

GSoC 2013

Project List

Apache Tajo (incubating) is a data warehouse system for Hadoop. Tajo is designed for low-latency and scalable ad-hoc queries, online aggregation and ETL on large-data sets by leveraging advanced database techniques. Tajo has gained much attention as one of the SQL-on-Hadoop projects. Tajo is in the alpha stage.

We are looking for competent students who want to participate in Apache Tajo project via GSoC 2013. It would be a great opportunity for you. Here is the project list.

Hybrid Hash Join Operator

- <https://issues.apache.org/jira/browse/TAJO-35>

Improve ExternalSortExec with N-merge sort and final pass omission

- <https://issues.apache.org/jira/browse/TAJO-36>

Outer Join

- <https://issues.apache.org/jira/browse/TAJO-34>
- Design document link: <https://sites.google.com/site/gsoc2013tajo34/home>

How to Apply

Please read [GSoC guide for student](#) to apply. It is highly recommend to discuss your interest before you apply. The best way to discuss is to comment on individual Jira or send mail to dev list.