



专注于商业智能BI和大数据的垂直社区平台

# Spark初级课程

讲师：平常心

# 目录



基于HDP的Apache Spark集群



基于CDH的Apache Spark集群


# 基于HDP的Apache Spark集群搭建

## 安装软件

scala-2.10.4.zip

spark-1.6.3-bin-hadoop2.6.tgz

← → ↻ ⓘ 192.168.91.135:8090

**Spark Master at spark://hdp1:7077**

URL: spark://hdp1:7077  
REST URL: spark://hdp1:6066 (*cluster mode*)  
Alive Workers: 3  
Cores in use: 3 Total, 0 Used  
Memory in use: 16.1 GB Total, 0.0 B Used  
Applications: 0 Running, 0 Completed  
Drivers: 0 Running, 0 Completed  
Status: ALIVE

### Workers

Worker Id	Address	State	Cores	Memory
<a href="#">worker-20161119043138-192.168.91.135-56928</a>	192.168.91.135:56928	ALIVE	1 (0 Used)	6.7 GB (0.0 B Used)
<a href="#">worker-20161119043142-192.168.91.136-35586</a>	192.168.91.136:35586	ALIVE	1 (0 Used)	4.7 GB (0.0 B Used)
<a href="#">worker-20161119043144-192.168.91.137-36309</a>	192.168.91.137:36309	ALIVE	1 (0 Used)	4.7 GB (0.0 B Used)

### Running Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------

### Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------

# 基于CDH的Apache Spark集群搭建


## 安装软件

scala-2.10.4.zip

spark-1.6.3-bin-hadoop2.6.tgz

← → ↺

cdh1:8080

 1.6.3

Spark Master at spark://cdh1:7077

URL: spark://cdh1:7077

REST URL: spark://cdh1:6066 (cluster mode)

Alive Workers: 3

Cores in use: 3 Total, 3 Used

Memory in use: 16.1 GB Total, 3.0 GB Used

Applications: 1 Running, 0 Completed

Drivers: 0 Running, 0 Completed

Status: ALIVE

Workers

Worker Id	Address	State	Cores	Memory
<a href="#">worker-20161119072404-192.168.91.141-57996</a>	192.168.91.141:57996	ALIVE	1 (1 Used)	4.7 GB (1024.0 MB Used)
<a href="#">worker-20161119072405-192.168.91.140-46201</a>	192.168.91.140:46201	ALIVE	1 (1 Used)	4.7 GB (1024.0 MB Used)
<a href="#">worker-20161119072406-192.168.91.142-44454</a>	192.168.91.142:44454	ALIVE	1 (1 Used)	6.7 GB (1024.0 MB Used)

Running Applications

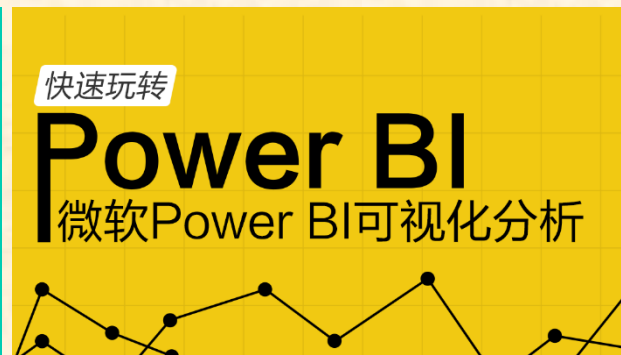
Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
<a href="#">app-20161119073906-0000</a>	(kill) <a href="#">SparkSQL::192.168.91.140</a>	3	1024.0 MB	2016/11/19 07:39:06	root	RUNNING	34 s

Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------



更多商业智能BI和大数据精品视频尽在 [www.hellobi.com](http://www.hellobi.com)



BI、商业智能  
数据挖掘 大数据  
数据分析  
R Python  
机器学习  
Tableau  
QLIKVIEW  
Hive Hadoop  
BIWORK  
BAO胖子 seng  
曹浩 贝克汉姆