

Apache Spark 1.4.1正式发布(稳定版)

Apache Spark于北京时间2015年07月16日05点正式发布。Spark 1.4.1主要是维护版本，包含了大量的稳定性修复。该版本是基于branch-1.4分支。社区推荐所有1.4.0使用升级到这个稳定版本。此版本有85位开发者参与。

Spark 1.4.1包含了大量的Bug修复，这些Bug出现在Spark的DataFrame、外部数据源支持以

[这里](#)查看。

Data Sources and DataFrames

SPARK-8804: Order of UTF8String is not consistent with String if there is any non-ascii character in it

SPARK-8406: Race condition when writing Parquet files

SPARK-8329: DataSource options parser no longer accepts '_'

SPARK-8368: ClassNotFoundException in closure for map

SPARK-8470: MissingRequirementError for ScalaReflection on user classes

SPARK-8358: DataFrame explode with alias and * fails

MLLib

SPARK-8151: Pipeline components should correctly implement copy

SPARK-8468: Some metrics in RegressionEvaluator should have negative sign

SPARK-8736: GBTRegressionModel shouldn't threshold predictions

SPARK-8563: IndexedRowMatrix.computeSVD() yields the U with wrong numCols

PySpark

SPARK-8202: Infinite loop during external sort

SPARK-8573: Trigger exceptions when invalid operators are used

SPARK-8766: Support non ASCII characters in columns

SparkR

SPARK-8506: Support for Spark packages when initializing SparkR

SPARK-8085: Support for user defined schemas when reading from data sources

下载地址

<https://spark.apache.org/downloads.html>

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: [【】](#) ()