
MAIN POINTS

Lesson 8 – Apache Hive

Enjoy greater efficiency and accomplish more

WHOLENESS OF THE LESSON

Apache Hive is a data warehouse infrastructure built on top of MapReduce for providing data summarization, query and analysis. It provides a very well-known SQL-like language called HiveQL with schema on read and transparently converts queries to MR, Apache Tez and Spark jobs. **Science & Technology of Consciousness:** Just like the infinite reservoir of creative intelligence at the source of thought can be open to experience, the infinite knowledge of SQL can be used to experience new insights from the big data. Thus, with Hive, there is no need to learn any new language to be more efficient.

MAIN POINTS

1. The Hive Metastore provides a simple mechanism to project structure onto large amounts of unstructured data by applying a table schema on top of the data. This table abstraction of the underlying data structures and file locations presents users with a relational view of data in the file systems and NoSQL databases. Structure is applied to data at time of read, so users don't need to worry about formatting the data when it is stored in their cluster. **Science & Technology of Consciousness:** Likewise, the huge variety of expressions in the manifest world is merely an abstraction of the hidden, unmanifest dynamics of pure consciousness.
2. Hive features an SQL-like programming interface called HiveQL to query data stored in various databases and file systems. HiveQL automatically translates SQL-like queries into batch MapReduce jobs. **Science & Technology of Consciousness:** Likewise, pure consciousness is the silent level of intelligence that underlies all expressions of intelligence in the form of thoughts and actions in life.
3. Managed and External tables are the two different types of tables in hive used to improve how data is loaded, managed and controlled. Use managed table when you don't want the data after deletion. Use external tables when the data is also used outside of Hive and is not managed by Hive. **Science & Technology of Consciousness:** External tables concept is reminiscent of the incorruptible quality of pure consciousness – "No weapon can cut it into pieces, nor can it be burned by fire, nor moistened by water, nor withered by the wind."
4. Hive organizes tables into partitions. It is a way of dividing a table into related parts based on the values of partitioned columns such as date, city, and department. Using partition, it is easy to query a portion of the data. **Science & Technology of Consciousness:** Pure consciousness manifests as individuals in space. So even though individuals look different, like different hive partitions, basically they are the same representation of the underlying reality of Unity consciousness.

UNITY CHART

CONNECTING THE PARTS OF KNOWLEDGE WITH THE WHOLENESS OF KNOWLEDGE:

Building High-Level Dataflows over Map-Reduce

1. Hive stores schema in a database and processed data into HDFS.
2. Hive is used for completely structured Data mainly by data analysts.

-
3. Transcendental consciousness: is the field beyond type and interaction; it is the field of unbounded awareness and infinite silence.
 4. Impulses within the Transcendental Field: Nature accomplishes what it needs by having its impulses in the transcendental field be as efficient as possible.
 5. Wholeness moving within itself: In Unity Consciousness, the pure intelligence located in TC is found pervading all of creation, from gross to subtle.

