**DATABASE**

**Oracle**

Oracle is the relational database management system.it is type of one

Database. Oracle is the first database designed for enterprise grid computing. It is

used for manage the information and applications.

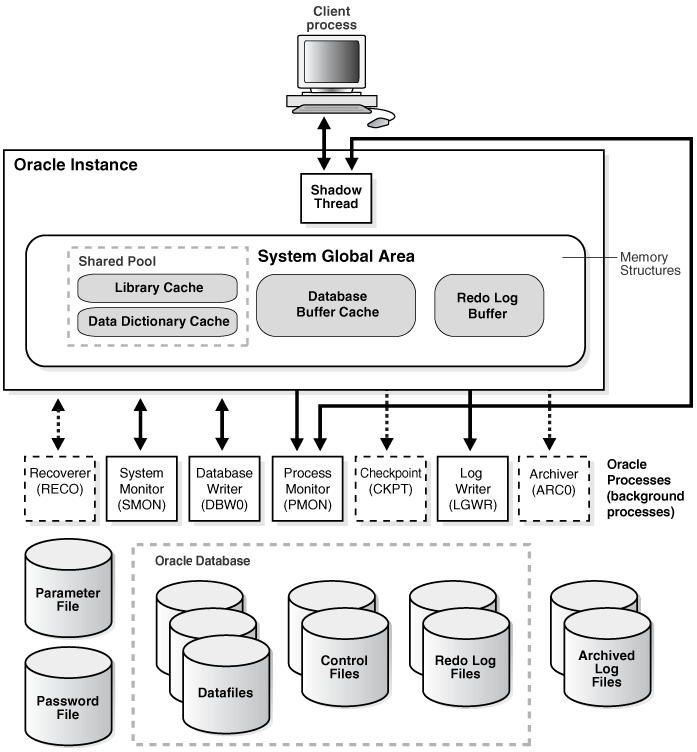
The enterprise grid computing provides the terms:

* Flexibility
* Cost effective

Uses:

* OLTP (Online Transaction Processing)
* DW (Data Warehousing)
* Mixed database workload

Oracle is the information of OS files or disks and also background process.



Features:

* Application Development
* Availability
* Business Intelligence
* Clustering
* Content Management
* Database Overall
* Database Security
* Grid Computing

**MySQL**

MySQL is the world most popular database that is open source and free. It is the open source relational DBMS.It is introduced in 1995 by oracle as a part of sun microsystem acquisition in 2009.

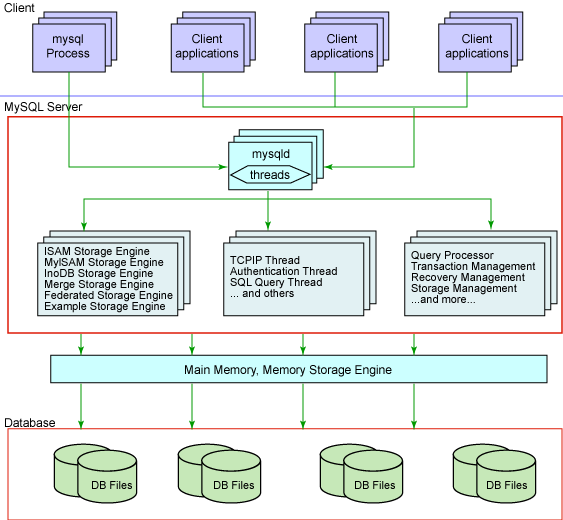
Microsoft runs a virtually all platform like as Linux, Unix and Windows. It is used for mostly web application and including Drupal, Joolma, phpBB and

WordPress.

The purpose of SQL is used for adding, accessing and managing the content.

Uses:

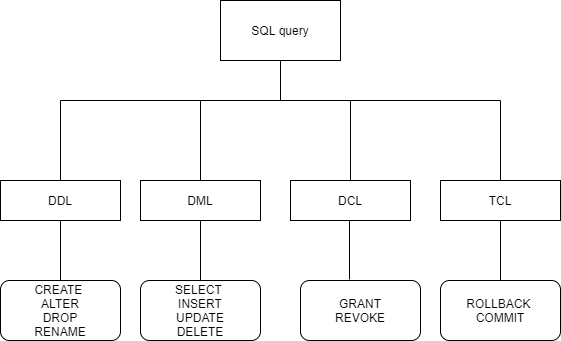
* Quick processing
* Proven reliability
* Easy
* Flexibility
* SQL key properties:
* Database management system.
* Relational.
* MySQL software is Open Source.
* Server works in client/server or embedded systems.
* Microsoft SQL server
* SQL server is the database server by Microsoft.SQL is handle data in a
* relational management system to designed using the special purpose programming languages.
* The data server is provide the data services to another programs or other computer.it is the client server model.it is support both 16 bit and 32 bit environments.



Features:

* Row level security
* Transfort level security
* COMMANDS OF SQL:
* DDL (Data Definition Languages)
* DML (Data Manipulation Languages)
* TCL (Transaction Control Languages)
* DCL (Data Control Languages)

Types of SQL commands



POSTGRE SQL

PostgreSQL is a general purpose and object-relational database management system, the most advanced open source database system

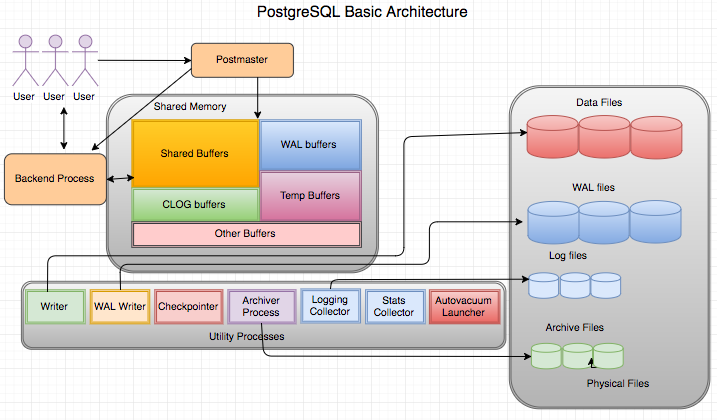
PostgreSQL was designed to run on UNIX-like platforms. However, PostgreSQL was then also designed to be portable so that it could run on various platforms such as Mac OS X, Solaris, and Windows

PostgreSQL, also known as Postgres, is a free and open-source relational database management system emphasizing extensibility and technical standards compliance. It is designed to handle a range of workloads, from single machines to data warehouses or Web services with many concurrent users.

PostgreSQL is designed to be extensible.

PostgreSQL is an open-source, object-relational database management system (ORDBMS) that is not owned or controlled by one company or individual. Because postgresSQL software is open-source, it is managed mostly through a coordinated online effort by an active global community of developers, enthusiasts and other volunteer.

It has become the first choice for corporations that perform complex and high-volume data operations due to its powerful underlying technology



Features:

* User-defined types
* Table inheritance
* Sophisticated locking mechanism
* Foreign key referential integrity
* Views, rules, subquery
* Nested transactions (savepoints)
* Multi-version concurrency control (MVCC)
* Asynchronous replication

**HBASE DATABASE:**

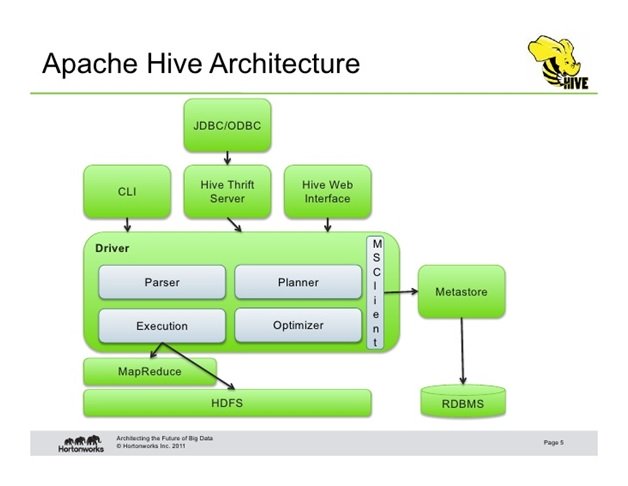
HBASE is called the Hadoop database.It is a NoSQL database runs on the top of the Hadoop.

HBase is an open-source non-relational distributed database model.

It is developed as part of Apache Software Foundation's Apache Hadoop project and runs on top of HDFS or Alluxio, providing Bigtable-like capabilities for Hadoop.

Apache HBase is used to random, real-time read/write access to Big Data. It hosts very large tables on top of clusters of commodity hardware. Apache HBase is a non-relational database modeled after Google's Bigtable. Bigtable acts up on Google File System, likewise Apache HBase works on top of Hadoop and HDFS.

It is highly scalable and can process a form of data types.



Features of HBase

* HBase is linearly scalable.
* Automatic failure support.
* Provides consistent read and writes.
* Integrates with Hadoop, both as a source and a destination.
* It has easy java API for client.
* It provides data replication across clusters.

**//HBase** is Called the Hadoop **database** because it is a NoSQL **database** that runs on top of Hadoop. It combines the scalability of Hadoop by running on the Hadoop Distributed File System (HDFS), with real-time **data** access as a key/value store and deep analytic capabilities of Map Reduce.

//It is developed as part of Apache Software Foundation's Apache Hadoop project and runs on top of HDFS or Alluxio, providing Bigtable-like capabilities for Hadoop

**HIVE :**

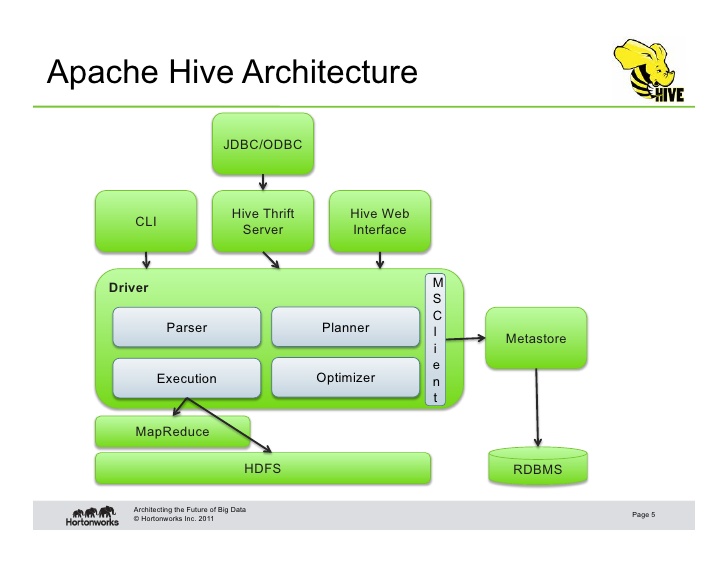
Hive is a data warehouse software project build on top of apache Hadoop and it providing data query and analysis. Hive is an ETL of Hadoop ecosystem and used for processing structured and semi structure data.

Hive is a database attending in Hadoop ecosystem perform DDL and DCL operations.

 Hive gives a SQL-like interface to query data stored in various databases and file systems that integrate with Hadoop.

// and it provides flexible query language such as HQL for better querying and processing of data.

//Apache **Hive works** by translating the input program written in the **hive** SQL like language to one or more Java map reduce jobs. It then runs the jobs on the cluster to produce an answer. It functions analogously to a compiler - translating a high level construct to a lower level language for execution.



REDIS:

Re-dis (Remote Dictionary Server) is an open source in-memory, key value, it is also used as database, cache and message broker

It supports data structures such as

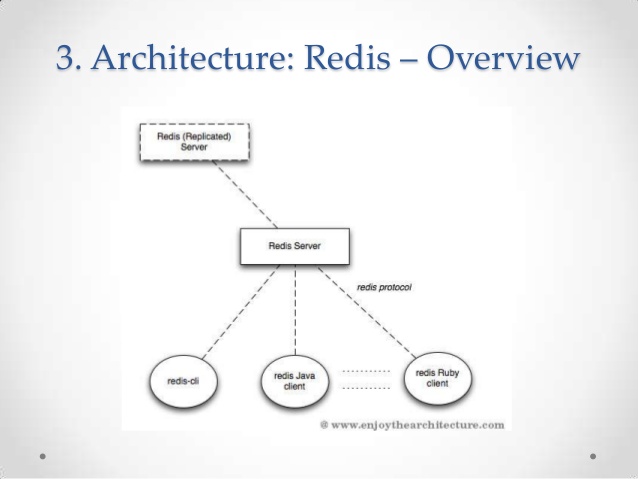
* + - String
    - Hashes
    - Lists
    - Sets
    - Sorted sets with range queries
    - Bitmaps
    - Hyper logs
    - Streams

Re-dis is powerful tool which has a ability to work with different types of data it is often used for cache management and speeding up web application

It also BSD licensed with in-memory key-value database, in-memory data structure project implementing a distribution with optional durability

The difference between Re-dis and other key-value database is Re dis's ability to store and manipulate high-level data types

Re-dis can accept keys in large range of formats, operation that improves on server side and reduce the client's workload.



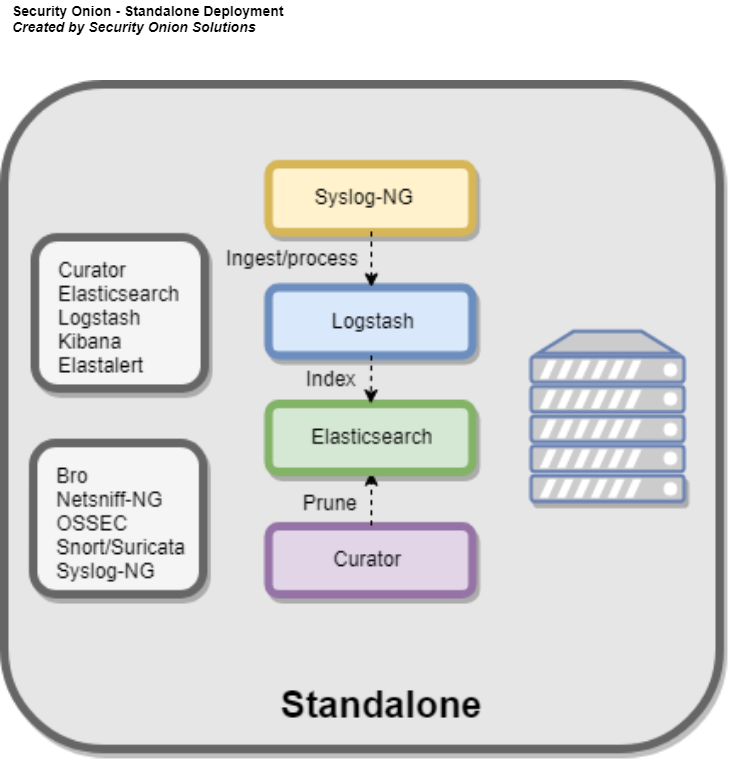
**ELASTICS SEARCH:**

Re-dis uses a searching techniques such as ELASTICS SEARCH

Elastic-search is highly scalable open-source full-text search and analytics engine.

This Elastic-search ES is one such NOSQL distributed database. It is a document-oriented database, retries, designed to store and manage document oriented or semi-structured data.

When Elastic-search it stores data in JSON document form.



**CASSANDRA DATABASE:**

Apache Cassandra has providing high availability with no single point of failures such as

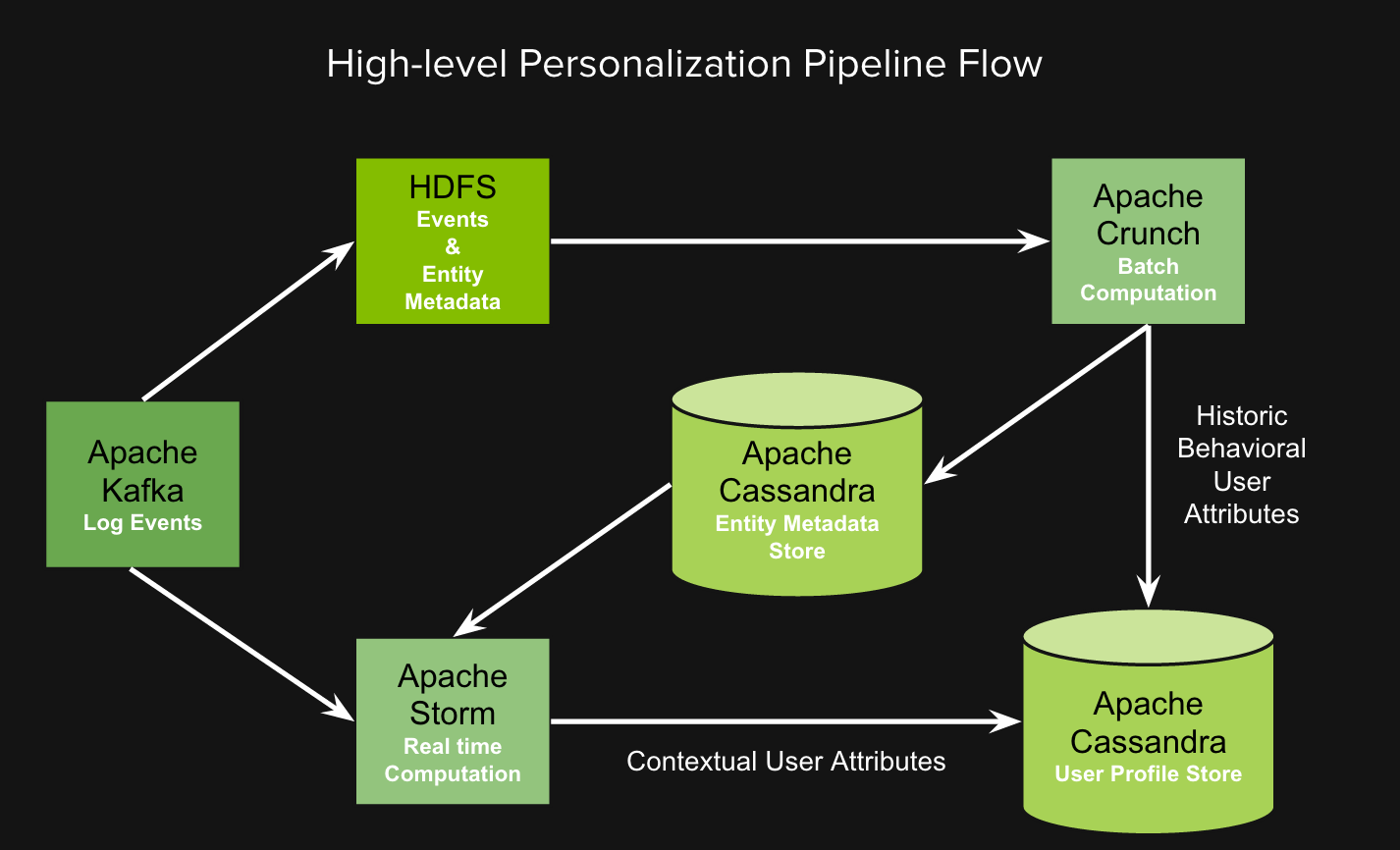
1: one that predicts misfortune or disaster

2: a daughter of Pram endowed with the gift of prophecy but fated to be believed

REASON:

There are five advantage of using Cassandra, they are

* It helps the user to solve complex tasks
* It has high fault tolerance and extreme resilience
* It is very speed



MARIA DB:

Maria DB is an open source software and also provides an SQL interface for Purpose of accessing the data. GIS and JSON features are including in the latest versions of MariaDB.

It is easy to starting the process and collecting the data.

MariaDB is a community developed and supported branch of the MySQL relational Database management system.

It is the open source software under the GNU General Public License.

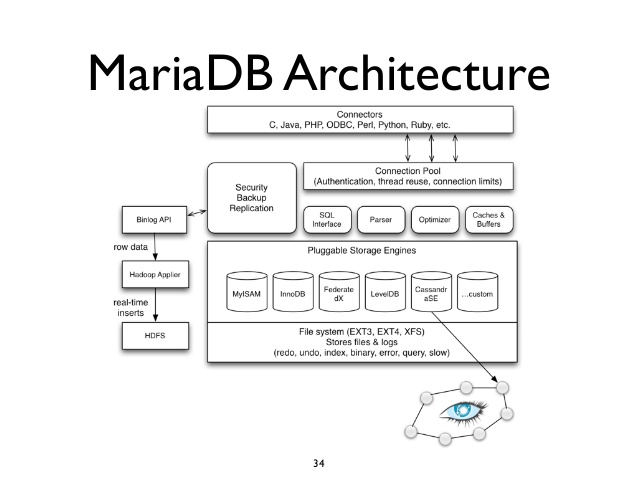
MariaDB databases with hardware capacity of resizable and cost efficient.

MariaDB server is open source relational database popular in worldwide and available In the standard repositories of all major Linux distributions.

MariaDB is stable version receives the bug fixes and security fixes periodically.

MariaDB is using the in Memory Caching easily retrieving the data from the temporarily Storage in the computer's main memory.

It is a secure solution and AWS free tier.

* + MIGRATE TO MARIA DB:
  + Development is more open and vibrant
  + More transparent security releases and Quicker
  + More cutting edge features
  + More storage engines
  + Better performance
  + Galena active-active master clustering
  + Oracle stewardship is uncertain
  + MariaDB has leapt in popularity
  + Compatible and easy to migrate
  + 

Features of MariaDB:

* + licensed under GPL, LGPL, or BSD.
  + includes a wide selection of storage engines and high-performance storage engines For working with other RDBMS data sources.
  + using standard and popular querying language.
  + runs on various operating systems and supports a wide variety of programming Languages.
  + supporting most popular web development languages of PHP
  + Provide Galera cluster technology.
  + MariaDB also offers many operations and commands unavailable in MySQL, and
  + Eliminates/replaces features impacting performance negatively.

MICROSOFT AZURE:

Microsoft Azure is introducing by Microsoft of the cloud computing services.

It is provides the services and supports many different programming languages, tools and frameworks including both Microsoft specific and third party software and system.

Microsoft Azure is perform the operation for building, testing, deploying, and managing Application, services through Microsoft managed data center.

Microsoft Azure is provides the services:

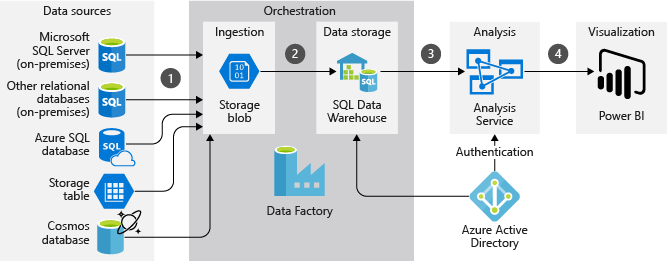
* Infrastructure as a services (Iaas)
  + Platform as s services (Paas)
* Software as a services (Saas)

Microsoft’s cloud strategy has long focused on the kind of hybrid cloud deployments that allow Enterprises to run workloads in a public cloud like Azure and in their own data centers. Azure Stack, its Project for bringing the core Azure services into the corporate data center, is the logical conclusion of This If developers can target a single platform for both the public and private cloud, the thinking goes, Then hybrid deployments become almost trivial.

There are many cloud computing platforms offered by different organizations. Windows Azure is one of them, which is provided by Microsoft.

Azure can be described as the managed data centers that are used to build, deploy, manage the applications and provide services through a global network.

The Services provided by Microsoft Azure are PaaS and IaaS. Many programming languages and Frameworks are supported by it.



FEATURES:

* Build websites with ASP.NET, PHP or Node.js.
* Deploy and run Windows Server and Linux virtual machine.
* Migrate applications and infrastructure.
* SQL Database.
* Caching.
* CDN.
* Virtual Network.
* Mobile Services.

GOOGLE CLOUD STORAGE:

Google cloud storage is an enterprise public cloud storage platform.

Google cloud platform infrastructure is storing and accessing the data and cloud Storage is a restful online file storage web services.

The service combines the performance and scalability of Google's cloud with Advanced security and sharing capabilities.

// It provides unified object storage for live or archived data. Objects stored in GoogleCloud Storage are grouped into buckets.

Google cloud storage is GCP's object data store that means accepts the any quantity Of data and most useful for users.

Google cloud platform provides the services:

* + Infrastructure as a service
  + Platform as a service
  + Server less computing environment

//Google Cloud Storage is GCP's object data store, meaning it accepts any quantity of data and Represents that data to its user in whatever manner is most useful -- for example, as files, a database, a Google Cloud Platform (GCP), offered by Google, is a suite of cloud computing services that runs on The same infrastructure that Google uses internally for its end-user products, such as Google Search and

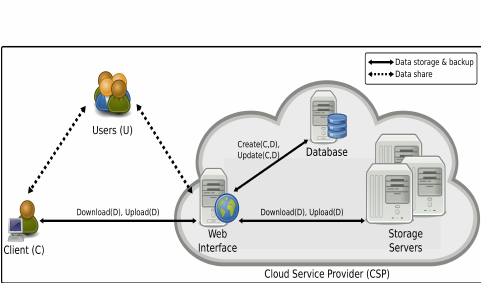
YouTube.[1] Alongside a set of management tools, it provides a series of modular cloud services Including computing, data storage, data analytics and machine learning.[2] Registration requires a credit

Card or bank account details.[3]

Google Cloud Platform provides infrastructure as a service, platform as a service, and server less Computing environments.

Google Cloud services are available in locations across North America, South America, Europe, Asia, and Australia. These locations are divided into regions and zones. You can choose where to locate Your applications to meet your latency, availability, and durability requirements.

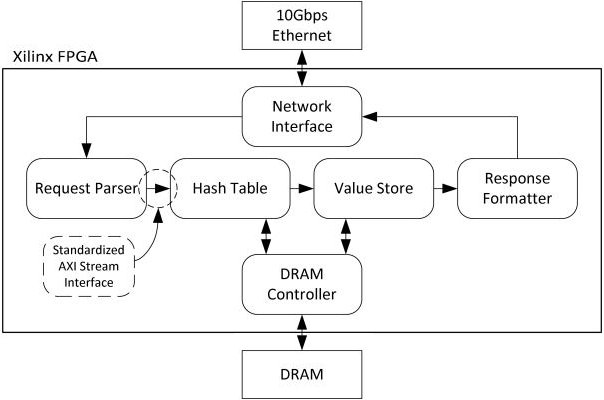
Google Cloud Platform (GCP), offered by Google, is a suite of cloud computing services that runsOn the same infrastructure that Google uses internally for its end-user products, such as Google Search And YouTube.



MEMCACHED DB:

Memcached is an open source, high-performance, distributed memory caching system intended to Speed up dynamic web applications by reducing the database load. It is a key-value dictionary of strings, Objects, etc., stored in the memory, resulting from database calls, API calls, or page rendering.

Memcached was developed by Brad Fitzpatrick for Live Journal in 2003. However, it is now being used By Net log, Facebook, Flickr, Wikipedia, Twitter, and YouTube among others.



Key features of Memcached

* + - * It is open source.
      * Memcached server is a big hash table.
      * It significantly reduces the database load
      * It is perfectly efficient for websites with high database load.
      * It is distributed under Berkeley Software Distribution (BSD) license.
      * It is a client-server application over TCP or UDP.

Memcached is not −

* + A persistent data store
  + A database
  + Application-specific
  + A large object cache
  + Fault-tolerant or highly available

Memcached could be preferable when caching relatively small and static data, such as HTML code Fragments. Memcached’s internal memory management, while not as sophisticated as that of Redis, is More efficient in the simplest use cases because it consumes comparatively less memory resources forMetadata. Strings (the only data type supported by Memcached) are ideal for storing data that is only Read, because strings require no further processing.

Memcached is pronounced as mem-cash-Dee or mem-cached. It is a free, open-source, high- Performance, distributed memory object caching system.

Memcached is used to speed up dynamic web Applications by reducing the database load. Memcached is used by all the major websites having huge Data for example, YouTube, Wikipedia, Twitter etc.

Memcached is used in memory caching software because it is very easy to install on any Windows or UNIX system. It offers API integration for all the major languages like PHP, Java, C/C++, Python, Ruby, Perl etc.

It stores the data based on key values for small arbitrary strings or objects including:

* + Results of database calls
  + API calls
  + Page rendering
  + Components of Memcached

Memcached is made up of 4 main components. These components allow the client and the server to

Work together in order to deliver cached data as efficiently as possible:.

* Client Software: It is used to give a list of available Memcached servers.
* A Client-based hashing algorithm: It chooses a server based on the key.
* Server Software: It is used to store values and their keys into an internal hash table.
* LRU: LRU stands for Least Recently Used. This determines when to throw out old data or reuse Memory.

Features of Memcached

* It is open source.
* It is very scalable; just add boxes with memory to spare.
* Memcached runs as a standalone service. So, if you take your application down, the
* Cached data will remain in memory as long as the service runs.
* Memcached server is a big hash table.
* It reduces the database load.
* It is very efficient for websites with high database load.
* The cache nodes are very ignorant: which means they have no knowledge about other
* Nodes participating. This handles the management and configuration of such a system
* Extremely easy.
* It is distributed under BSD (Berkeley Software Distribution) license.
* It is a client server application over UDP or TCP.

Disadvantages of Memcached

* It is not a fault- tolerant tool.
* Compared to in-memory cache, it is very slow, mostly because of serialization or
* Deserialization and network latency.
* The cache nodes are very ignorant: For example, there is no way to iterate over all of the
* Cached items.
* It is not a persistent data store.

HBASE DATABASE:

HBASE is called the Hadoop database. It is a NoSQL database runs on the top Of the Hadoop.

Base is an open-source non-relational distributed database model.

It is developed as part of Apache Software Foundation&#39;s Apache Hadoop project and runs on top of HDFS or Alluxio, providing big table-like capabilities for Hadoop.

Apache Base is used to random, real-time read/write access to Big Data. It Hosts very large tables on top of clusters of commodity hardware.

// Apache Base is a non-relational database modeled after Google&#39;s Big table. Big table acts up on Google File System, likewise Apache Base works on top of Hadoop and HDFS.

It is highly scalable and can process a form of data types.

Features of HBase

* data replication across clusters.
* HBase is linearly scalable.
* Automatic failure support.
* Provides consistent read and writes.
* Integrates with Hadoop, both as a source and a destination.
* It has easy java API for client.
* It provides

//HBase is called the Hadoop database because it is a NoSQL database that runs on top of

Hadoop. It combines the scalability of Hadoop by running on the Hadoop Distributed File

System (HDFS), with real-time data access as a key/value store and deep analytic capabilities

Of Map Reduce.

//It is developed as part of Apache Software Foundation&#39;s Apache Hadoop project and runs on top of

HDFS or Alluxio, providing Big table-like capabilities for Hadoop

HIVE:

Hive is a data warehouse software project build on top of apache Hadoop and It providing data query and analysis. Hive is an ETL of Hadoop ecosystem and

Used for processing structured and semi structure data.

Hive is a database attending in Hadoop ecosystem perform DDL and DCL

Operations.

 Hive gives a SQL-like interface to query data stored in various databases and file systems that integrate with Hadoop.

// and it provides flexible query language such as HQL for better querying and processing of

Data.

//Apache Hive works by translating the input program written in the hive SQL like

Language to one or more Java map reduce jobs. It then runs the jobs on the cluster toProduce an answer. It functions analogously to a compiler - translating a high level Construct to a lower level language for execution.

Features of Hive

* It stores schema in a database and processed data into HDFS.
* It is designed for OLAP.
* It provides SQL type language for querying called HiveQL or HQL.
* It is familiar, fast, scalable, and extensible.

SPLUNK:

Splunk is an absolutely fast engine and provides lightning fast result.

Splunk is effectively handle big data with no decrease in performance.

Splunk is used for extracting value out of machine-generated data. It can be thought of as a data mining tool for big data applications. The best part of Splunk

Is that it does not need any database to store its data as it extensively makes use of

Its indexes to store the data. The main advantage of using Splunk is that it does not need any database to Store its data, as it extensively makes use of its indexes to store the data.

Splunk is an advanced, scalable, and effective technology that indexes and

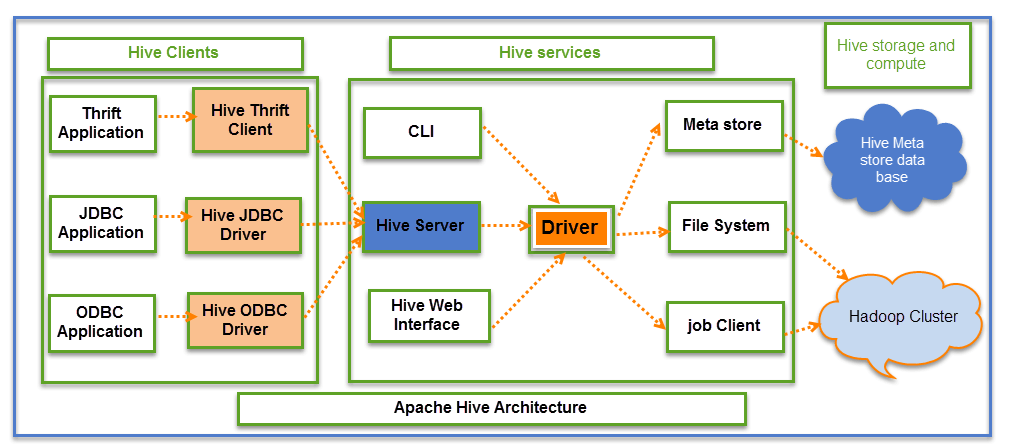
Searches log files stored in a system. The main advantage of using Splunk is that it does not need any database to store its data, as it extensively makes use of its indexes to store the data.

Splunk can be used as monitoring, reporting, analyzing, security information and event Management tool among other things.

Splunk takes the valuable machine-generated Data and converts it into powerful operational intelligence by delivering insights through Reports, charts and alerts.

Splunk is a software mainly used for searching, monitoring, and examining machine- Generated Big Data through a web-style interface.

Splunk performs capturing, indexing, and correlating the real-time data in a searchable container from which it can produce Graphs, reports, alerts, dashboards, and visualizations.za



Features of Hive

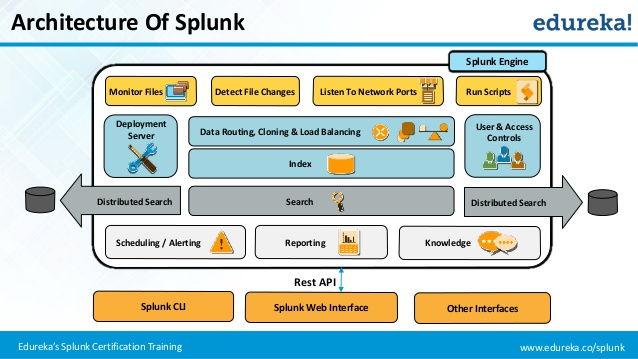
* It stores schema in a database and processed data into HDFS.
* It is designed for OLAP.
* It provides SQL type language for querying called HiveQL or HQL.
* It is familiar, fast, scalable, and extensible.

**SPLUNK:**

Splunk is an absolutely fast engine and provides lightning fast result. Splunk is effectively handle big data with no decrease in performance.

Splunk is used for extracting value out of machine-generated data. It can be thought of as a data mining tool for big data applications. The best part of Splunk is that it does not need any database to store its data as it extensively makes use of its indexes to store the data.

The main advantage of using Splunk is that it does not need any database to store its data, as it extensively makes use of its indexes to store the data.



Splunk is an advanced, scalable, and effective technology that indexes and searches log files stored in a system. The main advantage of using Splunk is that it does not need any database to store its data, as it extensively makes use of its indexes to store the data.

Splunk can be used as monitoring, reporting, analyzing, security information and event management tool among other things. Splunk takes the valuable machine-generated data and converts it into powerful operational intelligence by delivering insights through reports, charts and alerts.

**Splunk** is a software mainly used for searching, monitoring, and examining machine-generated Big Data through a web-style interface. **Splunk** performs capturing, indexing, and correlating the real-time data in a searchable container from which it can produce graphs, reports, alerts, dashboards, and visualizations.

**SLite:**

SQLite is an relational database management system and an embedded system database engine.

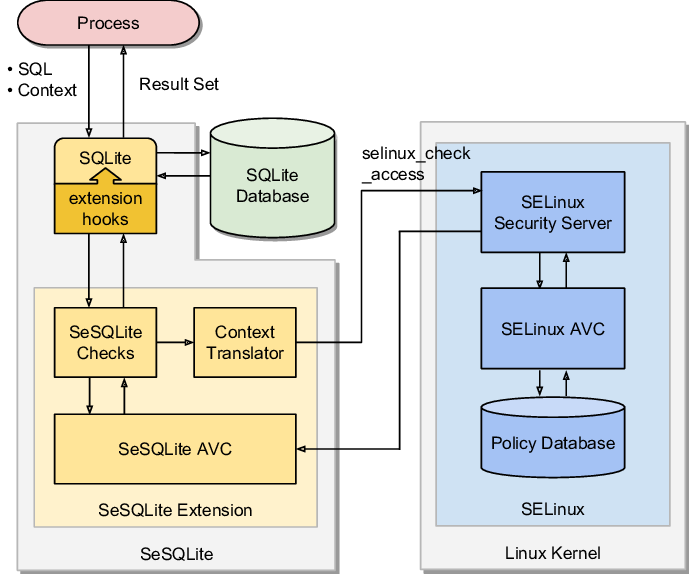
SQLite is contained a C library that’s provides a relational database management system.It is not client server database engine.

SQLite is perform the read and write operation directly to ordinary disk files.

SQLite has take the methods of create,delete,execute SQL commands and perform other common database management tasks.

It does not having any separate server process.

**//** The lite in **SQLite** means light weight in terms of setup, **database** administration, and required resource.



features:

SQLite is an in process library that implements

self-contained

serverless

zero-configuration

transactional.

//The code for SQLite is in the [public domain](https://www.sqlite.org/copyright.html) and is thus free for use for any purpose, commercial or private. SQLite is the [most widely deployed](https://www.sqlite.org/mostdeployed.html) database in the world with more applications than we can count, including several [high-profile projects.](https://www.sqlite.org/famous.html)

Database names must be unique within an application, not across all applications.

**MONGO DB:**

MongoDB stores data in flexible, JSON-like documents, meaning fields can vary from document to document and data structure can be changed over time

The document model maps to the objects in your application code, making data easy to work with

Ad hoc queries, indexing, and real time aggregation provide powerful ways to access and analyze your data

MongoDB is a distributed database at its core, so high availability, horizontal scaling, and geographic distribution are built in and easy to use.

MongoDB is a cross-platform document-oriented database program. Classified as a NoSQL database program, MongoDB uses JSON-like documents with schema. MongoDB is developed by MongoDB Inc. and licensed under the Server Side Public License

## Why Use MongoDB?

* **Document Oriented Storage** − Data is stored in the form of JSON style documents.
* Index on any attribute
* Replication and high availability
* Auto-sharding
* Rich queries
* Fast in-place updates
* Professional support by MongoDB

## Where to Use MongoDB?

* Big Data
* Content Management and Delivery
* Mobile and Social Infrastructure
* User Data Management
* Data Hub
* MongoDB is a document-oriented NoSQL database used for high volume data storage. MongoDB is a database which came into light around the mid-2000s. It falls under the category of a NoSQL database.