hid-sp18-523.bib bibempty

0.1 H20

"H20 is a open source platform that can execute highly advanced and complex machine learning algorithms in faster and scalable way,regardless of the size, format and location of the data" [**?**]."It achieves this by serializing the data faster between nodes and clusters that stores huge amount of data. Data processing is done in memory thus providing faster response" [**?**]."It uses fine grain parallelism technique for processing of distributed data archiving 100x faster speed as compared to traditional mapreduce without compromising on accuracy" [**?**].H2O4GPU, Sparking Water and driverless AI are popular products of H2O.Many companies across different domain such as banks, insurance, online sales are using H2O platform for their machine learning and AI related research.

0.2 Google Dremel

With vast amount of publicly available data over the internet / cloud , there was a need of technological system/framework that is deployed on cloud which can execute on demand queries in faster and scalable way for read only multi level nested data. Along with that a system that uses structured query language, which is widely adapted and extensively used by the developers for writing queries to avoid the learning curve of new language. To fill this gap Google came up with Dremel. It is a interactive ad hoc query system that lets the user query the large dataset providing them results with much faster speed compared to traditional technologies [**?**]. "By combining multi-level execution trees and columnar data layout, it is capable of running aggregation queries over trillion-row tables in seconds" [**?**]. "Dremel is capable of scaling up to thousands of CPUs and petabytes of data" [**?**]. MapReduce framework and technologies thar are built over it such as Pig, Hive etc has latency issue between running the job and getting output. Dremel on the other hand took a different approach, it uses execution engine based on on aggregating trees algorithm that provides output almost realtime for queries.

0.3 Google Genomics

With the size of medical data getting increased exponentially from petabytes to exabytes rapidly, Google came up with Google Genomics as extension to Google cloud platform.It helps the life science community organize the world's genomic information and make it accessible and useful [**?**]. Researchers are able to apply Google powerful technologies such as Google Search and Maps to securely store, process, explore, and share large, complex genomics datasets [**?**]. "Multiple genome repositories data can be processed using Google Genomics within seconds as it is backed by Google bigtable and Spanner technologies" [**?**]. "It is based on open standard from Global Alliance of Genomics and Health achieving higher level of interoperability for genomics data" [**?**]. It is fully integrated with Google cloud virtual machine, storage and SQL/NoSQL databases [**?**]."It helps analysing Genomic data in real-time with BigQuery, in literate programming style with Cloud Datalab, in batch with GATK on Google Genomics, with Apache Spark or Cloud Dataflow, or with a Grid Engine cluster" [**?**].

0.4 Google Vision

"Google Cloud Vision API has made the herculean task of correct labeling/classification of images simple. With exponential increase in different types of data including images, voice , video are transformed into digital form, stored and transmitted over network. There was a dire need of automated technology solution that can correctly classify / label images with high level of confidence; Google Vision API provides such platform to researchers and developers. It quickly classify images into thousands of predefined meaning ful categories" [**?**]. It does this by encapsulating powerful machine learning models (KNN , Regression) etc for classification of

be developed in C sharp, Go, Java, Node.js, PHP, Python, and Ruby. Already existing JDBC driver with popular third-party tools can be used to connect with spanner" [**?**]. It is purposely built for global transactional consistency.

References

[heading=bibempty]