

Real-time BigBasket

on Feb 6th 2019



About me

- With BigBasket since 2016.
- Platform/Site-Reliability-Engineering Team.
- A constant learner.

My twitter : @none_daa

My email : nandakishore.b@bigbasket.com



Real-time Usecases

- Querying Geo-JOSM data in Real-time.
- Real-time User Behaviour Analytics.
- Real-time View of flow of Data.
- Switching Database connections in Real-time.
- And some (Near) Real-time usecases.



Querying Geo-JOSM data in Real-time

- Given a polygon, how many orders are happening over a period of time ?
- Useful for deciding if hubs are running at optimal capacity.
- Query **Aerospike** (persistent-cache) containing 70 Lakh addresses and get response in a few seconds.



Real-time User Behaviour Analytics

- Know user's behavior in real-time.
 - Useful for segmentation of Users.
 - Useful for suggesting Promos in real-time.
- Client side and server side tracking.
- Using Snowplow, (AWS) Kinesis, Stream-Enrich combo, enriched and processed events data is pushed to (AWS)Redshift for all analytics and business needs.



Real-time View of flow of Data

- Debezium (built on Kafka)
- Turns your existing databases into event streams
- Applications can now see and respond immediately to each row-level change in the databases.



Switching Database Connections in Real-time.

- Mission: Always serve fresh data for the consumer.
 - Latest view of (delivery)slots.
 - Latest order status.
 - Latest view of product sales.
- But databases, especially Replicas, are bound to have lag problems.
- A heavy query on Master, or a heavy query on Replica could be the cause of the problem usually.



(Near) Real-time Usecases

- Compiling the aggregated data of all products as and when changes happen to products.
- High Buying days:
 - Autoscaling of K8s Pods based on [requests count](#) for a given (micro)service.
 - Autoscaling of (AWS)EC2 instances based on [App Response time](#).





We are hiring.