

华东师范大学数据科学与工程学院实验报告

课程名称： 分布式模型与编程 年级： 2016 级 上机实践成绩：
指导教师： 徐辰 姓名： 张宏伟
上机实践名称： Giraph 部署与编程 学号： 10165101180 上机实践日期：
上机实践编号： 15 组号： 上机实践时间：

一、实验目的

熟悉 Giraph API

二、实验任务

单机部署 Giraph，运行 Giraph 最短路径程序

三、使用环境

Hadoop2.5.1 + giraph1.3.0

四、实验过程

1. 创建/tmp/tiny_graph.txt，输入以下作为输入

```
[0,0,[[1,1],[3,3]]]
[1,0,[[0,1],[2,2],[3,1]]]
[2,0,[[1,2],[4,4]]]
[3,0,[[0,3],[1,1],[4,4]]]
[4,0,[[3,4],[2,4]]]
```

2. 将该文件上传到 hdfs 中 `$ /usr/local/hadoop/bin/hadoop dfs -put /tmp/tiny_graph.txt /tiny_graph.txt`

3. 提交 jar 包运行

```
$ /usr/local/hadoop/bin/hadoop jar /usr/local/giraph/giraph-examples/target/giraph-examples-1.3.0-SNAPSHOT-for-hadoop-2.5.1-jar-with-dependencies.jar
org.apache.giraph.GiraphRunner
org.apache.giraph.examples.SimpleShortestPathsComputation -vif
org.apache.giraph.io.formats.JsonLongDoubleFloatDoubleVertexInputFormat -vip
/tiny_graph.txt -vof org.apache.giraph.io.formats.IdWithValueTextOutputFormat -op
/shortestpaths -w 1 -ca giraph.SplitMasterWorker=false
```

输入为 hdfs 中的文件/tiny_graph.txt 输出也为 hdfs 的/shortestpath，执行后 hdfs 目录如图

```
Found 26 items
drwxr-xr-x - root root 4096 2019-01-03 16:33 /bin
drwxr-xr-x - root root 4096 2018-04-08 21:07 /boot
drwxr-xr-x - root root 3740 2019-01-03 16:01 /dev
drwxr-xr-x - root root 4096 2019-01-07 22:12 /etc
drwxr-xr-x - root root 4096 2016-04-13 04:14 /home
-rw-r--r-- 1 root root 34389550 2018-04-08 21:07 /initrd.img
-rw-r--r-- 1 root root 34381530 2018-03-26 12:05 /initrd.img.old
drwxr-xr-x - root root 4096 2019-01-03 22:57 /lib
drwxr-xr-x - root root 4096 2018-03-26 12:05 /lib64
drwx----- - root root 16384 2018-03-26 12:00 /lost+found
drwxr-xr-x - root root 4096 2018-03-26 12:00 /media
drwxr-xr-x - root root 4096 2017-08-01 19:16 /mnt
drwxr-xr-x - root root 4096 2017-08-01 19:16 /opt
dr-xr-xr-x - root root 0 2019-01-03 16:01 /proc
drwx----- - root root 4096 2019-01-13 20:52 /root
drwxr-xr-x - root root 640 2019-01-13 20:53 /run
drwxr-xr-x - root root 12288 2018-04-08 21:07 /sbin
drwxr-xr-x - root root 4096 2019-01-13 21:04 /shortestpaths
drwxr-xr-x - root root 4096 2017-08-01 19:16 /srv
dr-xr-xr-x - root root 0 2019-01-13 20:57 /sys
-rw-r--r-- 1 root root 112 2019-01-13 20:54 /tiny_graph.txt
drwxrwxrwt - root root 4096 2019-01-13 21:05 /tmp
drwxr-xr-x - root root 4096 2018-03-26 12:00 /usr
drwxr-xr-x - root root 4096 2018-03-26 12:00 /var
-rw----- 1 root root 7145008 2018-03-13 21:44 /vmlinuz
-rw----- 1 root root 7141424 2018-02-13 08:57 /vmlinuz.old
root@iZuf6ari54gn0wytwb125Z: /usr/local/hadoop-2.5.1#
```

4. 查看运行的输出结果

```
$ /usr/local/hadoop/bin/hadoop dfs -cat /shortestpaths/p* | less
```

```
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.
```

```
0      1.0  
1      0.0  
2      2.0  
3      1.0  
4      5.0
```

```
(END)
```

五、总结

对比非面向图分布式运算的系统，giraph 的运算速度快不少。