

# 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. Vedio

---

- 2.1 2018实时大数据Meetup——Spark、Flink、CarbonData的技术与实践, 2018-09-08 09:30 - 12:20: <http://www.itdks.com/eventlist/detail/2599>.
- 2.2 Open Sourcing CarbonData, a distributed global co- operation, development and success story, 2018-10-20 14:10-14:50(2018中国开源年会 COSCon'18):
  - <https://www.itdks.com/dakalive/detail/17003>.
  - <https://v.qq.com/x/page/c0797395o2n.html>
- 2.3 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: <http://www.itdks.com/dakalive/detail/11395>
- 2.4 Apache CarbonData成长故事, 2017-11-18 14:00-14:30: <http://www.itdks.com/dakalive/detail/7727>
- 2.5 Apache Carbondata: An Indexed Columnar File Format for Interactive Query by Jacky Li/Jihong Ma, Spark summit EAST 2017: [https://www.youtube.com/watch?v=lhsAg2H\\_GXc](https://www.youtube.com/watch?v=lhsAg2H_GXc).
- 2.6 Introducing Apache CarbonData : A new hadoop-native file format for faster data analysis, O'Reilly Open Source Convention: OSCON, May 16 - 19, 2016: <https://www.youtube.com/watch?v=VEckmJuU47g>
- 2.7 Apache Carbondata: An indexed columnar file format for interactive query with Spark SQL: <https://www.youtube.com/watch?v=yya8-GzRW5M>

## 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](https://www.sohu.com/a/193921606_470008)
- 3.4 CarbonData+Spark SQL的应用实践和调优经验分享:
  - [https://www.sohu.com/a/190681267\\_470008](https://www.sohu.com/a/190681267_470008),
  - InfoQ Link: <https://www.infoq.cn/article/2017%2F09%2FCarbonData-Spark-huawei>
- 3.5 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](#)
- 3.6 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.Slides

---

- 4.1 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](#)
- 4.2 Apache CarbonData Shanghai Meetup PPT (2nd Sep,2017)
  - [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](#)
- 4.3 China Open Source Conference ShenZhen(20 Oct. 2018)
  - [ApahceCarbonDataOpenSourceJourneyofBecoming TopLevelProjectRagunandan S](#)
- 4.4 Apache CarbonData Beijing Meetup PPT (29th Oct, 2016)
  - [CarbonData An Indexed Columnar File Format For Interactive Query](#)
- 4.5 Spark Summit East 2017-Apache CarbonData (8th Feb,2017)
  - [Apache CWiki Link](#)

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

## 5.Reference

---

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