

CS523DE BDT Final Project

Batdelger Otgonbaatar (612330)

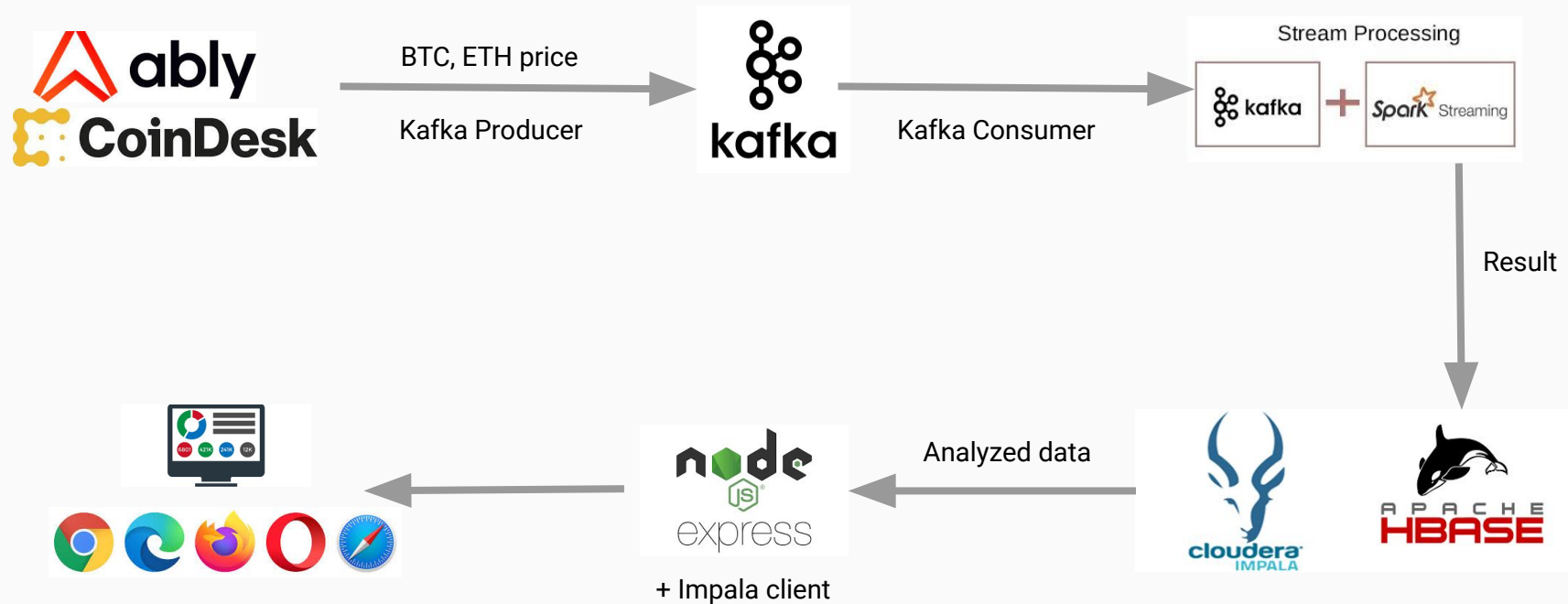


Crypto price real-time streaming

Intro

This is a tiny and simple project designed to demonstrate real-time streaming with big data technology in its whole life cycle. Basically, the whole process of the project is ingesting crypto coin price data from a Crypto Price Streaming service to a big data cluster, then showing a small real-time analysis on the data.

Process flow



Environment

- Ubuntu 18.04 (Google cloud VM)
- Cloudera v5.7 docker
 - Hadoop hdfs v2.6
 - Spark v1.6
 - HBase v1.2
 - Impala v2.5
 - Java v1.8
- Apache Kafka v2.1.1

Github

<https://github.com/batdelger/cs-5-2-3-de>

- Kafka producer
 - Spring Boot kafka + Ably API
- Data ingestion
 - Spark Streaming + Kafka -> HBase
- Client web application
 - Nodejs + express.js + node-impala

Demo

