# 1 Apache Atlas

Apache atlas is the new scalable platform which includes the core set of the functional governance services.The Apache atlas enables the enterprises to effectively meet the requirements within the hadoop.It also provides the integration of the whole enterprise data ecosystem. The data scientists, data analysts, and the data governance team can benefit from the open metadata management and the governance capabilities can be used for the organizations to develop and create the catalog of their data assets. These assets can be classified and collaborated within the enterprise easily [2].

# 2 Talend

Talend is the open studio tool for the data integration. The java code can be embedded in the existing libraries and along with that we can create our own components in order to leverage the community components and the also code to extend the project. The talend open studio enables to work with the latest cloud applications and platforms or the traditional databases and applications. It also provides the capabilities to design and deploy efficiently with the pre-built components and the connectors [8].

Features Design and productivity tools:

Eclipse-based developer tooling and job designer, ETL and ELT support, versioning, export and execute standalone jobs in runtime environment.

Free open source Apache license.

Data Preparation [5].

Data Quality and Profiling [6].

Components to open, move, compress, and decompress files without scripting, to control data flows with master jobs, and to map, aggregate, sort, enrich, and merge data.

# 3 Zmanda

Zmanda, Inc. delivers cloud backup, open source backup and recovery software and services to small and mid-size enterprises. It offers Amanda Enterprise software which is an open source backup and recovery software to protect Linux, Solaris, Windows, and Mac OS X environments; It also provides the Zmanda Recovery Manager (ZRM) for MySQL for scheduling full and incremental backups of MySQL database [9].

Amanda Enterprise is the world's most cost effective and prevalent commercial open source backup and recovery software. This provides the optimized solution in order to simplify the backup of the diverse systems, databases and the applications. It also securely protects the critical data and enables the IT professionals to guard against the mishaps and helps to respond swiftly to backup and recovery requests tasks [1].

# 4 AppFog

AppFog, CenturyLink's Platform-as-a-Service (PaaS) based on Cloud Foundry, empowers developers to center on writing great cloud-based applications without having to stress around overseeing the basic foundation. The result is expanded deftness and efficiency, more proficient use of resources and low operational overhead.

Instead of investing time provisioning servers, setting up databases, designing web servers or updating firewalls, AppFog clients essentially convey their cloud-native applications to a blazingly quick, resilient, multi-region PaaS. AppFog underpins the most prevalent runtimes and systems, programmed stack adjusting, easy-to-use application scaling, automatic load balancing and more. Platform-as-a-Service guarantees tremendous benefits to engineers who require accelerated time-to-market to meet todays showcase requests, and AppFog delivers [3].

# 5 Appscale

AppScale is an open-source cloud computing stage that consequently sends and scales unmodified Google App Engine applications over open and private cloud frameworks, as well as on-premises clusters. AppScale is modeled on the App Motor APIs and has bolster for Go, Java, PHP, and Python applications.The objective of AppScale is to supply designers with a fast, API-driven improvement stage that can run applications on any cloud infrastructure.AppScale decouples application rationale from its benefit biological system to grant engineers and cloud administrators control over app deployment, information capacity, asset utilization, backup and migration [4].

# 6 OrientDB

The orientDB is a distributed Multi-Model NoSQL database with a Graph Database Engine. The multi-model feature of the orientDB has enabled it to work with the Graph, Document, key-value, geo-Spatial and reactive models. The database provides the support for the searching the data of any type and the user domain modelling supports the object oriented concepts enabling the easy extension.

In addition to the multi-model orientDB has the improved auditing and authentication, password SALT and data-at-rest encryption methods making this a secure Open-Source NoSQL database.The orientDB studio provides the peer-to-peer architecture and new modules such as the teleporter to sync with the relational database and simplify the migrations to the orientDB.

The database can be spanned across the multiple servers. It also additionally supports the multi-master replication on the distributed systems [7].

# References


[1] Amanda enterprise and zmanda recovery manager. Web. URL: http://zmanda.com/webinars.html.

[2] Apache atlas. Web. URL: http://atlas.apache.org/.

[3] Appfog. Web. URL: https://www.ctl.io/appfog/.

[4] Appscale. Web. URL: https://en.wikipedia.org/wiki/AppScale.

[5] Data preparation. Web. URL: https://sourceforge.net/projects/talend-data-preparation.

[6] Data quality and profiling. Web. URL: https://sourceforge.net/projects/talendprofiler.

[7] Orientdb. Web. URL: http://orientdb.com/orientdb/.

[8] Talend open studio for data integration. Web. URL: https://sourceforge.net/projects/talend-studio/.

[9] Zmanda. Web. URL: https://www.crunchbase.com/organization/zmanda.