

Spark稳定版0.9.2版本发布

Spark 0.9.2于昨天（2014年07月23日）发布。对，你没看错，是Spark 0.9.2。Spark 0.9.2是基于0.9的分枝，修复了一些bug，推荐所有使用0.9.x的用户升级到这个稳定版本。有28位开发者参与了这次版本的开发。虽然Spark已经发布了Spark 1.0.x，但是里面有不少的bug，这次的Spark是稳定版。



如果想及时了解Spark、Hadoop或者Hbase相关的文章，欢迎关注微信公共帐号：[iteblog_hadoop](#)

全文如下：

You can download Spark 0.9.2 as either a [source package](#) (6 MB tgz) or a prebuilt package for Hadoop 1 / CDH3 (156 MB tgz), CDH4 (161 MB tgz), or Hadoop 2 / CDH5 / HDP2 (168 MB tgz). Release signatures and checksums are available at the official [Apache download site](#).

Fixes

Spark 0.9.2 contains bug fixes in several components. Some of the more important fixes are highlighted below. You can visit the Spark issue tracker for the full list of fixes.

Spark Core

1. ExternalAppendOnlyMap doesn't always find matching keys. (SPARK-2043)
2. Jobs hang due to akka frame size settings. (SPARK-1112, SPARK-2156)
3. HDFS FileSystems continually pile up in the FS cache. (SPARK-1676)
4. Unneeded lock in ShuffleMapTask.deserializeInfo. (SPARK-1775)
5. Secondary jars are not added to executor classpath for YARN. (SPARK-1870)

PySpark

1. IPython won't run standalone Python script. (SPARK-1134)
2. The hash method used by partitionBy doesn't deal with None correctly. (SPARK-1468)
3. PySpark crashes if too many tasks complete quickly. (SPARK-2282)

MLlib

1. Make MLlib work on Python 2.6. (SPARK-1421)
2. Fix PySpark's Naive Bayes implementation. (SPARK-2433)

Streaming

1. SparkFlumeEvent with body bigger than 1020 bytes are not read properly. (SPARK-1916)

GraphX

1. GraphX triplets not working properly. (SPARK-1188)

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】（）](#)