TajoProposal

Abstract

Tajo is a distributed data warehouse system for Hadoop.

Proposal

Tajo is a relational and distributed 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. It supports SQL standards. Tajo is inspired by Dryad, MapReduce, Dremel, Scope, and parallel databases. Tajo uses HDFS as a primary storage layer, and it has its own query engine which allows direct control of distributed execution and data flow. As a result, Tajo has a variety of query evaluation strategies and more optimization opportunities. In addition, Tajo will have a native columnar execution and and its optimizer. Tajo will be an alternative choice to Hive/Pig on the top of MapReduce.

Background

Big data analysis has gained much attention in the industrial. Open source communities have proposed scalable and distributed solutions for ad-hoc queries on big data. However, there is still room for improvement. Markets need more faster and efficient solutions. Recently, some alternatives (e.g., Cloudera's Impala and Amazon Redshift) have come out.

Rationale

There are a variety of open source distributed execution engines (e.g., hive, and pig) running on the top of MapReduce. They are limited by MR framework. They cannot directly control distributed execution and data flow, and they just use MR framework. So, they have limited query evaluation strategies and optimization opportunities. It is hard for them to be optimized for a certain type of data processing.

Initial Goals

The initial goal is to write more documents to describe Tajo's internal. It will be helpful to recruit more committers and to build a solid community. Then, we will make milestones for short/long term plans.

Current Status

Tajo is in the alpha stage. Users can execute usual SQL queries (e.g., selection, projection, group-by, join, union and sort) except for nested queries. Tajo provides various row/column storage formats, such as CSV, RowFile (a row-store file we have implemented), RCFile, and Trevni, and it also has a rudimentary ETL feature to transform one data format to another data format. In addition, Tajo provides hash and range repartitions. By using both repartition methods, Tajo processes aggregation, join, and sort queries over a number of cluster nodes. To evaluate the performance, we have carried out benchmark test using TPC-H 1TB on 32 cluster nodes.

Meritocracy

We will discuss the milestone and the future plan in an open forum. We plan to encourage an environment that supports a meritocracy. The contributors will have different privileges according to their contributions.

Community

Big data analysis has gained attention from open source communities, industrial and academic areas. Some projects related to Hadoop already have very large and active communities. We expect that Tajo also will establish an active community. Since Tajo already works for some features and is in the alpha stage, it will attract a large community soon.

Core Developers

Core developers are a diverse group of developers, many of which are very experienced in open source and the Apache Hadoop ecosystem.

- Eli Reisman <ereisman AT apache DOT org>
- Henry Saputra <hsaputra AT apache DOT org>
- Hyunsik Choi <hyunsik AT apache DOT org>
- Jae Hwa Jung <jhjung AT gruter DOT com>
- Jihoon Son <ghoonson AT gmail DOT com>
- Jin Ho Kim <jhkim AT gruter DOT com>
- Roshan Sumbaly <rsumbaly AT gmail DOT com>
- Sangwook Kim <swkim AT inervit DOT com>
- Yi A Liu <yi DOT a DOT liu AT intel DOT com>

Alignment

Tajo employs Apache Hadoop Yarn as a resource management platform for large clusters. It uses HDFS as a primary storage layer. It already supports Hadoop-related data formats (RCFile, Trevni) and will support ORC file. In addition, we have a plan to integrate Tajo with other products of Hadoop ecosystem. Tajo's modules are well organized, and these modules can also be used for other projects.

Known Risks

Orphaned Products

Most of codes have been developed by only two core developers, who are Hyunsik Choi and Jihoon Son. It may be a risk of being orphaned. However, they are guaranteed to have enough time to develop Tajo for years. As you can see the commit history, they have participated in this project for about two years. In addition, the initial committers are diverse, and Tajo has been supported by two IT companies in South Korea. So, the risk of being orphaned is very low. Later, we will be eager to recruit additional committers in order to eliminate this risk.

Inexperience with Open Source

Most of the initial committers have experience working on open source projects. In particular, Eli, Henry, and Hyunsik have experience as committers and PMC members on other Apache projects.

Homogeneous Developers

Although they are a diverse group of developers, what a half of core developers are in South Korea may be a risk. This is because their offline activities are limited due to their location. Since we surely recognize this risk, we will write more complete documents and presentation materials in order to disseminate Tajo's internal and users guide. In addition, to mitigate this risk we will be eager to recruit additional committers around the world.

Reliance on Salaried Developers

It is expected that Tajo development will occur on both salaried time and on volunteer time. Hyunsik and Jihoon belong to Database lab., Korea Univ. They will be paid by the lab to contribute Tajo for years. Jin Ho and Sangwook are paid by their employer to contribute to this project. Other developers will contribute to this project on volunteer time. In addition, we will be eager to recruit additional committers including salaried and non-salaried developers.

Relationships with Other Apache Products

Tajo has some overlapping function with Apache Incubator Drill. However, Tajo is even more mature than Drill. In addition, there are some significant differences. Drill is a distributed system specialized for low-latency query processing by using column operations and intermediate data streaming. Drill has very simple query optimizer. However, some queries including big-big table join and sort are not available in that manner. Drill will support some of query types.

In contrast, Tajo has advanced query optimization system. Tajo mainly aims at scalable and efficient processing on all query types. By using the query optimizer, Tajo will only chase low latency query processing for some query types that can be executed in online aggregation manner.

Besides, Tez has some overlapping functions with Tajo. However, Tez is in the pre-alpha stage and may be a prototype. When Tez becomes feasible, Tajo could use Tez as an underlying framework according to the applicability. However, Tajo will still use its row/native columnar execution engine and its optimizer. Tajo may be potentially the first application of Tez.

A Excessive Fascination with the Apache Brand

We believe that the Apache brand will help us to find contributors and to grow the community. The community and development process will make this project more stable and help establish ubiquitous APIs. In addition, Tajo depends other project in Apache Hadoop ecosystem. We expect that cooperative work occurs with other projects in the same place.

Documentation

Tajo's demonstration paper was accepted to IEEE ICDE 2013. Since this conference will be held in April 2013, we cannot publicly show the paper. Instead, we attached some presentation material. Checkout this Slide.

In addition, some documents (e.g., getting started) are available at http://tajo-project.github.com/tajo/

Initial Source

The initial source code has been developed in the Database Lab. Korea Univ. This is implemented in Java and has almost 100,000 lines except for parser and protobuf generated codes. Currently, initial source code is already available on GitHub at https://github.com/tajo-project/tajo.

Source and Intellectual Property Submission Plan

We intend the entire code base to be licensed under the Apache License, Version 2.0.

External Dependencies

The required dependencies are all Apache compatible licenses. The following components with non-Apache licenses are enumerated:

- Google Guava
- Google Protocol Buffer
- Antlr
- Mockito
- JLine2

Cryptography

Tajo will depend on secure Hadoop that can optionally use Kerberos.

Required Resources

Mailling List

- tajo-private (with moderated subscriptions)
- tajo-dev
- tajo-commits

Subversion Directory

https://git-wip-us.apache.org/repos/asf/tajo.git

Issue Tracking

Jira Tajo (TAJO)

Other Resources

- Continuous Integration
  - Jenkins
- Wiki
  - http://wiki.apache.org/tajo

Initial Committers

- Eli Reisman <ereisman AT apache DOT org>
- Henry Saputra <hsaputra AT apache DOT org>
- Hyunsik Choi <hyunsik AT apache DOT org>
- Jae Hwa Jung <jhjung AT gruter DOT com>
- Jihoon Son <ghoonson AT gmail DOT com>
- Jin Ho Kim <jhkim AT gruter DOT com>
- Roshan Sumbaly <rsumbaly AT gmail DOT com>
- Sangwook Kim <swkim AT inervit DOT com>
- Yi A Liu <yi DOT a DOT liu AT intel DOT com>

Affiliations

- Eli Reisman (Hortonworks)
- Henry Saputra (Platfora)
- Hyunsik Choi (Database Lab., Korea University)
- Jae Hwa Jung (Gruter)
- Jihoon Son (Database Lab., Korea University)
- Jin Ho Kim (Gruter)
- Roshan Sumbaly (LinkedIn)
- Sangwook Kim (Inervit)

- Yi A Liu (Intel)

The nominated mentors are employees of NASA JPL, LinkedIn, and Hortonworks.

- Chris Mattmann - NASA JPL
- Jakob Homan - LinkedIn
- Owen O'Malley - Hortonworks
- Alex Karasulu

Sponsors

Champion

- Jakob Homan <jghoman AT apache DOT org>

Nominated Mentors

- Chris Mattmann <chris DOT a DOT mattmann AT jpl DOT nasa DOT gov>
- Jakob Homan <jghoman AT apache DOT org>
- Owen O'Malley <omalley AT apache DOT org>
- Alex Karasulu <akarasulu AT apache DOT org>

Sponsoring Entity

Apache Incubator