

Apache TAJO Home

[Apache Tajo](#) is a robust big data relational and distributed data warehouse system for Apache Hadoop. Tajo is designed for low-latency and scalable ad-hoc queries, online aggregation, and ETL (extract-transform-load process) on large-data sets stored on HDFS (Hadoop Distributed File System) and other data sources. By supporting SQL standards and leveraging advanced database techniques, Tajo allows direct control of distributed execution and data flow across a variety of query evaluation strategies and optimization opportunities.

General Information

- [Official Apache Tajo Website: source code, issue-tracking, mailing-lists, etc.](#)
- [Overview of Tajo](#)
- [Powered By](#)
- [Presentations](#)
- [Architecture of Tajo](#)
- [Logos of Tajo](#)

Developer Documentation

- [Roadmap](#)
- [Tajo Internal](#)
- [How To Contribute](#)
- [How To Setup Your Development Environment](#)
- [TPC-H Benchmark](#)
- [TPC-DS Benchmark](#)
- [How to update Apache Tajo website](#)
- [Coding Style](#)
- [UnitTests](#)
- [Major Release Announcement Template](#)
- [How to write user documentations](#)

User Documentation

User documentations is located at <http://tajo.apache.org/docs/current/index.html>.

Others

- [Google Summer of Code 2013](#)