

Spark Release 1.0.2发布

Spark Release 1.0.2于2014年8月5日发布，Spark 1.0.2 is a maintenance release with bug fixes. This release is based on the branch-1.0 maintenance branch of Spark. We recommend all 1.0.x users to upgrade to this stable release. Contributions to this release came from 30 developers.



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](#)

You can download Spark 1.0.2 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 1.0.2 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

Avoid pulling in the entire RDD or PairRDDFunctions in various operators (SPARK-2534)
RangePartitioner's binary search does not use the given Ordering (SPARK-2598)
Exception in accumulator update should not crash DAGScheduler and SparkContext (SPARK-2323)

SQL

Slave node throws NoClassDefFoundError \$line11.\$read\$ when executing a Spark QL query on HDFS CSV file (SPARK-2576)
Concurrent initialization of various DataType objects causes exceptions (SPARK-2498)
Multiple instances of an InMemoryRelation in a single plan results in recaching (SPARK-2405)

PySpark

Make hash of None consistant cross machines (SPARK-2494)

MLlib

mutable.BitSet in ALS not serializable with KryoSerializer (SPARK-1977)

fix bin offset in DecisionTree node aggregations (SPARK-2152)

Streaming

Ability to limit the Receiver data rate to prevent overloading of data and Spark crash (SPARK-1341)

File stream processes existing files in a directory even if newFilesOnly = true (SPARK-2362)

QueueInputDStream with oneAtATime=false does not dequeue items (SPARK-2343)

GraphX

VertexPartition is not serializable (SPARK-2455)

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: [【】（）](#)