# 👷 Data Engineer - Data Platform Case Study

Overview

BookMyEvent Inc, an online service company that specializes in event planning and management. BookMyEvent provides a platform where users can search for and book event spaces, vendors, and services, all in one convenient location.

Architecture

BookMyEvent entire systems runs on a generic cloud and consists of multiple microservices, written in Python and Golang. These microservices - for instance, the `UserManagement` service, are packaged as Docker containers and run on a Kubernetes cluster. BookMyEvent operations team uses a GitOps approach to build, test and rollout new versions of those microservices. A high level overview of the existing architecture is displayed below:
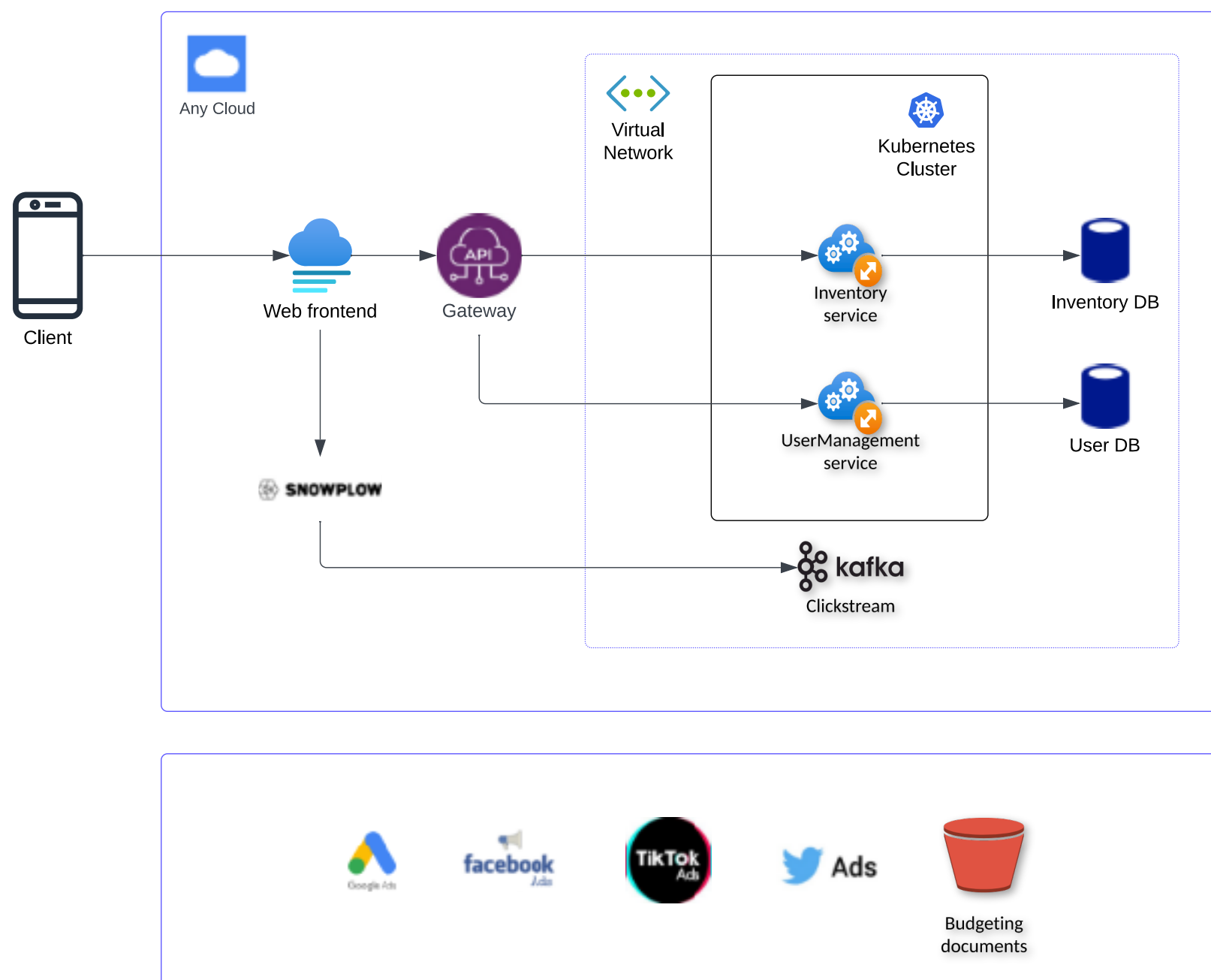
Mobile and desktop browsers can access the storefront at www.bookmyevent.com. The backend services - running on Kubernetes - are served via an API Gateway. The `UserManagement` microservice is written in Python and responsible for users lifecycle and authenticating. The `User DB` contains BookMyEvent customer data. The `Inventory` microservice is written in Go and responsible for the purchasing process. The `Inventory DB` not only contains BookMyEvent products, but also the purchase history.

There are clickstream event from the frontend, which is used for tracking user activity for the further analytics. Those are gathered via `Snowplow` in `Kafka`.

As for the external marketing sources: `Google`, `Facebook`, `Tiktok`, `Twitter Ads`.

Your Job

The current analytics in this company are built on Excel sheets. They would like to have an analytical tool that would allow them to have a reporting system, A/B testing, ad-hoc research and possibly ML models.

Your job is to propose a data platform solution on **any cloud** which can help BookMyEvent Inc better understand their business and make data-driven decisions.

There is a list of soft requirements that you can consider when implementing the solution:

💡 **Soft requirement 1:** Data sources might be huge, so that would be great if the solution can be efficient and scalable.

💡 **Soft requirement 2:** Would be nice if the data that used for the reporting were as consistent as possible.

💡 **Soft requirement 3:** Since most tables in the DBs are not append-only, analysts may want to have their historical changes. Would be great if a solution could help them with this.

Deliverables

BookMyEvent expects a 30 minutes presentation where you use a slide deck to explain the solution architecture, motivate design choices, explain how you tackled the 3 main requirements as well as answer questions from the BookMyEvent team.

Final Remarks

Assumptions If something is vague or unclear in the description above, make reasonable assumptions. You need to document those assumptions and be prepared to explain and defend them.

Presentation You need to bring your own laptop to present the solution architecture. It can be slides only, no implementation necessary. You can use the official cloud icons to create an architecture diagram for your presentation.