



# Release Notes

**Version 0.2.0**

**Published: 11/21/2016**

CarbonData

# Introduction

Apache CarbonData (incubating) is an open source project of The Apache Software Foundation (ASF). CarbonData is a new big data native file format for faster interactive query using advanced columnar storage, index, compression, and encoding techniques to improve computing efficiency, in turn it will help speedup queries an order of magnitude faster over PetaBytes of data.

This release notes provides information on the new features, improvements, and bug fixes of this release. There are more than 30+ new feature and improvements , more than 80+ bug fixes to provide a stable and reliable package. All links provided in this document will guide you to get the latest package, documentations, and more information about this release, and details.

# What's New in Version 0.2.0?

In this version of CarbonData, there are major performance improvements like blocklets distribution, support BZIP2 compressed files, and so on added to enhance the CarbonData performance significantly. Along with performance improvement, there are new features added to enhance compatibility and usability of CarbonData.

## CarbonData Files Readable through Spark/MapReduce Program

Now CarbonData files are readable through Spark/MapReduce programs.
**Category:** Compatibility
**Benefits:** Standardized format for easy integration with BigData eco system.

For more details, see [CARBONDATA-257].

## Support of SQLContext to Read CarbonData

Support SQLContext to read CarbonData file through DataSource API.
**Category:** Compatibility
**Benefits:**

- This improves Spark data source API support.
- Now the CarbonData can be used as Spark data source.
- Now we can read the CarbonData files without creating CarbonContext, we can directly SQLContext to read the files through DataSource API.

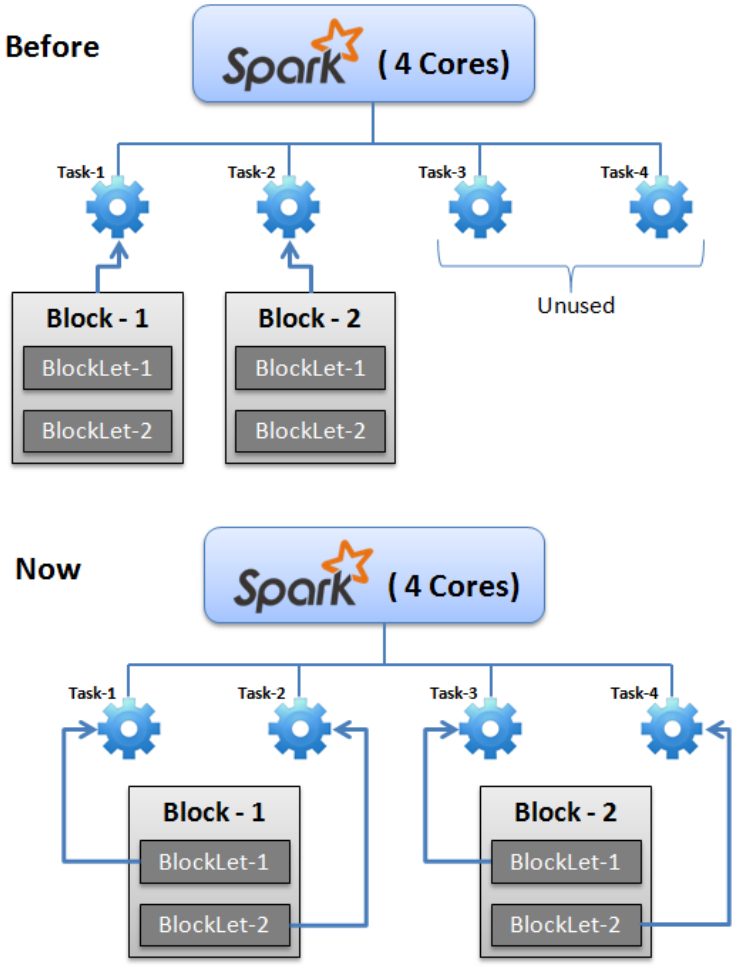
For more details, see [CARBONDATA-212].

## Optimum Resource Usage by BlockLet Distribution

Now instead of using blocks, we further divide blocks to blocklets when there are no sufficient blocks identified. These blocklets can be scanned parallel to utilize all executors.

**Category:** Performance improvement

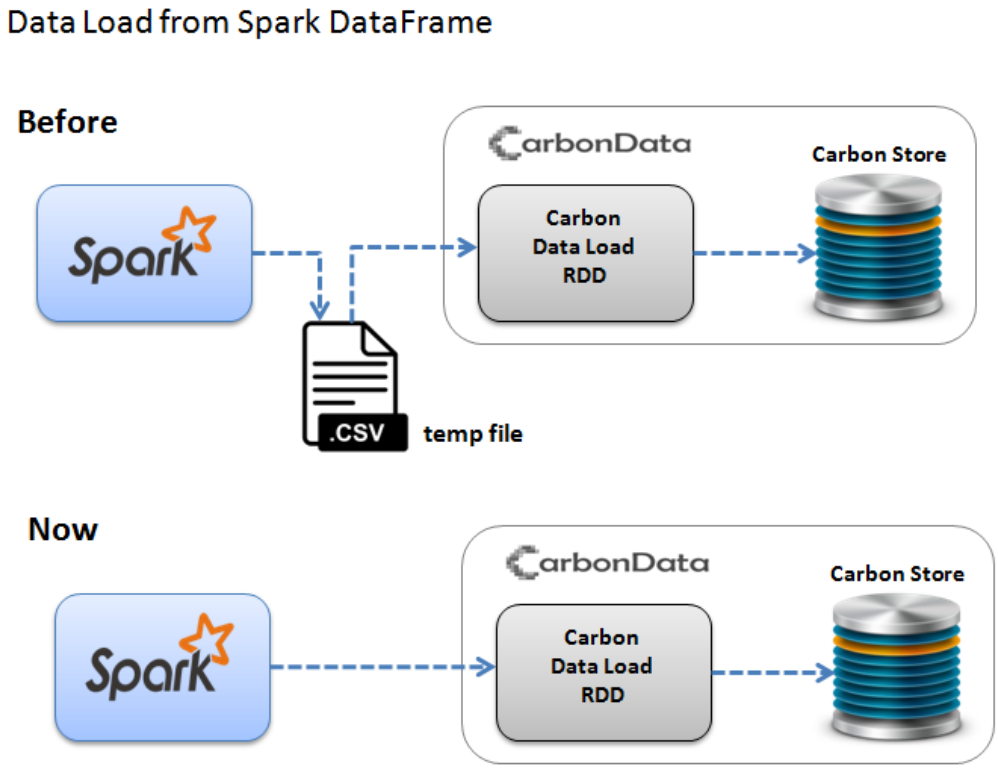
**Benefits:** This improves resource usage and query response time. The efficiency is improved up to three times.




For more details, see [CARBONDATA-117].

### Save DataFrame without Writing into Temporary File

Save a DataFrame to CarbonData file without writing temporary CSV file. Here DataFrame is iterated to write data directly to Carbon.
**Category:** Performance improvement
**Benefits:** Faster and reduced I/O during data load time.




For more details, see [CARBONDATA-279].

### Support BZIP2 Compressed Files

Support loading BZIP2 compressed CSV files.
**Category:** Performance improvement and Compatibility
**Benefits:** The input files are compressed taking less disk space. (half size)

For more details, see [CARBONDATA-210].

### Enhanced Filter Support

Filters with IS NULL and IS NOT NULL are pushed down to Carbon.
**Category:** Performance improvement
**Benefits:** This filters applied at the Carbon reader will result in faster filtering.

For more details, see [CARBONDATA-278].

### Remove Thrift Complier Dependency

Now you can build CarbonData without installing Thrift, unless you want to change the Carbon Thrift file format.

**Category:** Usability
**Benefits:** This helps in reducing build time as no longer requires downloading Thrift compiler.

For more details, see [CARBONDATA-213].

### Support Append Mode

Support Append mode when writing DataFrame to CarbonData.
**Category:** Compatibility
**Benefits:** This helps in appending incremental data from DataFrame.

For more details, see [CARBONDATA-286].

## Fixed Issues

In the CarbonData release version 0.2.0, more than 80+ issues are fixed to provide a stable package.

For complete list of bug fixes, click.

## Release download

https://www.apache.org/dyn/closer.lua/incubator/carbondata/0.2.0-incubating

## Other Resources & Links

- For complete documentation on CarbonData, click.
- To know how to contribute to Carbon data, click.

## Feedback

Please mail you valuable feedback and query to *dev@carbondata.incubator.apache.org*