**Datawarehousing Options**

Open Source Datawarehouse services

Apache Doris

PostgreSQL Datawarehouse

Apache Hive

Apache Kylin

Cloud Datawarehouse Services

Azure Synapse Analytics

AWS Redshift

**Data Lake Options**

Apache Hudi

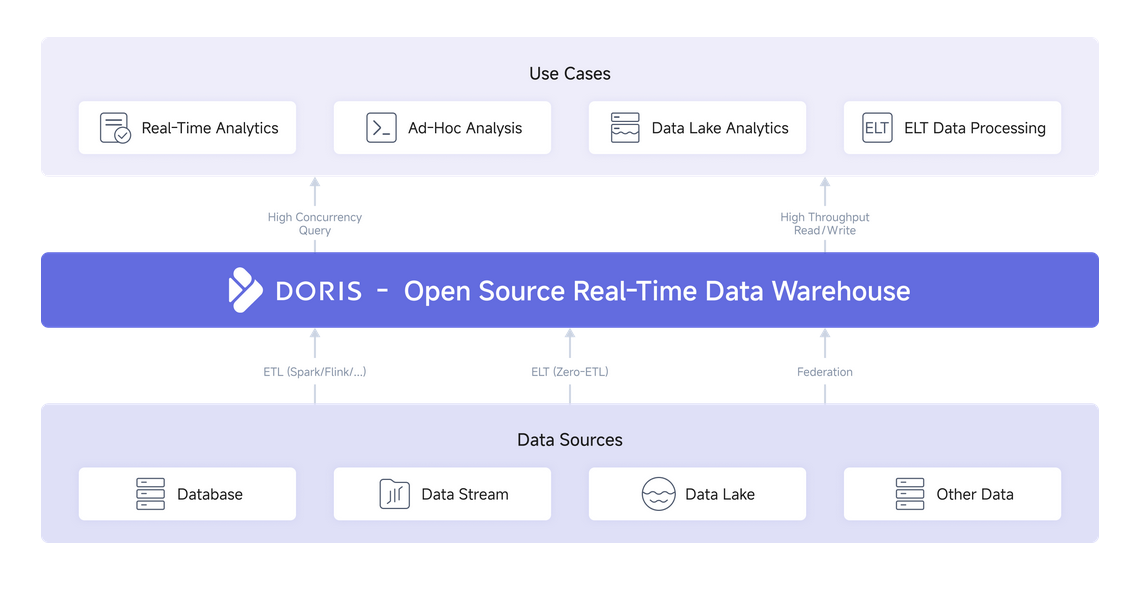
Apache Iceberg

Cloud data lakes

**Apache Doris**

Apache Doris is a high-performance, real-time analytical database based on MPP architecture, known for its extreme speed and ease of use. It only requires a sub-second response time to return query results under massive data and can support not only high-concurrent point query scenarios but also high-throughput complex analysis scenarios. All this makes Apache Doris an ideal tool for scenarios including report analysis, ad-hoc query, unified data warehouse, and data lake query acceleration. On Apache Doris, users can build various applications, such as user behaviour analysis, AB test platform, log retrieval analysis, user portrait analysis, and order analysis.

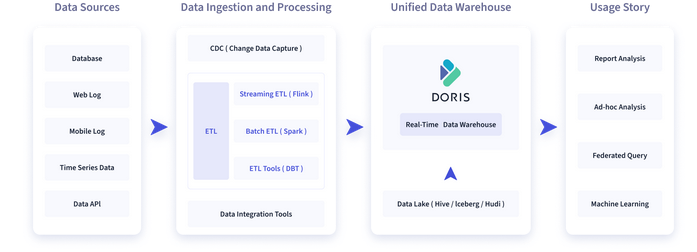
**High Level Architecture**

****

**Capabilities**

* Real time Data Ingestion
* MPP architecture – fast query results
* Data Lake and Database support
* Supports Semi- Structured data
* Open Source

**Usage Scenario**



Real-time dashboards

Ad-Hoc Query

Unified Data Warehouse Construction

Data Lake Query

**Installation Instructions on Kubernetes**

<https://doris.apache.org/docs/install/k8s-deploy/operator-deploy/>

**Apply the** [**Custom Resource Definition(CRD)**](https://kubernetes.io/docs/concepts/extend-kubernetes/api-extension/custom-resources/#customresourcedefinitions)

kubectl apply -f https://raw.githubusercontent.com/selectdb/doris-operator/master/config/crd/bases/doris.selectdb.com\_dorisclusters.yaml

**Install Doris-Operator**

kubectl apply -f <https://raw.githubusercontent.com/selectdb/doris-operator/master/config/operator/operator.yaml>

kubectl apply -f operator.yaml

Start Doris on Kubernetes

1. Initialize Doris Cluster

kubectl apply -f https://raw.githubusercontent.com/selectdb/doris-operator/master/doc/examples/doriscluster-sample.yaml

1. Validate Doris Cluster Status

kubectl get pods

Use Doris Cluster

kubectl -n {namespace} get svc -l "app.doris.ownerreference/name={dorisCluster.Name}

**Use SQL Client for Access**

mysql -h 10.152.183.37 -uroot -P9030