**<<流式实时分布式计算框架**

**Spark的研究与应用>>**

**测试报告**



北京航空航天大学

2016-05-26

版本变更历史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 提交日期 | 主要编制人 | 审核人 | 版本说明 |
| V1.0 | 2016/5/26 | 阳艳红、武一杰、王铖成、于思民 | 于思民 | 初稿 |

目 录

[1. 测试概述 1](#_Toc452132733)

[1.1 标识 1](#_Toc452132734)

[1.2 文档说明 1](#_Toc452132735)

[1.3 编写目的 1](#_Toc452132736)

[1.4 测试范围 1](#_Toc452132737)

[1.5 参考资料 1](#_Toc452132738)

[2. 测试配置 2](#_Toc452132739)

[2.1 测试软硬件环境配置 2](#_Toc452132740)

[2.2 测试人员配置 2](#_Toc452132741)

[3. 测试报告 3](#_Toc452132742)

[3.1 测试结果报告 3](#_Toc452132743)

[3.1.1 数据编程模型测试 3](#_Toc452132744)

[3.1.1.1 创建RDD的测试 3](#_Toc452132745)

[3.1.1.2 转换RDD的测试 4](#_Toc452132746)

[3.1.1.3 RDD的Action操作测试 5](#_Toc452132747)

[3.1.1.4 RDD的缓存测试 5](#_Toc452132748)

[3.1.1.5 RDD检查点测试 6](#_Toc452132749)

[3.1.2 数据存储测试 7](#_Toc452132750)

[3.1.2.1 Master启动测试 7](#_Toc452132751)

[3.1.2.2 Slave注册测试 7](#_Toc452132752)

[3.1.2.3 Master向Slave发送控制命令测试 8](#_Toc452132753)

[3.1.2.4 Master从Slave获取状态测试 8](#_Toc452132754)

[3.1.2.5 Slave向Master发送状态更新测试 9](#_Toc452132755)

[3.1.3 集群部署测试 10](#_Toc452132756)

[3.1.3.1 Local模式测试 10](#_Toc452132757)

[3.1.3.2 Standalone模式测试 10](#_Toc452132758)

[3.1.3.3 Mesos模式测试 11](#_Toc452132759)

[3.1.3.4 YARN模式测试 12](#_Toc452132760)

[3.1.4 作业调度测试 13](#_Toc452132761)

[3.1.4.1 用户作业提交测试 13](#_Toc452132762)

[3.1.4.2 创建DAG测试 14](#_Toc452132763)

[3.1.4.3 Stage划分测试 14](#_Toc452132764)

[3.1.4.4 生成TaskSet测试 15](#_Toc452132765)

[3.1.4.5 Task调度测试 16](#_Toc452132766)

[3.1.5 非功能性测试 17](#_Toc452132767)

[3.1.5.1 运行速度测试 17](#_Toc452132768)

[3.1.5.2 容错性测试 18](#_Toc452132769)

[3.1.5.3 安全性测试 19](#_Toc452132770)

[3.2 测试问题报告 20](#_Toc452132771)

[3.3 测试结果分析报告 20](#_Toc452132772)

[3.3.1 需求覆盖分析 20](#_Toc452132773)

[3.3.2 缺陷分析 21](#_Toc452132774)

[4. 综合评价 21](#_Toc452132775)

[4.1 软件能力 21](#_Toc452132776)

[4.2 意见与建议 21](#_Toc452132777)

表目

[表1 测试软硬件环境配置表 2](#_Toc452132778)

[表2 测试人员配置表 2](#_Toc452132779)

[表3 创建RDD的测试用例 3](#_Toc452132780)

[表4 转换RDD的测试 4](#_Toc452132781)

[表5 RDD的Action操作测试 5](#_Toc452132782)

[表6 RDD的缓存测试 5](#_Toc452132783)

[表7 RDD的检查点测试 6](#_Toc452132784)

[表8 Master启动测试 7](#_Toc452132785)

[表9 Slave注册测试 7](#_Toc452132786)

[表10 Master向Slave发送控制命令测试 8](#_Toc452132787)

[表11 Master从Slave获取状态测试 8](#_Toc452132788)

[表12 Slave向Master发送状态更新测试 9](#_Toc452132789)

[表13 Local模式测试 10](#_Toc452132790)

[表14 Standalone模式测试 10](#_Toc452132791)

[表15 Mesos模式测试 11](#_Toc452132792)

[表16 YARN模式测试 12](#_Toc452132793)

[表17 用户作业提交测试 13](#_Toc452132794)

[表18 创建DAG测试 14](#_Toc452132795)

[表19 Stage划分测试 14](#_Toc452132796)

[表20 生成TaskSet测试 15](#_Toc452132797)

[表21 Task调度测试 16](#_Toc452132798)

[表22 运行速度测试 17](#_Toc452132799)

[表23 容错性测试 18](#_Toc452132800)

[表24 安全性测试 19](#_Toc452132801)

[表25 测试问题报告表 20](#_Toc452132802)

[表26 需求覆盖分析表 20](#_Toc452132803)

[表27 缺陷分析表 21](#_Toc452132804)

# 测试概述

## 标识

Spark版本号：Spark1.2.1

版本：V1.0

## 文档说明

本文档用于针对本组实验项目进行的相关软件测试进行报告，内容主要包括软件测试需求用例实际结果、软件测试中出现的Bug、对测试结果的分析和解决问题的方案。

## 编写目的

本测试报告为流式实时分布式框架Spark的测试报告，目的在于总结测试阶段的测试情况以及分析测试结果，描述系统是否符合需求文档，是否已经达到预期的功能目标和非功能目标，并对测试质量进行分析评估。

本测试报告提供给用户、测试人员、开发人员、项目管理者、其他管理人员和需要阅读本报告的技术人员阅读。

## 测试范围

测试工作主要依据《G-系统需求规格说明书-V3.2》文档和《G-测试需求说明书-V2.2》文档中的需求用例和测试需求用例进行系统测试，本测试主要采用黑盒测试的方法针对测试用例进行测试。

## 参考资料

1. Spark官方指导文档翻译：《spark-grogramming-guide-zh-ch.pdf》

2. Spark官方开发者文档翻译：《spark-developer-guide.pdf》

3.《G-系统需求规格说明书-V3.2》 2016.5.12

4.《G-测试需求说明书-V2.2》 2016.5.20

# 测试配置

## 测试软硬件环境配置

此次软件测试环境配置如表-1所示。

表1 测试软硬件环境配置表

|  |  |  |
| --- | --- | --- |
| 资源名称/类型 | 配置 | 备注 |
| 处理器（CPU） | Itel(R) Core(TM) i7-2670QM CPU @ 2.20GHZ | 只使用1个核心 |
| 运行内存（Memory） | 3072MB |  |
| 操作系统（OS） | CentOS7 |  |
| JDK | 1.8.0\_77 |  |
| Scala | 2.11.2 |  |
| Hadoop | 2.5.2 |  |
| Spark | 1.2.1 |  |

## 测试人员配置

此次软件测试人员配置如表-2所示。

表2 测试人员配置表

|  |  |  |
| --- | --- | --- |
| 成员 | 主要职责 | 测试用例编号 |
| 于思民 | 做数据编程模型相关测试需求，非功能性需求，汇总测试需求，编写测试需求规格说明书 | TC001、TC002、TC003、TC004、TC005、TC020、TC021、TC022 |
| 阳艳红 | 做集群部署模块相关测试需求，非功能性需求 | TC011、TC012、TC013、TC014、  TC020 |
| 武一杰 | 做数据存储模块相关测试需求，非功能性需求 | TC006、TC007、TC008、TC009、TC010、TC021 |
| 王铖成 | 做作业调度模块相关测试需求，非功能性需求 | TC015、TC016、TC017、TC018、TC019、TC020 |

# 测试报告

## 测试结果报告

### 数据编程模型测试

#### 创建RDD的测试

表3 创建RDD的测试用例

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC001 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地和远程的文件与目录、无效本地和远程的文件与目录、已有RDD对象 | | | | |
| **Procedure** | 1. 启动Master节点； 2. 启动Slave节点； 3. 执行spark-shell命令进入命令行； 4. 使用SparkContext对象从某一文件uri读取数据； 5. 输出RDD对相关属性及数据，检查结果正确性。 | | | | |
| **Pre-Results** | RDD对象各属性正确，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | file://noExist.txt(无效本地文件) | | Exception | | Normal，只有执行Action操作时才产生异常。 |
| <file://noDir>(无效本地目录) | | Exception | | Normal，只有执行Action操作时才产生异常。 |
| hdfs://192.168.108.211/noExist.txt(无效远程文件) | | Exception | | Normal，只有执行Action操作时才产生异常。 |
|  | hdfs://192.168.108.211/noDir(无效远程文件) | | Exception | | Normal，只有执行Action操作时才产生异常。 |
| **有效等价类** | <file://home/test.txt>(有效本地文件) | | Normal | | Normal |
| <file://home/dir>(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/test.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir(有效远程目录) | | Normal | | Normal |
| 已存在RDD对象 | | Normal | | Normal |

#### 转换RDD的测试

表4 转换RDD的测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC002 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地文件与目录、已有RDD对象 | | | | |
| **Procedure** | 1. 启动Master节点； 2. 启动Slave节点； 3. 执行spark-shell命令进入命令行； 4. 使用SparkContext对象从某一文件uri读取数据； 5. 对得到的RDD对象调用某种转换算子操作，并得到新的RDD对象； 6. 输出新转换的RDD对象的属性及数据，检查其结果的正确性 | | | | |
| **Pre-Results** | RDD对象各属性正确，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | <file://home/test.txt>(有效本地文件) | | Normal | | Normal |
| <file://home/dir>(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/test.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir(有效远程目录) | | Normal | | Normal |
| 已存在RDD对象 | | Normal | | Normal |

#### RDD的Action操作测试

表5 RDD的Action操作测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC003 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 已有RDD对象 | | | | |
| **Procedure** | 1. 启动Master节点； 2. 启动Slave节点； 3. 执行spark-shell命令进入命令行； 4. 使用SparkContext对象从某一文件uri读取数据； 5. 对得到的RDD对象调用某种转换算子操作，并得到新的RDD对象； 6. 对RDD执行Action算子操作； 7. 将操作结果输出，分析输出结果的正确性。 | | | | |
| **Pre-Results** | 输出结果符合Action算子操作。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 已存在RDD对象 | | Normal | | Normal |

#### RDD的缓存测试

表6 RDD的缓存测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC004 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 已有RDD对象 | | | | |
| **Procedure** | 1. 启动Master节点； 2. 启动Slave节点； 3. 执行spark-shell命令进入命令行； 4. 使用SparkContext对象从某一文件uri读取数据； 5. 将得到的RDD对象缓存起来； 6. 调用API，检测是否有该RDD对象的缓存； 7. 如果RDD尚未缓存，则缓存失败；如果RDD已缓存，则成功。 | | | | |
| **Pre-Results** | 内存中有该RDD对象的缓存。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 已存在RDD对象 | | Normal | | Normal |

#### RDD检查点测试

表7 RDD的检查点测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC005 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 已有RDD对象 | | | | |
| **Procedure** | 1. 启动Master节点； 2. 启动Slave节点； 3. 执行spark-shell命令进入命令行； 4. 使用SparkContext对象设置检查点目录； 5. 使用SparkContext对象从某一文件uri读取数据； 6. 为得到的RDD对象创建检查点； 7. 调用API，检测是否有该RDD对象的检查点，及检查点的相关属性（时间、大小等）的正确性。 | | | | |
| **Pre-Results** | 检查点的相关属性与该RDD对象相一致 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 已存在RDD对象 | | Normal | | Normal |

### 数据存储测试

#### Master启动测试

表8 Master启动测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC006 | **Tester** | | 武一杰 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure** | 1. 设置Master节点配置信息； 2. 发送启动Master命令； 3. 读取Master启动信息。 | | | | |
| **Pre-Results** | Master启动信息显示Master节点正常。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | Normal | | Normal |

#### Slave注册测试

表9 Slave注册测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC007 | **Tester** | | 武一杰 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure** | 1. 读取Master节点信息； 2. 启动Master节点； 3. 读取Slave节点信息； 4. Slave向Master发送状态更新信息,进行注册； 5. Master获取Slave节点的启动信息。 | | | | |
| **Pre-Results** | Master读取到Slave节点的注册信息（启动信息） | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | Normal | | Normal |

#### Master向Slave发送控制命令测试

表10 Master向Slave发送控制命令测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC008 | **Tester** | | 武一杰 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure** | 1. 读取Master节点信息； 2. 启动Master节点； 3. 读取Slave节点信息； 4. Slave向Master的特定端口发送状态更新信息,进行注册； 5. Master获取Slave节点的启动信息； 6. Master向Slave发送控制命令，并监听Slave的状态更新。 | | | | |
| **Pre-Results** | Slave的状态更新结果与控制命令相一致 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | Normal | | Normal |

#### Master从Slave获取状态测试

表11 Master从Slave获取状态测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC009 | **Tester** | | 武一杰 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure** | 1. 读取Master节点信息； 2. 启动Master节点； 3. 读取Slave节点信息； 4. Slave向Master的特定端口发送状态更新信息,进行注册； 5. Master获取Slave节点的启动信息，并获取Slave状态。 | | | | |
| **Pre-Results** | 获取Slave的状态，与Salve的状态一致。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | Normal | | Normal |

#### Slave向Master发送状态更新测试

表12 Slave向Master发送状态更新测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC010 | **Tester** | | 武一杰 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure** | 1. 读取Master节点信息； 2. 启动Master节点； 3. 读取Slave节点信息； 4. Slave向Master的特定端口发送状态更新信息,进行注册； 5. Master监听特定端口，获取状态更新信息。 | | | | |
| **Pre-Results** | Master获取到状态更新信息 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | Normal | | Normal |

### 集群部署测试

#### Local模式测试

表13 Local模式测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC011 | **Tester** | | 阳艳红 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure** | 1. 启动Master节点；  2. 启动Slave节点；  3.在用户程序中应用的部署方设置为lcoal，即将SparkConf.setMaster(“local”)；  4. 提交用户作业运行；  5. 执行JConsole可执行程序，查看已经启动的java 线程 ；  6. 查看作业运行结果 | | | | |
| **Pre-Results** | Java线程显示本地运行，且程序结果正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | /home/yangel/Scripts/SougouLogAnalysis/data | | Normal | | Normal |

#### Standalone模式测试

表14 Standalone模式测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC012 | **Tester** | | 阳艳红 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure** | 1. 启动Master节点；  2. 启动Slave节点；  3.在用户程序中应用的部署方设置为standalone(默认), 即将SparkConf.setMaster(“spark://master:7077”)；  4. 提交用户作业运行；  5. 执行JPS命令；  6. 查看作业运行结果 | | | | |
| **Pre-Results** | JPS命令显示各节点进程正常运行，且程序结果正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | Exception | |  |
| **有效等价类** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | Normal | |  |

#### Mesos模式测试

表15 Mesos模式测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC013 | **Tester** | | 阳艳红 | |
| **Time** | 2016/05/26 | **Status** | | NOT PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure** | 1. 启动Master节点；  2. 启动Slave节点；  3.在用户程序中应用的部署方设置为mesos，即将SparkConf.setMaster(“mesos://HOST:5050”)；  4. 提交用户作业运行；  5. 执行JPS命令；  6. 查看作业运行结果 | | | | |
| **Pre-Results** | JPS命令显示各节点进程正常运行，且程序结果正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** |  | | Exception | |  |
| **有效等价类** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | Normal | |  |

#### YARN模式测试

表16 YARN模式测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC014 | **Tester** | | 阳艳红 | |
| **Time** | 2016/05/26 | **Status** | | NOT PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure** | 1.设置Spark的配置信息，设置Hadoop的配置信息；  2.启动Hadoop集群；  3.启动Master节点；  4.启动Slave节点；  5.在用户程序中应用的部署方设置为mesos，即将SparkConf. setMaster(“yarn-client”)或 SparkConf.setMaster(“yarn-cluster”)；  6. 提交用户作业运行；  7. 执行JPS命令；  8. 查看作业运行结果 | | | | |
| **Pre-Results** | 集群部署已经完成 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | 无 | | 无 | | 无 |

### 作业调度测试

#### 用户作业提交测试

表17 用户作业提交测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC015 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地和远程的文件与目录、无效本地和远程的文件与目录、已有RDD对象 | | | | |
| **Procedure** | **1.用户向Spark集群提交作业**  **2.作业输出到DAGScheduler**  **3.生成Job**  **4.DAGScheduler将作业输出到事件处理器**  **5.从事件处理器获取事件处理结果** | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD,RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | [file://home/sched/jobsubmit.txt](file://home/test.txt)(有效本地文件) | | Normal | | Normal |
| <file://home/dir>/sched/(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/sched/jobsubmit.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir/sched/(有效远程目录) | | Normal | | Normal |
| 已存在RDD对象 | | Normal | | Normal |

#### 创建DAG测试

表18 创建DAG测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC016 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** |  | | | | |
| **Procedure** | **1.获取计算任务**  **2.读取RDD信息**  **3.将RDD转化成DAG**  **4.查看生成的DAG** | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | [file://home/sched/dagcrt.txt](file://home/test.txt)(有效本地文件) | | Normal | | Normal |
| <file://home/dir>/sched/(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/sched/dagcrt.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir/sched/(有效远程目录) | | Normal | | Normal |
| 已存在RDD对象 | | Normal | | Normal |

#### Stage划分测试

表19 Stage划分测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC017 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地和远程的文件与目录、无效本地和远程的文件与目录、已有RDD对象 | | | | |
| **Procedure** | **1.获取触发Action操作的RDD1**  **2.获取该RDD所依赖的父RDD2**  **3.将与RDD2与依赖关系的分为同一STAGE**  **4.将生成的STAGE转化为DAG**  **5.查看结果** | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | [file://home/sched/stage.txt](file://home/test.txt)(有效本地文件) | | Normal | | Normal |
| <file://home/dir>/sched/(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/sched/stag.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir/sched/(有效远程目录) | | Normal | | Normal |
| 已存在RDD对象 | | Normal | | Normal |

#### 生成TaskSet测试

表20 生成TaskSet测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC018 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地和远程的文件与目录、无效本地和远程的文件与目录、已有RDD对象 | | | | |
| **Procedure** | **1.判断需要进行计算Partion**  **2.依次为每个Partition创建对应的Task**  **3.将这些Task封装成对应的Task**  **4.向TaskScheduler提交任务**  **5.查看提交结果** | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | [file://home/sched/ts.txt](file://home/test.txt)(有效本地文件) | | Normal | | Normal |
| <file://home/dir>/sched/(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/sched/ts.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir/sched/(有效远程目录) | | Normal | | Normal |

#### Task调度测试

表21 Task调度测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC019 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 有效本地和远程的文件与目录、无效本地和远程的文件与目录、已有RDD对象 | | | | |
| **Procedure** | **1.提交任务**  **2.分配资源**  **3.发送到执行模块**  **4.执行作业**  **5.查看结果** | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **无效等价类** | 无 | | 无 | | 无 |
| **有效等价类** | [file://home/sched/sched.txt](file://home/test.txt)(有效本地文件) | | Normal | | Normal |
| <file://home/dir>/sched/(有效本地目录) | | Normal | | Normal |
| hdfs://192.168.108.211/sched/sched.txt(有效远程文件) | | Normal | | Normal |
| hdfs://192.168.108.211/dir/sched/(有效远程目录) | | Normal | | Normal |

### 非功能性测试

#### 运行速度测试

表22 运行速度测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC019 | **Tester** | | 王铖成 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure** | 1.设置Spark的配置信息，设置Hadoop的配置信息；  2.启动Hadoop集群；  3.启动Master节点；  4.启动Slave节点；  5.提交Job-1（SougouLogAnalysisApp-assembly-1.0.jar）至Spark集群运行；  6.Job-1结束后，提及Job-2（SougouLogAnalysisAppV1-assembly-1.0.jar）至Spark集群运行；  7.Job-2运行结束后，提交Job-3（SougouLogAnalysisAppV2-assembly-1.0.jar）至Spark集群运行；  8.Job-3运行结束后，提交Job-4（SougouLogAnalysisAppV3-assembly-1.0.jar）至Spark集群运行；  9.Job-4运行结束后，提交Job-5（SougouLogAnalysisAppV5-assembly-1.0.jar）至Spark集群运行；  10.提交Job-6（）至Hadoop集群运行；  10.所有作业运行结束后，统计各作业运行时间。 | | | | |
| **Pre-Results** | 5Min | | | | |
| **程序编号** | **用例数据** | | **期望结果** | | **实际结果** |
| **Job-1** | 同Data | | 5Min | | 2.8Min |
| **Job-2** | 同Data | | 5Min | | 53Sec |
| **Job-3** | 同Data | | 5Min | | 2.8Min |
| **Job-4** | 同Data | | 5Min | | 58Sec |
| **Job-5** | 同Data | | 5Min | | 1.2Min |
| **Job-6** | 同Data | | 5Min | | 19.2Min |

#### 容错性测试

表23 容错性测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC019 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | /home/yangel/Scripts/SougouLogAnalysis/data（搜狗用户月度搜索日志） | | | | |
| **Procedure1** | 1. 设置Spark的配置信息，并将其中某项配置设置为异常值; 2. 启动Master、Slavle节点; 3. 观察各节点是否正常启动; | | | | |
| **Procedure2** | 1.设置Spark的配置信息，将集群部署模式配置为YARN（与YARN适配的相关配置信息）；  2.启动Hadoop集群；  3.启动Master节点；  4.启动Slave节点；  5.提交Spark提供的例子作业；  6.将Hadoop集群的某个节点关闭。  7.查看作业执行结果 | | | | |
| **Procedure3** | 1.设置Spark的配置信息，将集群部署模式配置为YARN（与YARN适配的相关配置信息）；  2.启动Hadoop集群；  3.启动Master节点；  4.启动Slave节点；  5.提价Spark提供的例子作业；  6.将Hadoop集群中执行任务的节点关闭；  6.查看作业执行结果 | | | | |
| **Pre-Results** | 环境搭建正确,集群部署操作成功,需要初步的RDD，RDD数据正确。 | | | | |
| **输入等价类** | **用例数据** | | **期望结果** | | **实际结果** |
| **Procedure1** | 同data | | Exception | | Exception |
| **Procedure2** | 同data | | Normal | | Normal |
| **Procedure3** | 同data | | Normal | | Normal |

#### 安全性测试

表24 安全性测试

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Program** | Spark研究与应用 | **Version** | | V1.0 | |
| **Test Case** | TC019 | **Tester** | | 于思民 | |
| **Time** | 2016/05/26 | **Status** | | PASS | |
| **Data** | 无 | | | | |
| **Procedure1** | 1. 设置Spark的配置信息，将spark.ui.filters参数设置为用户A，启用javax.servlet.filters进行UI安全设置。 2. 启动Master、Slavle节点   用户B登录到Spark，并查看用户A的UI界面 | | | | |
| **Procedure2** | 1. 设置Spark的配置信息； 2. 启动Master、Slave节点；   用户A提交用户B某目录下的作业（A用户与B用户不属于相同的组）到Spark集群 | | | | |
| **Pre-Results** | 1. 用户B无法获取UI界面数据。 2. Spark拒绝作业的执行。 | | | | |
| **Procedure** | **用例数据** | | **期望结果** | | **实际结果** |
| **Procedure1** | 无 | | Exception | | Exception |
| **Procedure2** | 无 | | Exception | | Exception |

## 测试问题报告

表25 测试问题报告表

|  |  |  |  |
| --- | --- | --- | --- |
| 测试用例 | 问题描述 | 解决状态 | 备注 |
| TC001 | RDD创建过程并不检测数据URL的合法性 | 未解决 |  |
| TC004 | RDD的缓存过程出现重复操作 | 未解决 |  |
| TC006-TC010 | 缺少Master与Slave节点间通信实时检测模块。 | 未解决 |  |
| TC011-TC014 | 在Spark1.2.1中，集群虽然有多种部署模式，但是缺少可视化，使得很难检测部署模式的运行时特征。 | 未解决（Spark1.4版本已解决） |  |
| TC015 | 再将作业运行结果持久化到本地时，输出目录有很多小的结果文件。 | 未解决 |  |
| TC015 | Spark读取本地或远程文件时，无法指定文件的编码格式，使得读取含有中文的数据时，可能出现乱码问题。 | 已解决（未彻底解决） |  |

## 测试结果分析报告

### 需求覆盖分析

表26 需求覆盖分析表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 需求/功能 | 用例个数 | 执行个数 | 未执行 | 未/漏测分析原因 |
| 数据编程模型 | 5 | 5 | 0 | 无 |
| 数据存储 | 5 | 5 | 0 | 无 |
| 集群部署 | 4 | 2 | 2 | 无 |
| 作业调度 | 5 | 5 | 0 | 无 |
| 非功能性需求 | 3 | 3 | 0 | 无 |

### 缺陷分析

表27 缺陷分析表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 严重级别  需求 | A-严重影响系统运行的错误 | B-功能缺陷，影响系统运行 | C-不影响运行但必须修改 | D-合理的建议 | 总数 |
| 数据编程模型 | 0 | 0 | 0 | 2 | 2 |
| 数据存储 | 0 | 0 | 0 | 1 | 1 |
| 集群部署 | 0 | 0 | 0 | 1 | 1 |
| 作业调度 | 0 | 0 | 0 | 1 | 1 |
| 总数 | 0 | 0 | 0 | 5 | 6 |

# 综合评价

## 软件能力

Spark集群的分布式能力与性能已经得到业界的普遍认同，但是就细枝末节问题，Spark还有待完善与提高。

## 意见与建议

1. 增加可视化模块，对系统运行时、作业运行时进行可视化管理；
2. 增加实时检测模块，可针对系统的流量、通信等数据进行检测。