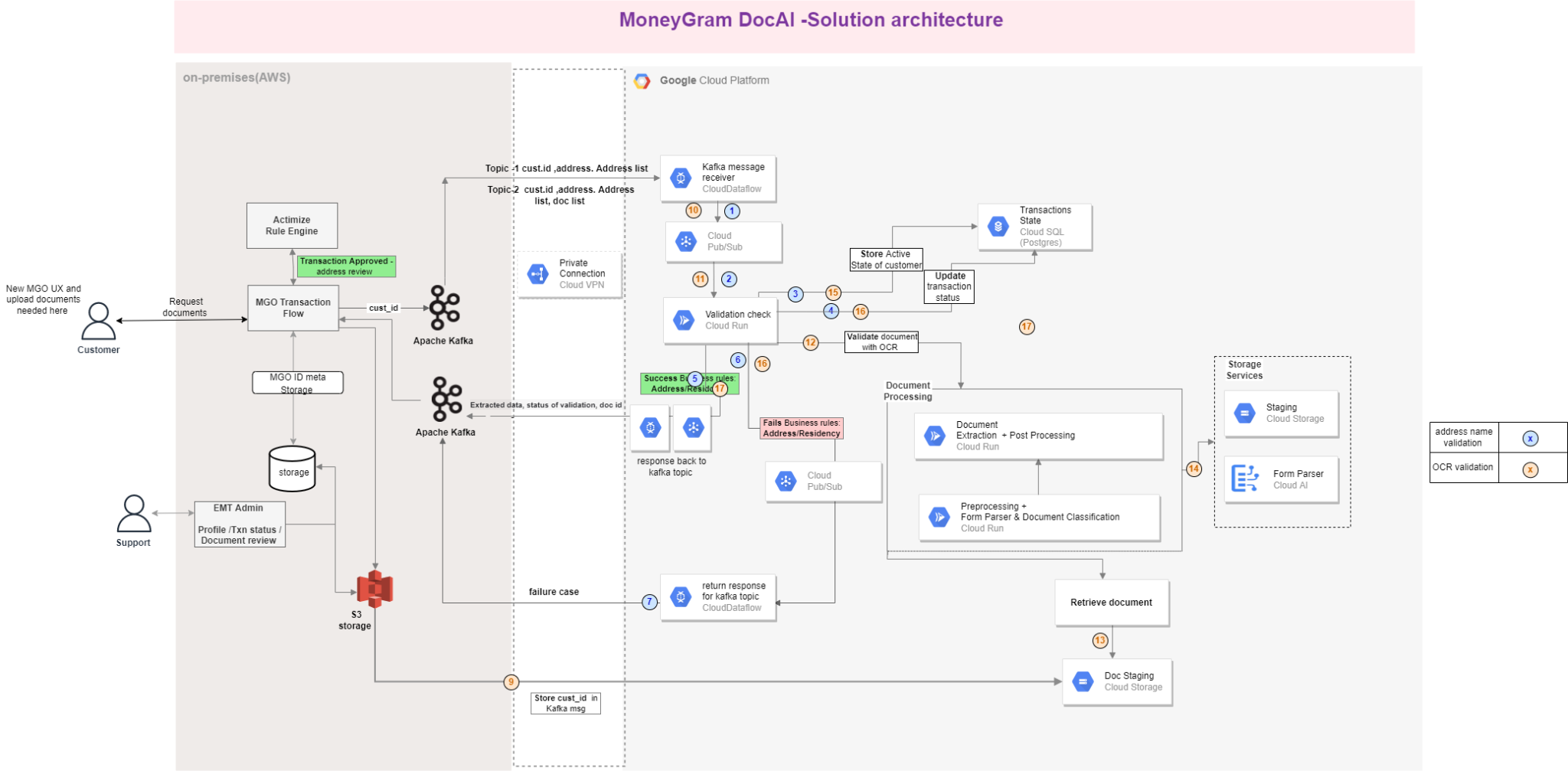
**Dataflow Email**

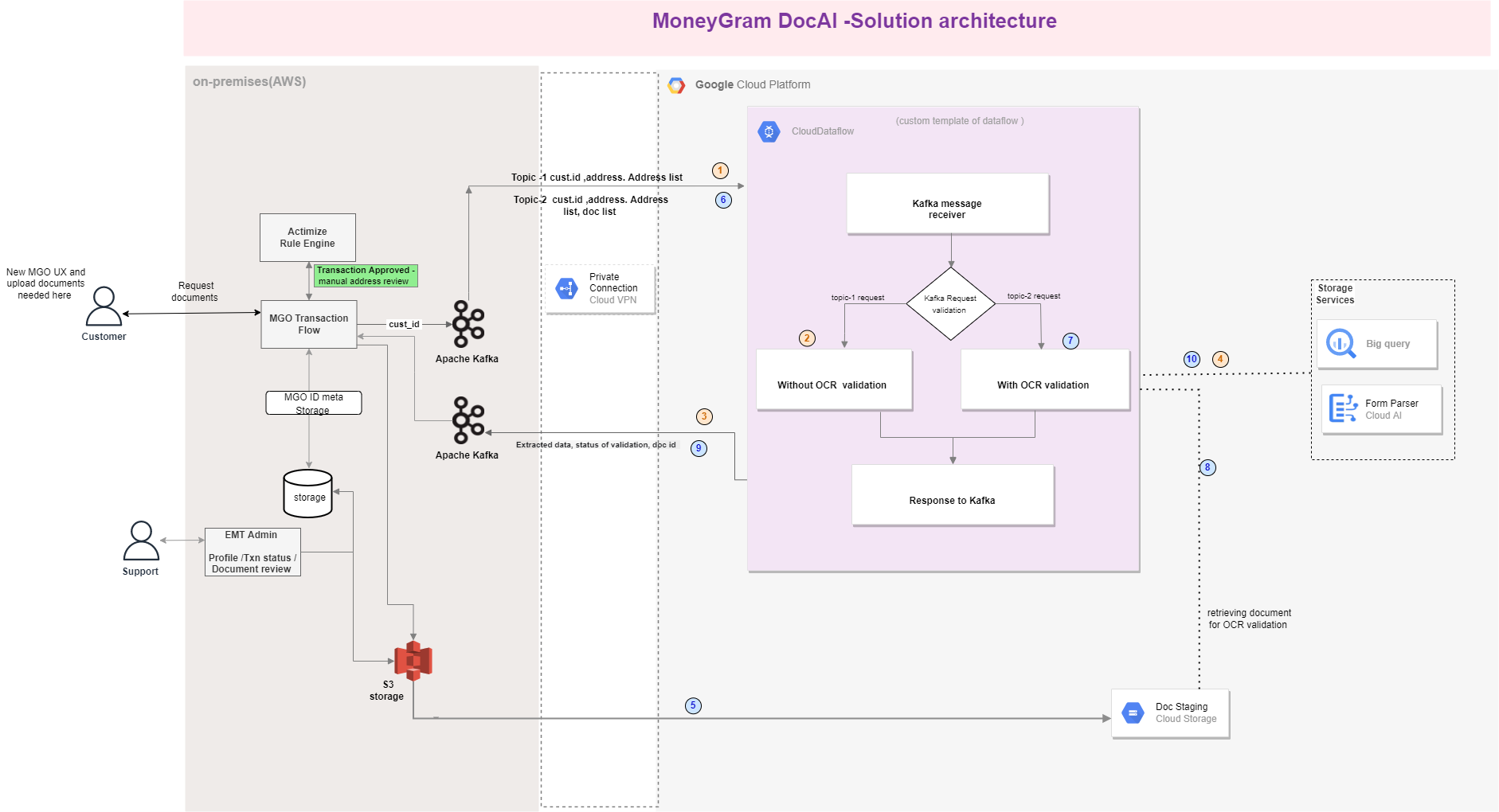
Hi Simran and Anil,

As per discussion on call, please find the final updated solution architecture below:

1. Based on Cloud Run Approach



1. Based on the Data flow Approach (We are aligned on the comment/question shared by Andy above in **purple.** After yesterday’s call,it is incorporated in the below diagram)



**With Dataflow SA steps**

| **Process Number** | **Description** |
| --- | --- |
| 1 | Quantiphi solution receives the customer id and address, address list as a Kafka topic from DataFlow |
| 2 | Validate data for the Customer ID received from dataflow ( existing profile data and OCR data) & update the validation status for customer ID in GCP database |
| 3 | For successful & failed validation, report the status back from cloud dataflow to Kafka, complete the transaction |
| 4 | Store customer state in the GCP storage service |
| 5 | MoneyGram will ingest the documents into a GCS bucket from S3 storage(AWS ) using Lambda |
| 6 | MoneyGram sends the customer id and address ,address list ,and doc\_id list ,as a Kafka topic to dataflow from MoneyGram for OCR validation of customer uploaded document |
| 7 | Trigger the OCR Validate as per the received parameter from cloud dataflow |
| 8 | Based OCR validation preference (like cust \_id , docu\_name) retrieve the document from GCS |
| 9 | For successful & failed validation, report the status back from cloud dataflow to Kafka,(OCR artifacts, decision results ) complete the transaction. |
| 10 | Store customer state in the GCP storage service |

As per our research we have found that Cloud Run is more suitable for ML modeling when custom packages are involved. When we use Cloud Run, we can run our packages as containers without complexities as microservices.

When we use Dataflow we see the following challenges.

1. Integration with other APIs and services using Apache beam SDK libraries
2. Sending back multiple parameters to Kafka as a response to dataflow incoming message based on validation

We would like to have input from Anil to build a solution based on Dataflow.

—----x—----