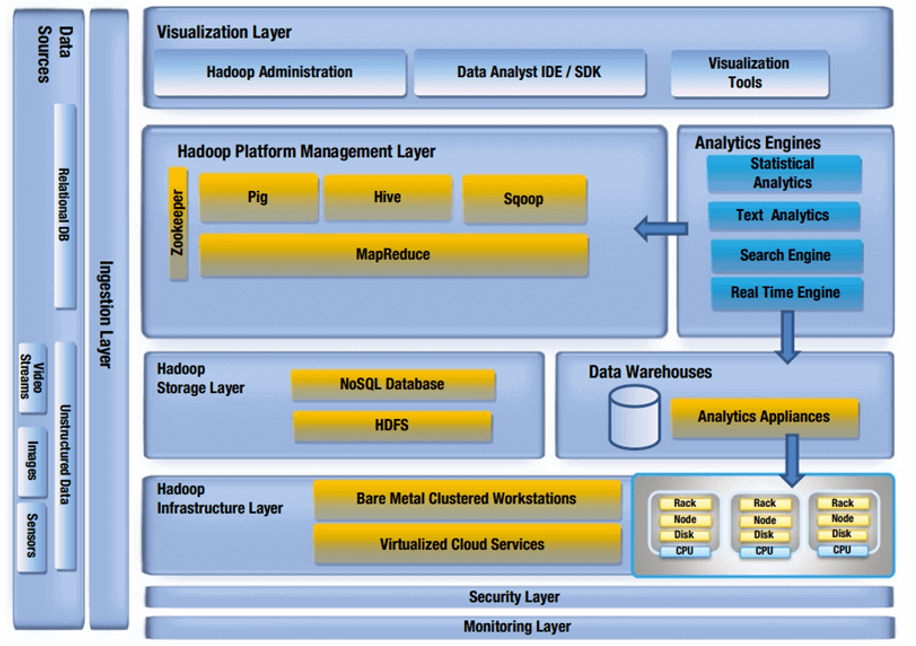
**Architecture Big Data**

Big data merupakan kumpulan dataset yang besar baik dari sisi volume, kecepatan maupun variasi data yang beragam baik secara terstruktur maupun tidak terstruktur yang sulit untuk disimpan pada databae tradisional. Big data biasanya dimanfaatkan untuk membantu kegiatan bisnis dan dapat menghasilkan prediksi, analisis, dan lain-lain

Arsitektur big data merupakan suatu model/struktur yang logis dalam menangani seberapa banyak data yang akan diakses, dikelola maupun disimpan. Arsitektur big data ini menjadi salah satu referensi awal dalam merancang infrastruktur big data besera solusi yang akan digunakan nantinya dalam menangani kasus-kasus tertentu.



Dalam merancang arsitektur big data, terdapat beberapa layer yaitu :

* 1. Data Source

Data source merupakan kumpulan data yang akan diolah (input) yang terdiri dari berbagai macam tipe data baik yang terstruktur maupun tidak terstruktur.

* 2. Ingestion Layer

Merupakan layer sebagai jembatan antara data source dengan layer berikutnya dimana pada layer ini semua data inputan yang masuk akan disortir terlebih dahulu untuk menghindari terjadinya data yang tidak konsisten.

* 3. Hadoop Storage Layer

Hadoop merupakan sebuah framework open source yang memungkinkan kita untuk dapat menyimpan volume data yang cukup besar dengan cara mendistribusikannya melalui sistem-sistem yang tersedia.

* 4. Hadoop Infrastructure Layer

Data source merupakan kumpulan data yang akan diolah (input) yang terdiri dari berbagai macam tipe data baik yang terstruktur maupun tidak terstruktur.

* 5. Hadoop Platform Management Layer

Pada layer ini menyediakan tools maupun query untuk mengakses database menggunakan sistem penyimpanan file HDFS yang berada pada lapisan infrastruktur fisik Hadoop.

* 6. Security Layer

Merupakan sebuah layer untuk menjaga keamanan data yang disimpan ataupun diproses untuk menghindari adanya ancaman dari luar.

* 7. Monitoring Layer

Merupakan layer untuk melakukan monitoring system terhadap kinerja sistem yang sedang dijalankan.

* 8. Visualization Layer

Merupakan layer dimana akan menampilkan hasil dari analisis yang telah dijalankan secara visual baik secara static, dinamik, ataupun interaktif.

Dari layer-layer yang telah disebutkan sebelumnya, semua data yang akan diproses dapat diolah ke analytics engines berupa analisis statistic, analisis teks, sistem search, maupun analisis secara realtime.