

<<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
1.1 标识.....	1
1.2 文档说明.....	1
1.3 编写目的.....	1
1.4 测试范围.....	1
1.5 参考资料.....	1
2. 测试配置.....	2
2.1 测试软硬件环境配置.....	2
2.2 测试人员配置.....	2
3. 测试报告.....	3
3.1 测试结果报告.....	3
3.1.1 数据编程模型测试.....	3
3.1.1.1 创建 RDD 的测试.....	3
3.1.1.2 转换 RDD 的测试.....	4
3.1.1.3 RDD 的 Action 操作测试.....	5
3.1.1.4 RDD 的缓存测试.....	5
3.1.1.5 RDD 检查点测试.....	6
3.1.2 数据存储测试.....	7
3.1.2.1 Master 启动测试.....	7
3.1.2.2 Slave 注册测试.....	7
3.1.2.3 Master 向 Slave 发送控制命令测试.....	8
3.1.2.4 Master 从 Slave 获取状态测试.....	8
3.1.2.5 Slave 向 Master 发送状态更新测试.....	9
3.1.3 集群部署测试.....	10
3.1.3.1 Local 模式测试.....	10
3.1.3.2 Standalone 模式测试.....	10

3.1.3.3 Mesos 模式测试	11
3.1.3.4 YARN 模式测试	12
3.1.4 作业调度测试	13
3.1.4.1 用户作业提交测试	13
3.1.4.2 创建 DAG 测试	14
3.1.4.3 Stage 划分测试	14
3.1.4.4 生成 TaskSet 测试	15
3.1.4.5 Task 调度测试	16
3.1.5 非功能性测试	17
3.1.5.1 运行速度测试	17
3.1.5.2 容错性测试	18
3.1.5.3 安全性测试	19
3.2 测试问题报告	20
3.3 测试结果分析报告	20
3.3.1 需求覆盖分析	20
3.3.2 缺陷分析	21
4. 综合评价	21
4.1 软件能力	21
4.2 意见与建议	21

表目

表 1 测试软硬件环境配置表	2
表 2 测试人员配置表	2
表 3 创建 RDD 的测试用例	3
表 4 转换 RDD 的测试	4
表 5 RDD 的 ACTION 操作测试	5
表 6 RDD 的缓存测试	5
表 7 RDD 的检查点测试	6
表 8 MASTER 启动测试	7
表 9 SLAVE 注册测试	7
表 10 MASTER 向 SLAVE 发送控制命令测试	8
表 11 MASTER 从 SLAVE 获取状态测试	8
表 12 SLAVE 向 MASTER 发送状态更新测试	9
表 13 LOCAL 模式测试	10
表 14 STANDALONE 模式测试	10
表 15 MESOS 模式测试	11
表 16 YARN 模式测试	12
表 17 用户作业提交测试	13
表 18 创建 DAG 测试	14
表 19 STAGE 划分测试	14
表 20 生成 TASKSET 测试	15
表 21 TASK 调度测试	16
表 22 运行速度测试	17
表 23 容错性测试	18
表 24 安全性测试	19
表 25 测试问题报告表	20
表 26 需求覆盖分析表	20

表 27 缺陷分析表	21
------------------	----

1. 测试概述

1.1 标识

Spark 版本号: Spark1.2.1

版本: V1.0

1.2 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 and 解决问题的方案。

1.3 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。

1.4 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。

1.5 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

2. 测试配置

2.1 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	

2.2 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

3. 测试报告

3.1 测试结果报告

3.1.1 数据编程模型测试

3.1.1.1 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

3.1.1.2 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

3.1.1.3 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

3.1.1.4 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

3.1.1.5 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

3.1.2 数据存储测试

3.1.2.1 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

3.1.2.2 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

3.1.2.3 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

3.1.2.4 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

3.1.2.5 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

3.1.3 集群部署测试

3.1.3.1 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 local，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

3.1.3.2 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	

3.1.3.3 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	

3.1.3.4 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

3.1.4 作业调度测试

3.1.4.1 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	<u>file://home/sched/jobsubmit.txt</u> (有效本地文件)	Normal	Normal
	<u>file://home/dir/sched/</u> (有效本地目录)	Normal	Normal
	<u>hdfs://192.168.108.211/sched/jobsubmit.txt</u> (有效远程文件)	Normal	Normal
	<u>hdfs://192.168.108.211/dir/sched/</u> (有效远程目录)	Normal	Normal
	已存在 RDD 对象	Normal	Normal

3. 1. 4. 2 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	<u>file://home/sched/dagcrt.txt</u> (有效本地文件)	Normal	Normal
	<u>file://home/dir/sched/</u> (有效本地目录)	Normal	Normal
	<u>hdfs://192.168.108.211/sched/_dagcrt.txt</u> (有效远程文件)	Normal	Normal
	<u>hdfs://192.168.108.211/dir/sched/</u> (有效远程目录)	Normal	Normal
	已存在 RDD 对象	Normal	Normal

3. 1. 4. 3 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 (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

3.1.4.4 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	<u>file://home/sched/ts.txt</u> (有效本地文件)	Normal	Normal
	<u>file://home/dir/sched/</u> (有效本地目录)	Normal	Normal
	<u>hdfs://192.168.108.211/sched/ts.txt</u> (有效远程文件)	Normal	Normal
	<u>hdfs://192.168.108.211/dir/sched/</u> (有效远程目录)	Normal	Normal

3.1.4.5 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	<u>file://home/sched/sched.txt</u> (有效本地文件)	Normal	Normal
	<u>file://home/dir/sched/</u> (有效本地目录)	Normal	Normal
	<u>hdfs://192.168.108.211/sched/sched.txt</u> (有效远程文件)	Normal	Normal
	<u>hdfs://192.168.108.211/dir/s</u>	Normal	Normal

	<u>ched/</u> (有效远程目录)		
--	-----------------------	--	--

3.1.5 非功能性测试

3.1.5.1 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

3.1.5.2 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、Slave 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. 提交 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

3.1.5.3 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、Slave 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

3.2 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)	

3.3 测试结果分析报告

3.3.1 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

3.3.2 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

4. 综合评价

4.1 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。

4.2 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。