



University of
South Australia

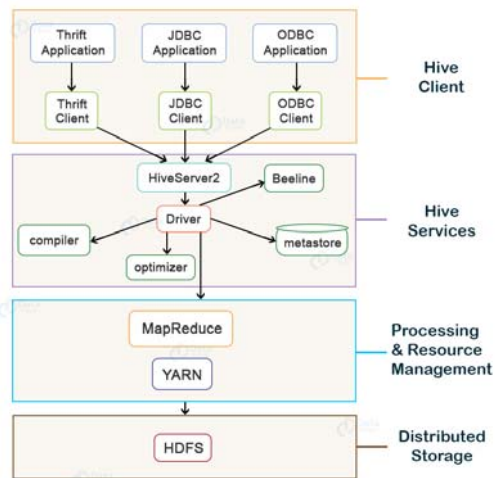
Apache Hive

1



Data warehouse software to
facilitate reading, writing and
managing large datasets residing
in distributed storage using SQL.

2



Hive Architecture & Its Components

<https://data-flair.training/blogs/apache-hive-architecture/>

- We can think of Hive as being split up into 3 or 4 segments
- We're familiar with the processing, resource management and distribute storage systems
- The driver acts like a controller
- The same concepts of compilation and optimization are used in Apache Pig

Hive > Beeline

```
!connect jdbc:hive2://localhost:10000 cloudera cloudera
```

WARNING

This material has been reproduced and communicated to you by or on behalf of the **University of South Australia** in accordance with section 113P of the *Copyright Act 1968 (Act)*.

The material in this communication may be subject to copyright under the Act. Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act.

Do not remove this notice