入下一测试阶段。</p>