

PoweredBy

Applications and organizations using Hive include (alphabetically):

Bizo

We use Hive for reporting and ad hoc queries.

Chitika

We use Hive for data mining and analysis on our 435M monthly global users.

CNET

We use Hive for data mining, internal log analysis and ad hoc queries.

Digg

We use Hive for data mining, internal log analysis, R&D, and reporting/analytics.

eHarmony

We use Hadoop to store copies of internal log and dimension data sources and use it as a source for reporting/analytics and machine learning. Currently have a 640 machine cluster with ~5000 cores and 2PB raw storage. Each (commodity) node has 8 cores and 4 TB of storage.

Grooveshark

We use Hive for user analytics, dataset cleaning, and machine learning R&D.

hi5

We use Hive for analytics, machine learning and social graph analysis.

HubSpot

We use Hive as part of a larger Hadoop pipeline to serve near-realtime web analytics.

Last.fm

We use Hive for various ad hoc queries.

MedHelp Find a Doctor

We implemented Hive to analyse large amounts of doctors across the United States, and for internal analytics for over 1M pageview/day.

NexR

- [Replacing RDB&DW with Hadoop and Hive](#)

We use hive for replacing Oracle DW, big data analysis and integrating R. We develop the enterprise Hive.

Papertrail

- [Log analytics with Hadoop and Hive](#)

We use Hive as a customer-facing analysis destination for our hosted syslog and app log management service.

Rocket Fuel

We use Hive to host all our fact and dimension data. Off this warehouse, we do reporting, analytics, machine learning and model building, and various ad hoc queries.

SaaSPulse

We use Hive for analytics, machine learning and customer interaction analysis of web applications.

Scribd

We use hive for machine learning, data mining, ad-hoc querying, and both internal and user-facing analytics

TaoBao

We use Hive for data mining, internal log analysis and ad-hoc queries. We also do some extensively developing work on Hive.

Trending Topics

Hot Wikipedia Topics, Served Fresh Daily. Powered by Cloudera Hadoop Distribution & Hive on EC2. We use Hive for log data normalization and building sample datasets for trend detection R&D.

VideoEgg

We use Hive as the core database for our data warehouse where we track and analyze all the usage data of the ads across our network.