

PistachioProposal

Pistachio

Abstract

Pistachio is a fault-tolerant low latency distributed storage system which enables simple embedding the computation to the storage layer to achieve best data locality. It evolves from Yahoo's global user profile storage system.

Proposal

Pistachio is a distributed key value store system with fault tolerance and consistency guarantee. It supports multiple local storage engine including in-memory, kyoto cabinet, rocks DB etc. Pistachio is being used as the user profile storage for massive scale global ads products in Yahoo storing 10+ billion user profiles. The performance and reliability has been well proven on production.

Pistachio can easily embed computation to the storage layer to achieve the best data locality to improve the computation performance significantly which is an innovative model comparing with the normal ways where the storage and compute are independent to each other.

Background

Pistachio is originally designed and optimized for Yahoo's large scale global open RTB(real-time bidding) use cases where latency is critical(the whole request needs to be finished within 100ms including network round trips). It stores 10+ billion user profiles in 8 data centers.

Then because of the great performance and the flexibility of local storage choices, we evolved it to do distributed compute. Rich call back interfaces are added to supports easy compute directly on top of the storage system local to the data partition. This model is totally different from the traditional distributed computation model where the storage and compute are separated and independent. In the new model we found data locality can be improved significantly and lots of data access round trips can be reduced in computation, and the performance can be improved significantly.

It was publicly announced in April 2015 and currently being hosted in Github.

Rationale

As a key value store system Pistachio is unique in terms of low latency access with fault tolerance and consistency guarantee. The reliability, scalability, fault tolerance and performance has been well proven in global large scale revenue supporting production system in Yahoo.

As a distributed computation system, it's a different model from the common models like Hadoop that the compute layer is introduced on top of the storage layer natively and naturally to optimize the data locality of computation. Comparing with the Hadoop model, intermediate results can always be stored in a partitioned way and the computation can be distributed to storage partitions to avoid data transfer in the shuffling phase between map reduce jobs in Hadoop to achieve much better throughput.

Operating the project in "apache way" greatly aligns with the long-term vision of this project and can greatly help the development of the community.

Current Status

Pistachio was open-sourced and announced in April 2015 and currently being hosted in Github, it was mainly being developed by the team from Yahoo and already attracted lots of external developers (20+ watches and forks on github).

Meritocracy

We plan to build an environment following the Apache meritocracy principles. Many companies including LinkedIn, GF securities, Microsoft and open source communities like deeplearning4j have already expressed interests or accepted the invitations to participate in this project.

Community

Since the announcement of Pistachio we received lots of interests. And the concept of embedding computation to storage also got lots of recognitions. We also started to work with other communities like deeplearning4j to build more application use cases with Pistachio. We believe the community will grow fast.

Core Developers

This project is created by Gavin Li. Core developers are currently mainly in Yahoo.

Alignment

Pistachio depends on many Apache projects and dependencies including Kafka, Helix, Zookeeper, Curator, Apache Commons, etc.

Known Risks

Orphaned Products

The risk of Pistachio being orphaned is small because Yahoo heavily invested in this system. It's the internal storage standard for Yahoo's global ads products and still being expanded. Migration cost from this project is very high. We are also working with external communities like deeplearning4j and other companies to expand the applications.

Inexperience with Open Source

Core developers are experienced open source contributors in many projects including Druid, Spark, Storm, etc. Pistachio committers will be guided by the mentors with strong Apache open source project backgrounds.

Homogeneous Developers

The initial committers include developers from several institutions including Microsoft, GF Securities, LinkedIn and Yahoo.

Reliance on Salaried Developers

We work on Pistachio on both salaried time and after hours. Many developers from other institutions already accepted the invitation to volunteer working on Pistachio.

Relationships with Other Apache Products

As mentioned earlier, Pistachio depends on apache kafka, helix, zookeeper, curator, etc.

A Excessive Fascination with the Apache Brand

Generating publicity is not the purpose of this proposal. We mainly want to join the ASF in order to increase our contacts and visibility in the open source world to attract great developers.

Document

Current documentation can be found here: <https://github.com/yahoo/Pistachio>.

Initial source

Initial source can be found here in the Github repo: <https://github.com/yahoo/Pistachio>.

External dependencies

To the best of our knowledge, here is the list of dependencies:

- Rocks DB
- ICU4j
- Apache Curator
- netty
- google http client
- codahale.metrics
- apache helix
- apache zookeeper
- apache commons
- apache thrift
- apache kafka
- kyoto cabinet (GNU GPL)
- google protocol buffer
- kryo
- slf4j

To the best of our knowledge, except kyoto cabinet others are all distributed under Apache compatible licenses:

- BSD
- ICU
- Apache License 2.0
- MIT

Kyoto cabinet is under GNU GPL, but it is not a hard necessary dependency to Pistachio, it's an optional pluggable storage engine. It's designed in the way that it's totally pluggable and very loosely coupled. We can easily remove it before our first Apache incubating release.

Required Resources

Mailing Lists

- dev@pistachio.incubator.apache.org
- commits@pistachio.incubator.apache.org
- private@pistachio.incubator.apache.org (for private PMC discussions)

Git

The Pistachio team prefers Git for source version control: [git://git.apache.org/pistachio](https://git.apache.org/pistachio)

Issue Tracking

JIRA Pistachio (PISTACHIO)

Other Resources

Jenkins continuous integration testing

Initial Committers

- Gavin Li <lyo.gavin@gmail.com>
- Lie Yang <lyang@yahoo-inc.com>
- Jay Kim <pitecus@yahoo-inc.com>
- Flavio Junqueira <fpj@apache.org>
- Chihong Liang <chihong.liang@gmail.com>
- Yong Liu <ly7110@gmail.com>
- Shengwu Yang <yangshengwu@gmail.com>
- Kishore Gopalakrishna <g.kishore@gmail.com>

Affiliations

- Gavin Li - Yahoo
- Flavio Junqueira - Microsoft
- Chihong Liang - GF securities
- Yong Liu - Yingmi Asset Management Corp.
- Lie Yang - Yahoo
- Jay Kim - Yahoo
- Shengwu Yang - LinkedIn China
- Kishore Gopalakrishna - LinkedIn

Sponsors

Champion

Nominated Mentors

- Jake Farrell <jfarrell@apache.org>
- Amareshwari Sriramadasu <amareshwari@apache.org>

Sponsoring Entity

The Apache Incubator