

CarbonData Learning Materials

1. introduction

Apache CarbonData is an indexed columnar data store solution for fast analytics on big data platform, e.g. Apache Hadoop, Apache Spark, etc.

- 1.1 官网: <http://carbondata.apache.org>
- 1.2 Github: <https://github.com/apache/carbondata>.
- 1.3 CarbonData cwiki: <https://cwiki.apache.org/confluence/display/CARBONDATA/>

2. Activity (Video and ppt)

- 2.1 Apache CarbonData ShenZhen Meetup PPT (1st Dec, 2018)
 - Videos:
 - 视频回顾: <http://zhibo.huaweicloud.com/watch/2590022>
 - News:
 - https://www.oschina.net/question/3802184_2290292
 - Slides
 - [What's New in Apache Spark 2.4](#)
 - [基于CarbonData构建万亿级数据仓库](#)
 - [QATCodec Past, Present and Future_rev1.0](#)
 - [CarbonData 的实践与调优](#)
 - [Apache CarbonData & Spark Meetup 主持稿](#)
- 2.2 2018实时大数据Meetup——Spark、Flink、CarbonData的技术与实践, 2018-09-08 09:30 - 12:20:
 - Videos:
 - <http://www.itdks.com/eventlist/detail/2599>.
 - Slides (Realtime BigData BeiJing Meetup(8th Sep, 2018)):
 - [Easy, Scalable, Fault-tolerant stream processing with Structured Streaming-TD](#)
 - [CarbonData技术原理及使用介绍-蔡强](#)
 - [美团点评Flink实践与应用-孙梦瑶](#)
 - [智能流计算Flink_Spark在华为云CloudStream中的应用实践-时金魁](#)
- 2.3 Apache CarbonData Shanghai Meetup PPT (2nd Sep, 2017)
 - Slides:
 - [Spark SQL_Past, Present and Future_范文臣](#)
 - [CarbonData应用实践和未来规划_李昆](#)

- [Cost-Based Optimizer Framework for Spark SQL_王振华](#)
- [CarbonData Partition 功能介绍与上汽集团 CarbonData实践分享_曹鲁](#)
- 2.4 Apache CarbonData Beijing Meetup PPT (29th Oct, 2016)
 - Slides
 - [CarbonData_An Indexed Columnar File Format For Interactive Query](#)
- 2.5 2018中国开源年会: Open Sourcing CarbonData, a distributed global co-operation, development and success story, 2018-10-20 14:10-14:50(COSCon'18):
 - Videos:
 - <https://v.qq.com/x/page/c0797395o2n.html>
 - <https://www.itdks.com/dakalive/detail/17003>.
 - Slides (China Open Source Conference ShenZhen(20 Oct. 2018)):
 - [ApahceCarbonDataOpenSourceJourneyofBecoming TopLevelProjectRaghunandan_S](#)
- 2.6 2017中国开源年会: Apache CarbonData成长故事, 2017-11-18 14:00-14:30:
 - Videos:
 - <http://www.itdks.com/dakalive/detail/7727>
 - Slides:
 - [Apache CarbonData成长故事-PPT.pdf](#)
- 2.7 华为云技术私享会杭州站: 华为云大数据存储技术创新-CarbonData解密及实践,2018-04-22 15:40-16:20:
 - Videos:
 - <https://bbs.huaweicloud.com/videos/21e577defdbb4b5997e464af3bf447be>
 - <http://www.itdks.com/dakalive/detail/11395>
 - [slides](#)
 - Slides:
 - [华为云技术私享会-华为云大数据存储技术创新-CarbonData解密及实践](#)
- 2.8 Spark summit EAST 2017: Apache Carbondata: An Indexed Columnar File Format for Interactive Query by Jacky Li/Jihong Ma:
 - Videos:
 - <https://v.qq.com/x/page/z0800eq6fhx.html>
 - <https://databricks.com/session/apache-carbondata-an-indexed-columnar-file-format-for-interactive-query-with-spark-sql>
 - https://www.youtube.com/watch?v=lhsAg2H_GXc.

- Slides (Spark Summit East 2017-Apache CarbonData (8th Feb,2017)):
 - [Apache CWiki Link](#)
- 2.9 O'Reilly Open Source Convention: Introducing Apache CarbonData : A new hadoop-native file format for faster data analysis, OSCON, May 16 - 19, 2016:
 - Videos:
 - <https://v.qq.com/x/page/j0800dgsdd8.html>
 - <https://www.youtube.com/watch?v=VEckmJuU47g>
- 2.10 Apache Carbondata: An indexed columnar file format for interactive query with Spark SQL:
 - Videos:
 - <https://v.qq.com/x/page/j0800u2ulaj.html>
 - <https://www.youtube.com/watch?v=yya8-GzRW5M>

详细请见: <https://github.com/xubo245/CarbonDataLearning/tree/master/docs/slides/>

3.Documentation

- 3.1 Spark、Flink、CarbonData 技术实践最佳案例解析:
 - <https://www.infoq.cn/article/spark-flink-carbondata-best-practice>
- 3.2 [单表千亿电信大数据场景，使用Spark+CarbonData替换Impala案例:
 - <https://my.oschina.net/u/4029686/blog/2878526?from=groupmessage>
- 3.3 Apache顶级项目CarbonData应用实践与2.0新技术规划介绍:
 - https://www.sohu.com/a/193921606_470008
- 3.4 云计算开源产业联盟发布融合数据白皮书，Apache CarbonData成为主流融合存储技术
 - <https://bbs.huaweicloud.com/blogs/2f27376be9e511e79fc57ca23e93a89f>
- 3.5 CarbonData+Spark SQL的应用实践和调优经验分享:
 - https://www.sohu.com/a/190681267_470008,
 - InfoQ Link: <https://www.infoq.cn/article/2017%2F09%2FCarbonData-Spark-huawei>
- 3.6 深度访谈：华为开源数据格式CarbonData项目，实现大数据即席查询秒级响应:
 - [AI前线 WeChat Link](#)
- 3.7 Apache CarbonData:大数据生态一种新的高性能数据格式:

- <https://my.oschina.net/u/3156280/blog/808506>

4. Performance report

- [CarbonData-1.5.1 TPCD Report](#)

5. Reference

- 5.1 <https://github.com/xubo245/CarbonDataLearning>
- 5.2 https://github.com/QiangCai/carbondata_guide
- 5.3 [Apache CarbonData 中文文档 - 过往记忆](#)

6. WeChat group(微信群)

Please add WeChat ID: xu601450868, I will add you into CarbonData WeChat group

7. 更多资料:

- 7.1 <https://github.com/xubo245/CarbonDataLearning>
- 7.2 <https://github.com/xubo245/CarbonDataLearning/blob/master/docs/learningMaterials/CarbonDataLearningMaterials.md>