MX2020 Big Data

GCG Mexico – Architecture Even Hub SourceConnector (Messaging)

June 05, 2019

Versión	Fecha	Descripción del Cambio	Autor/Departamento
1.0	10/07/2019	Creación del documento	[Big Data Architecture]

MX2020 Big Data

Table of Contents

Definición del documento

Purpose

The purpose of this document is to graphically represent the proposed technical solution of Event Hub.

The principles of architecture that are served:

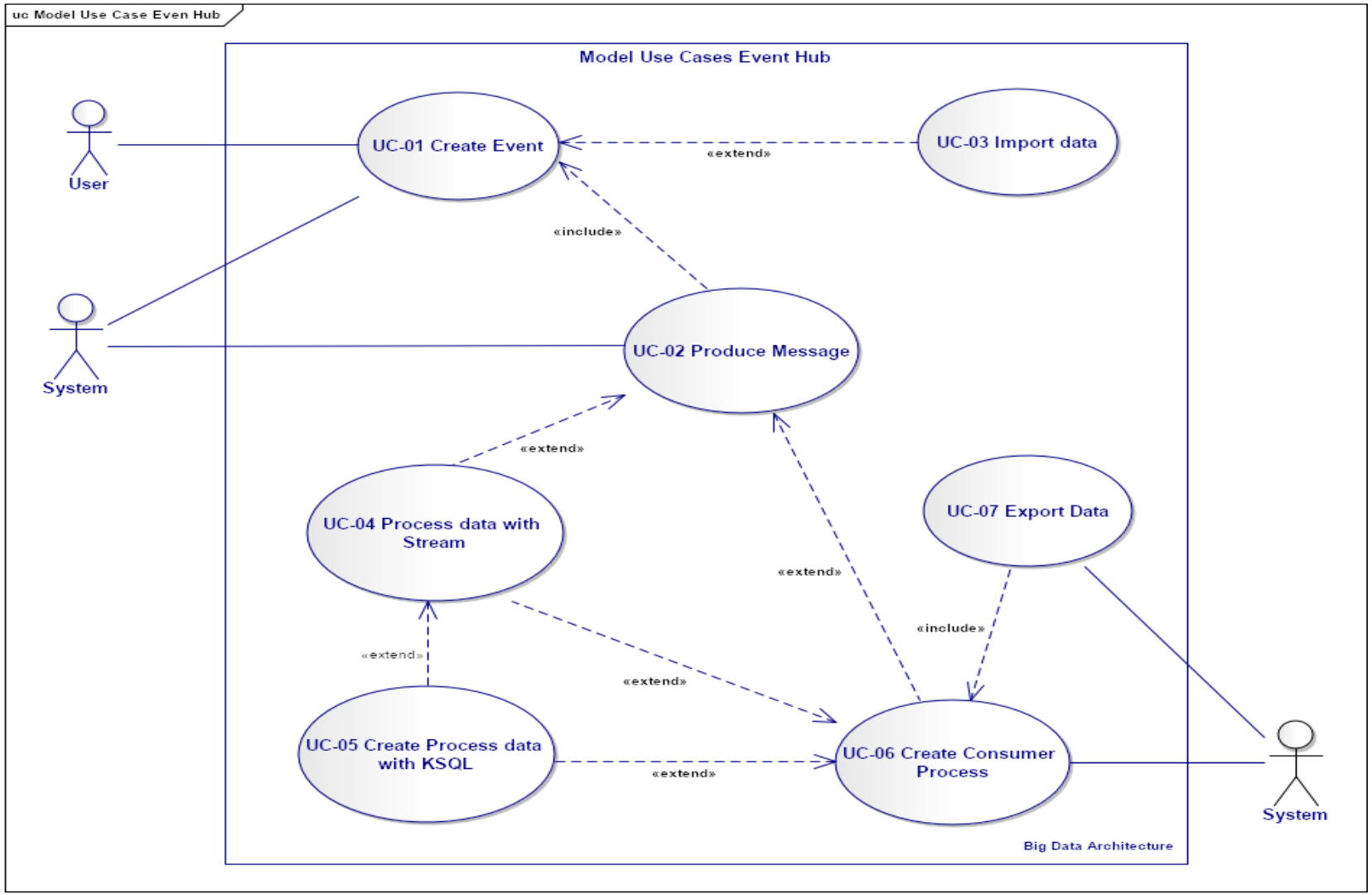
- Separation of responsibilities at the component level, without losing the essence of an Even Hub that is the integration with the dispersed systems, in order to obtain the latest events (data), generate transformations that give value to business in near real time.
- Build real-time data pipes and transmission applications. Integrate data from multiple sources and locations into a single central event transmission platform of Citibanamex.
- Simplify the connection of data sources to Kafka, and generate a common component.

Scope

Define a common structure to multiple sources, projecting a solution as a service

Even Hub

Functional analysis – Model Use Case Event Hub

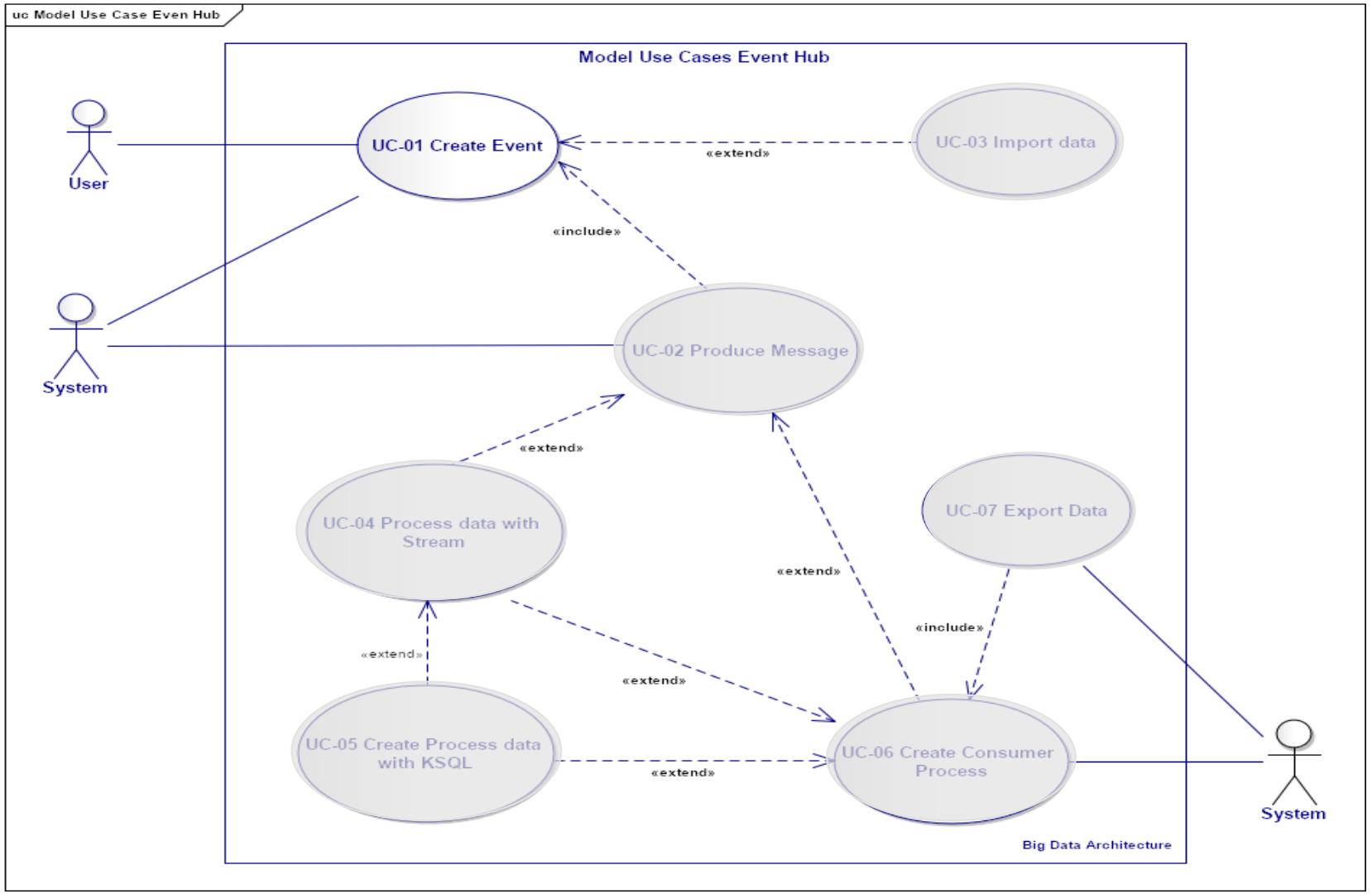


The following diagram describes the objectives of the users, the interactions between the users and the system, and the required behavior of the system to satisfy the functionalities.

CU-01 create Event
CU-02 Import data
CU-03 Produce Message
CU-04 Process data with Stream
CU-05 Process data with KSQL
CU-06 create consumer Process

Even Hub

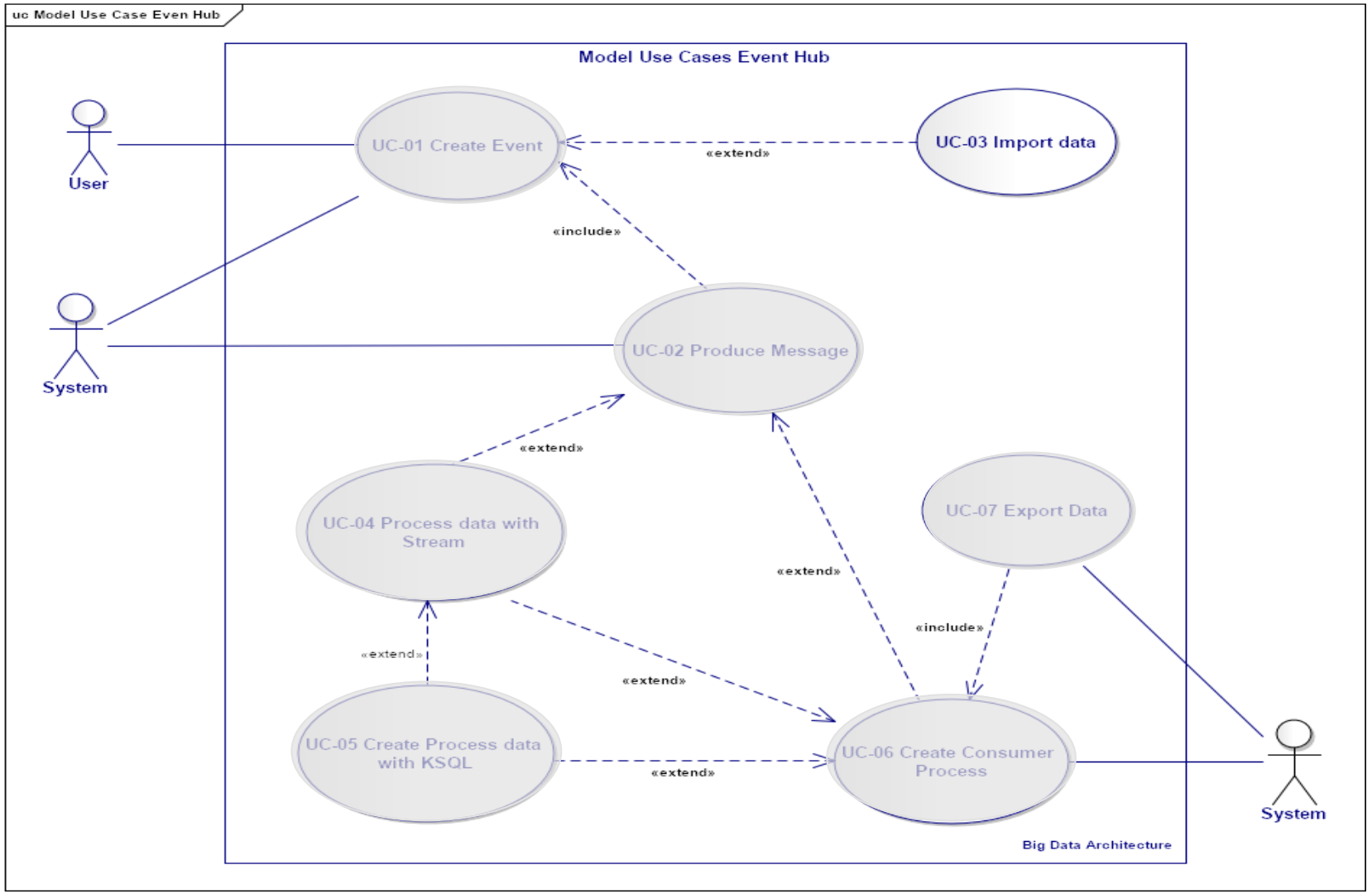
Functional analysis – Use Case 01 Create Event



CU-01 Create Event - Description

Even Hub

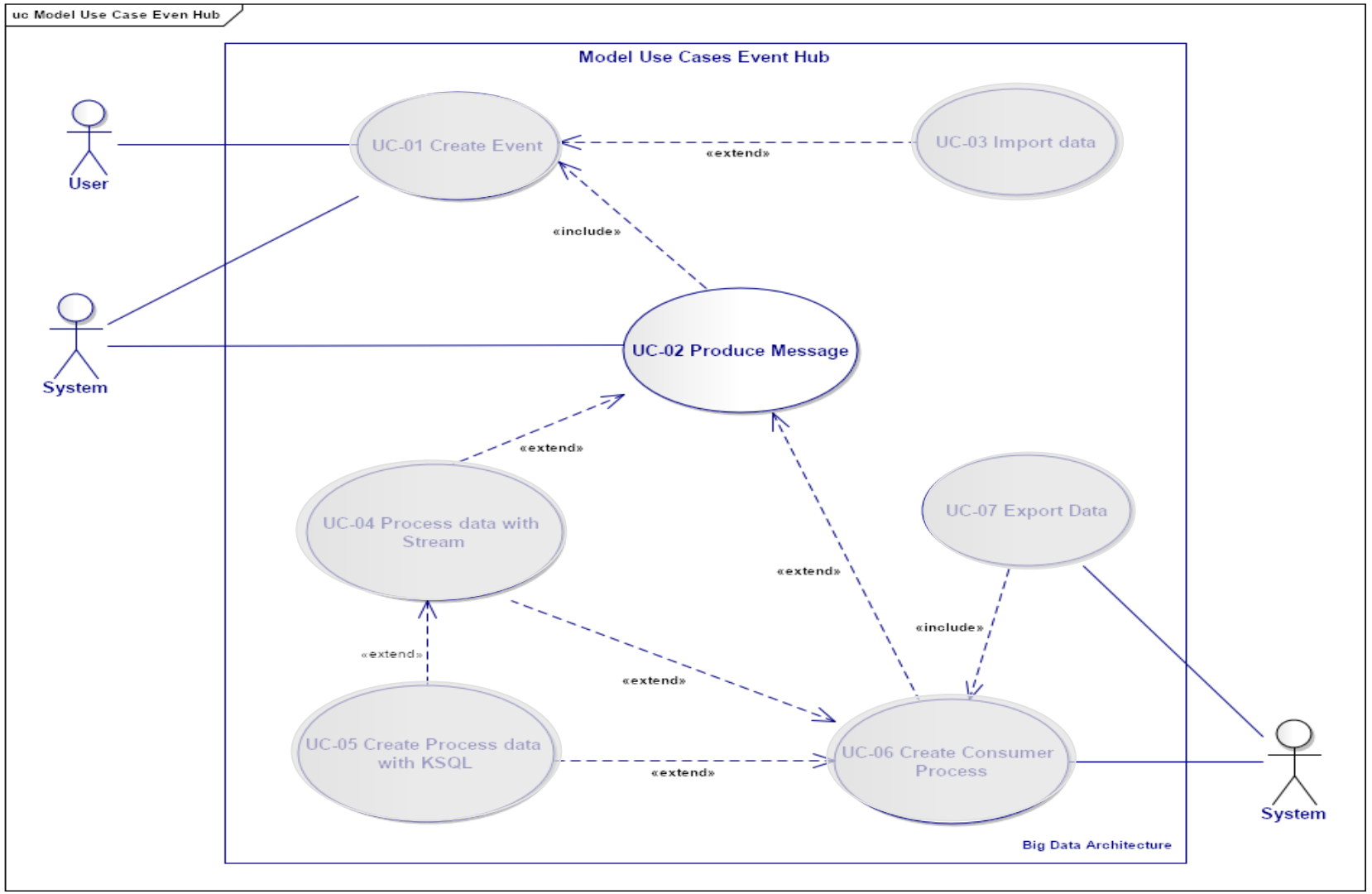
Functional analysis – Use Case 02 Import data



CU-03 create Event - Description

Even Hub

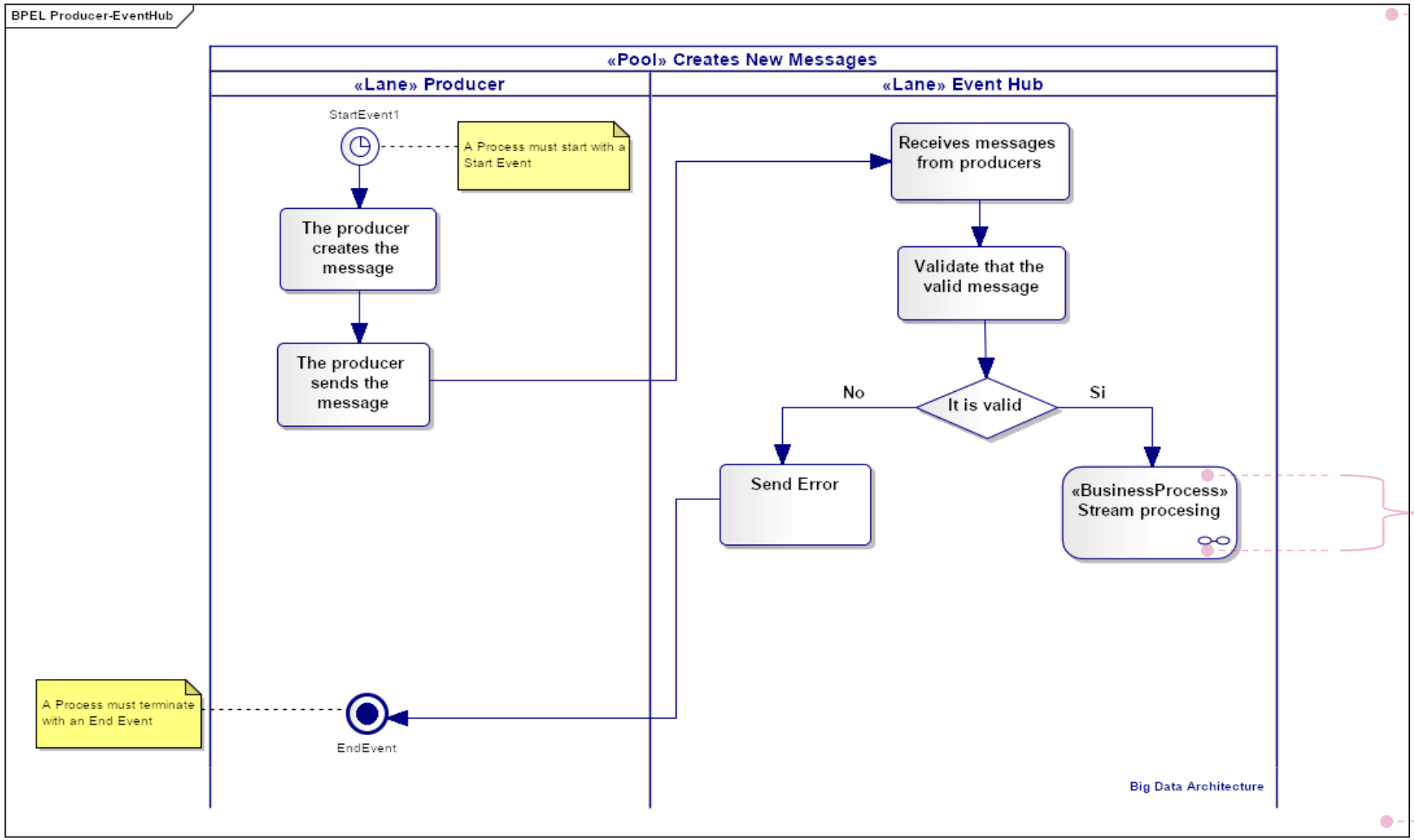
Functional analysis – Use Case 03 Create Produce Message



CU-03 Produce Message - The producer creates new messages. The producer will direct the messages to specific partitions. This is done using the message key and a partitioner that will generate a hash of the key and assign it to a specific partition. This ensures that all messages produced with a given key will be written to the same partition.

Even Hub

Functional analysis – Flow of Use Case Produce Message

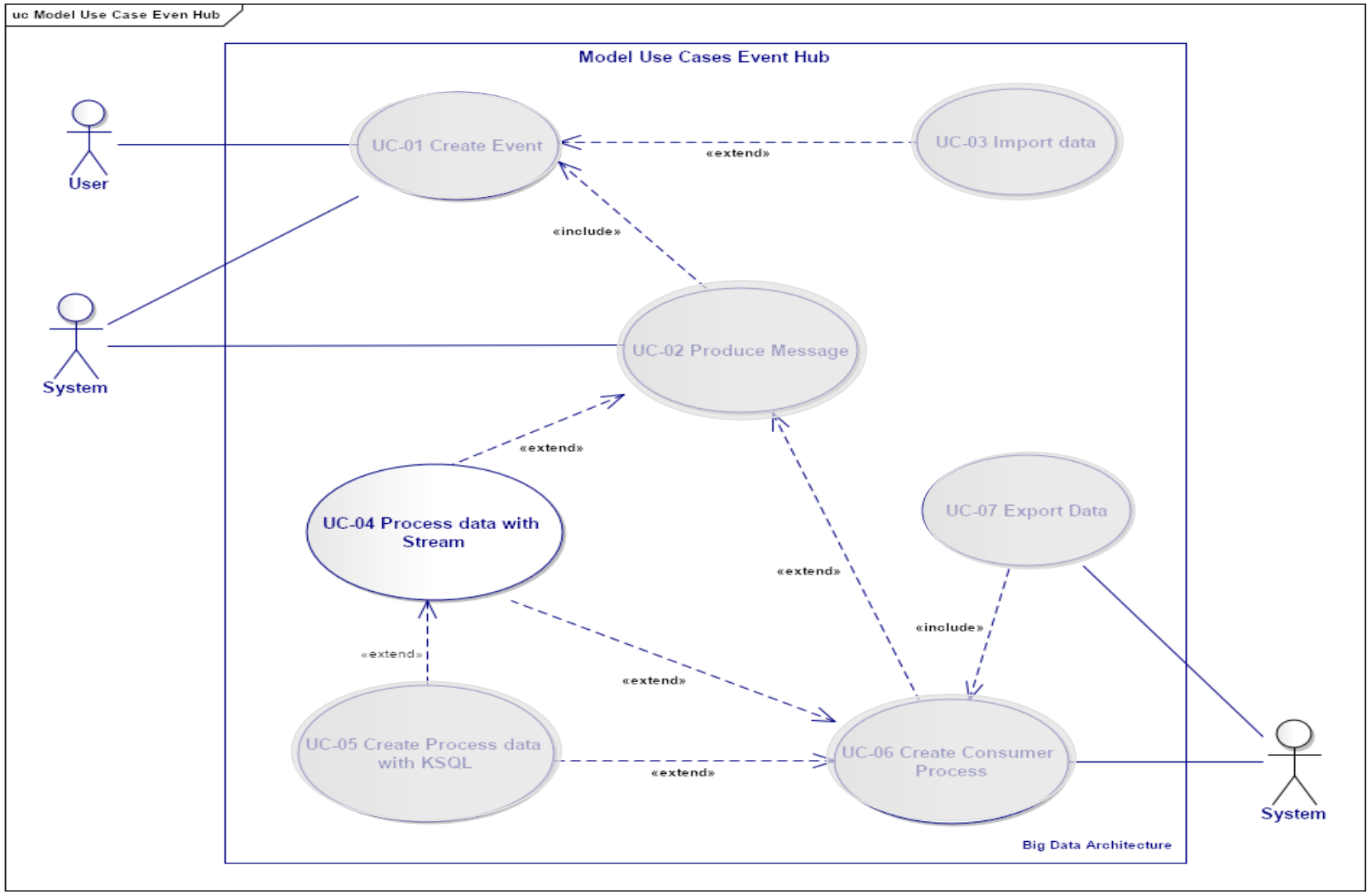


A call is made to the flow of Stream processing.

Describe the logic of the steps in a Produce Message use case process.

Even Hub

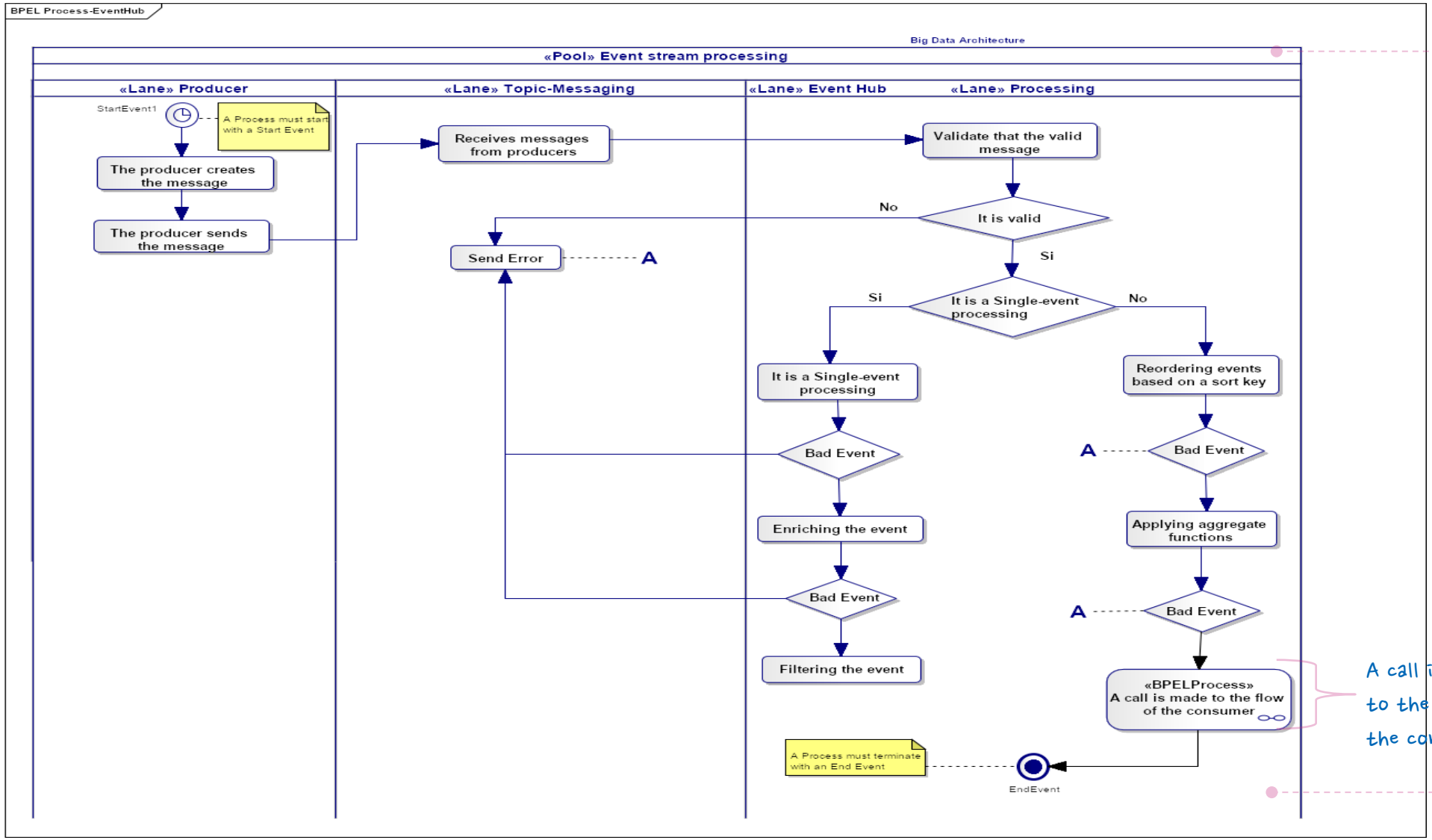
Functional analysis – CU-04 Process data with Steam API



cu-04 Process data with Steam API -

Even Hub

Functional analysis – Flow of Use Case Process data with Stream



Describe the activities of the use case of Process Message.

A call is made to the flow of the consumer

Even Hub

Functional analysis – Event modeling

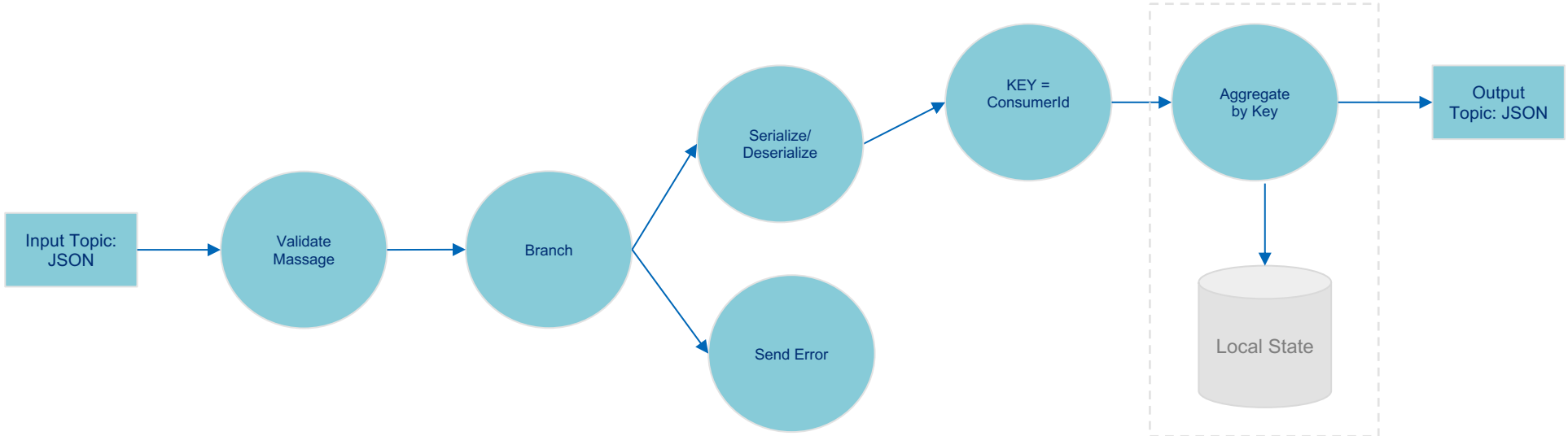
CustumerNumber	20741677
BusinessId	01
CountryId	MX
InputMessageId	MX010087A20190524135426l324
ProcessDate	24-052019 13:54:26
CommunictionType	A
EventId	101
ProductId	103
AccountNumber	5634250213811694
Amount	+55555555555.55
AuthorizationNumber	12345678
MerchantName	Starbucks Portal San Angel
CustumerName	AGUILR,ARREDONDO/ANTONIO
RepId	99
RepName	Antonio Aguilar Arredondo
ChannerlId	EMAIL
AlertValue	armando.antonio.aguilar@gmail.com
TemplateId	37
CommunicationContent	Deposito CUENTA PERFILES M.N. 760 monto $2000.00 el 18/05/19 02:46:00 PM. En operaciones con cheque valida tu saldo antes de realizar cualquier operación.
ProcessStatus	Error
ErrorMessage	Importe inferior al monto mínimo

The structure contains fields other than schema and payload, which is the envelope structure used by the JsonConverter with schemas.enable=true (the default).


```
{"schema": {"type": "struct","fields":
  [{"type": "int64", "optional": false, "field": "custumerNumber"},
   {"type": "int32", "optional": false, "field": "BusinessId"},
   {"type": "int32", "optional": false, "field": "countryId"},
   {"type": "int64", "optional": false, "field": "InputMessageId"},
   {"type": "TIMESTAMP", "optional": false, "field": "ProcessDate"},
   {"type": "int32", "optional": false, "field": "communictionType"},
   {"type": "int32", "optional": false, "field": "EventId"},
   {"type": "int32", "optional": false, "field": "ProductId"},
   {"type": "int64", "optional": false, "field": "AccountNumber"},
   {"type": "String", "optional": false, "field": "Amount"},
   {"type": "int64", "optional": false, "field": "AuthorizationNumber"},
   {"type": "String", "optional": false, "field": "MerchantName"},
   {"type": "String", "optional": false, "field": "custumerName"},
   {"type": "int32", "optional": false, "field": "RepId"},
   {"type": "String", "optional": false, "field": "RepName"},
   {"type": "String", "optional": false, "field": "channelId"},
   {"type": "String", "optional": false, "field": "channelvalue"},
   {"type": "int64", "optional": false, "field": "AlertValue"},
   {"type": "int32", "optional": false, "field": "TemplateId"},
   {"type": "String", "optional": false, "field": "communicationcontent"},
   {"type": "String", "optional": false, "field": "ProcessStatus"},
   {"type": "String", "optional": false, "field": "ErrorMessage"}],
  "optional": false, "name": "ksql.messaging"},
 "payload": {
 "custumerNumber": 20741677,
  "BusinessId": 01,
  countryId: MX,
  InputMessageId: MX010087A20190524135426l324,
  ProcessDate: 24-052019 13:54:26,
  communictionType: A,
  EventId: 101,
  ProductId: 103,
  AccountNumber: 5634250213811694,
  Amount: +55555555555.55,
  AuthorizationNumber: 12345678,
  MerchantName: Starbucks Portal San Angel,
  custumerName: AGUILR,ARREDONDO/ANTONIO,
  RepId: 99,
  RepName: Antonio Aguilar Arredondo,
  channerlId: EMAIL,
  AlertValue: armando.antonio.aguilar@gmail.com,
  TemplateId: 37,
  communicationcontent: Deposito CUENTA PERFILES M.N. 760 monto $2000.00 el 18/05/19 02:46:00 PM.
En  operaciones con cheque valida tu saldo antes de realizar cualquier operación.,
  ProcessStatus: Error,
  ErrorMessage: Importe inferior al monto mínimo}}
```

Even Hub

Global Architecture – Graph (topology) of the Massaging App

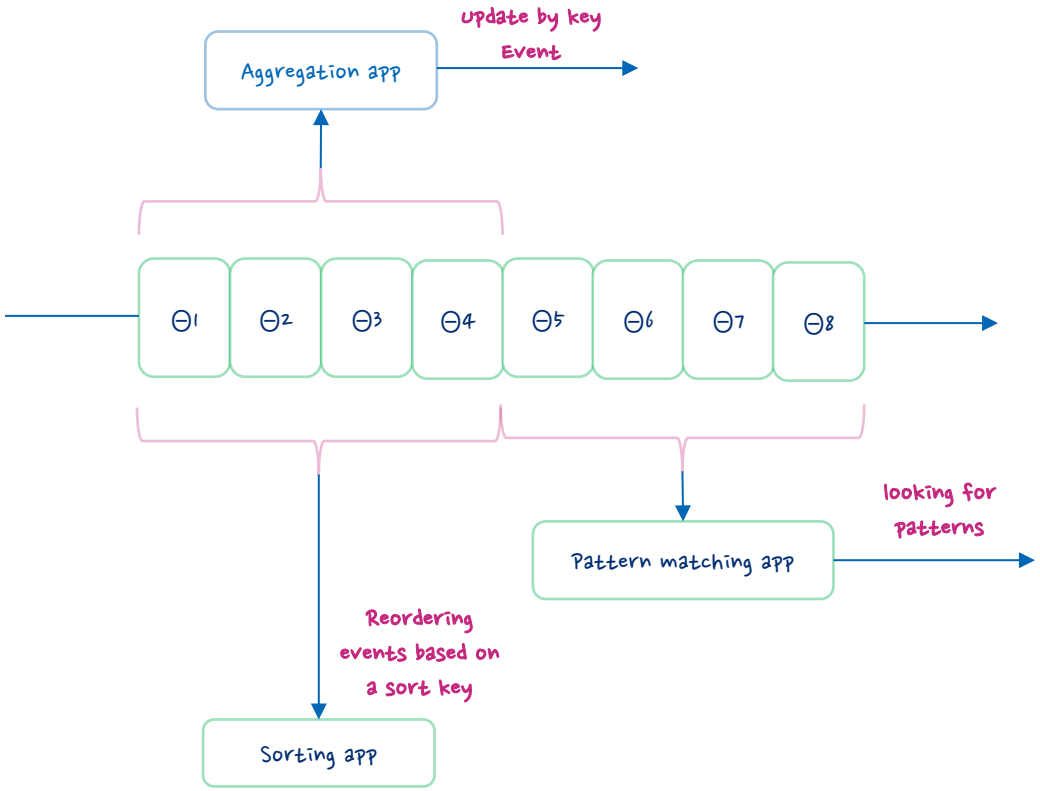


Even Hub

Functional analysis – Multiple-event processing

For the processing of the transmission it is visualized that the multiple events of a client are given the task of reading and processing the information presented in the following scenario, to generate a type of output.

- Use Aggregating
- Pattern matching
- Sorting

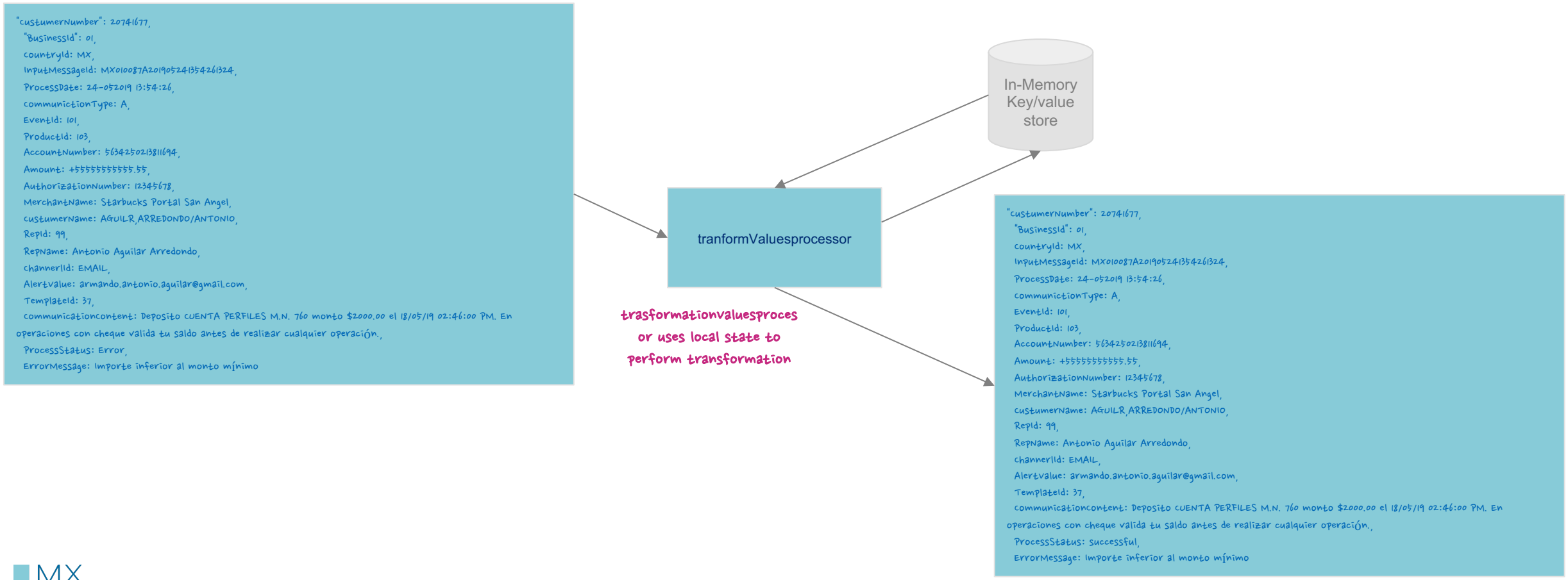


Even Hub

Functional analysis – Streams and state

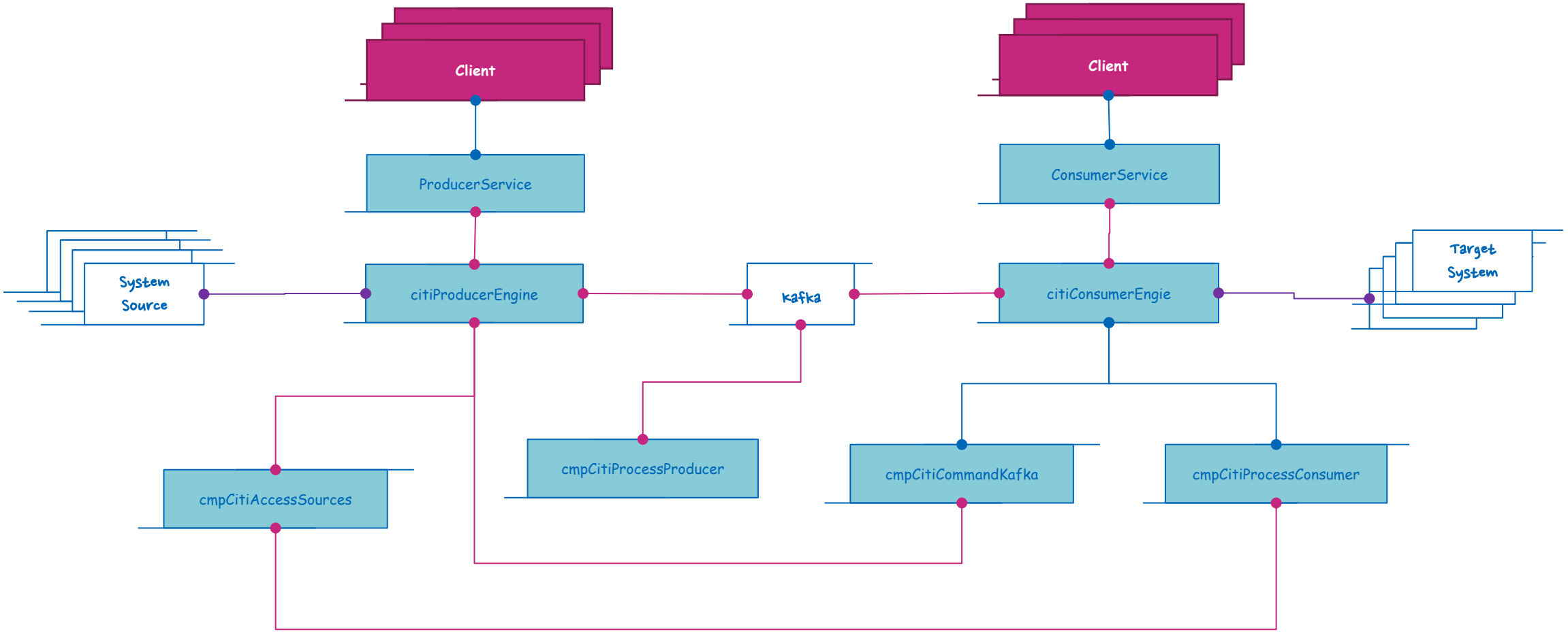
We contemplate the use of state to Streams Processing, to follow up each client updated the status of the mail. The consumer verifies the client's status and sends the results to the Data Lake.

KStream.transformValues, will allow us to have states in the processes.



Even Hub

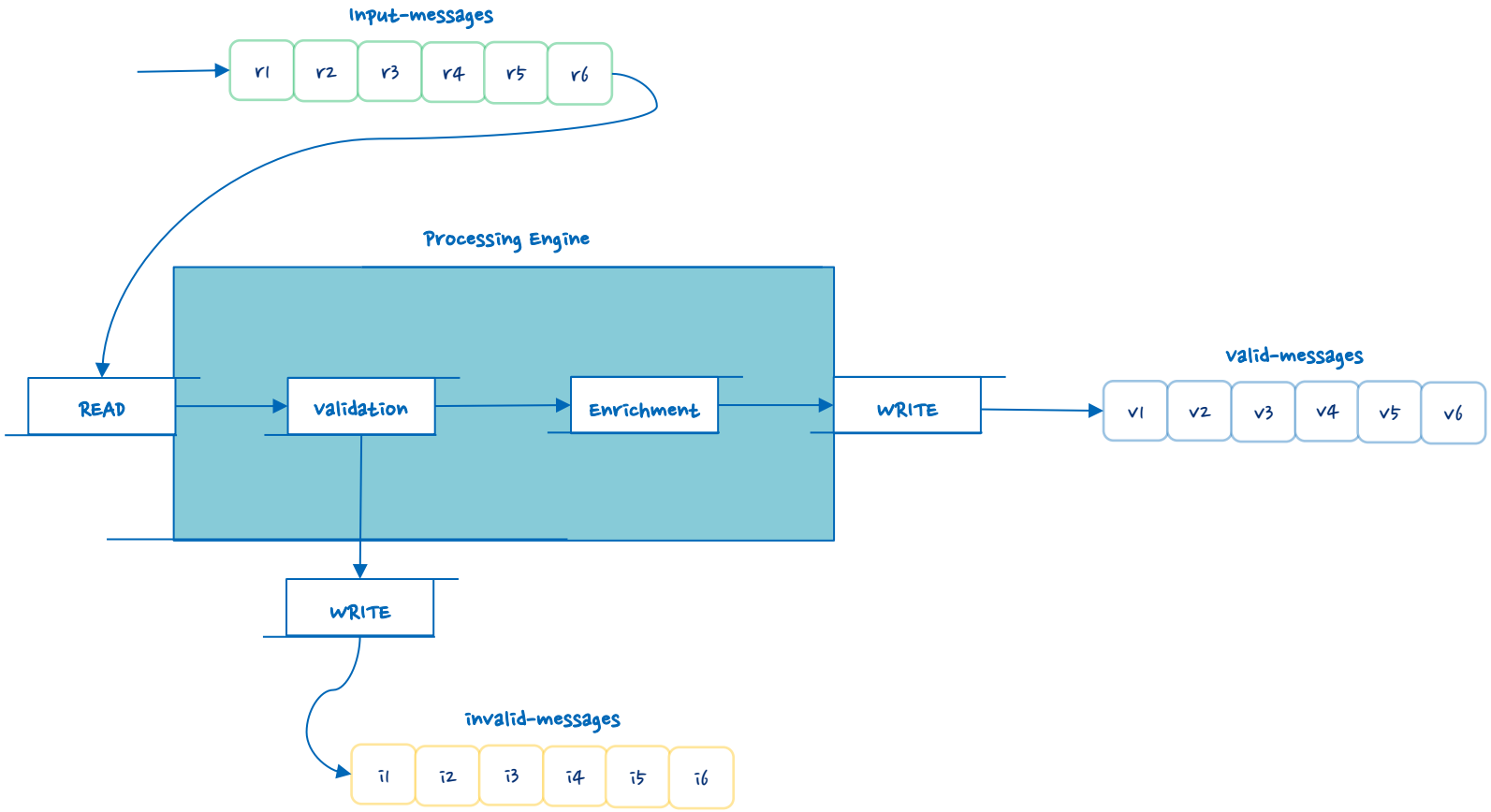
Global Architecture



Even Hub

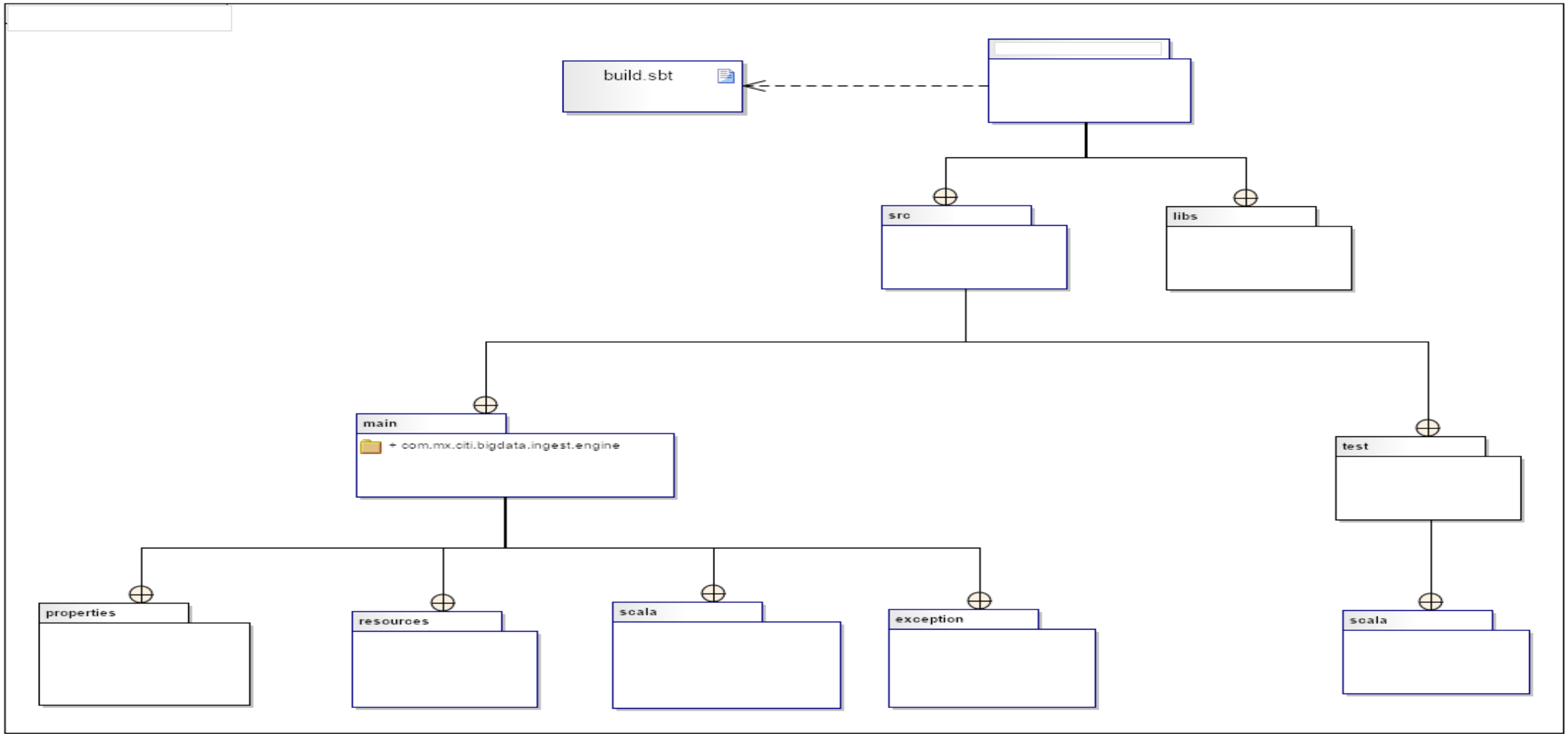
Global Architecture - The steps of the architecture for the processing engine

- Read the individual events of a Kafka theme (input messages)
- Validate the message and send any defective event to a topic dedicated to Kafka called invalid messages
- Write the enriched messages in a Kafka theme called valid messages



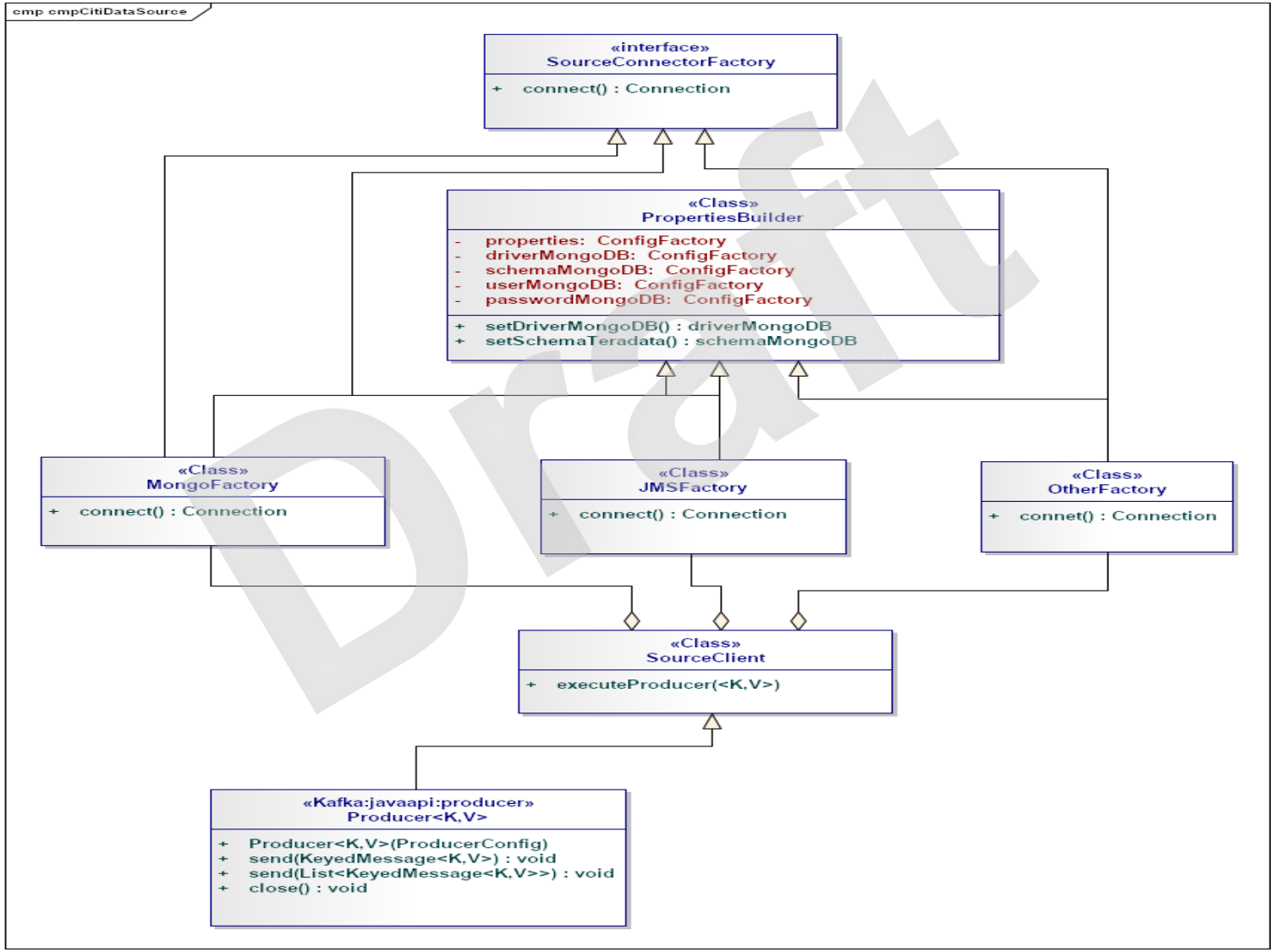
Even Hub

Architecture Internal Structure Of cmpCitiDataSource



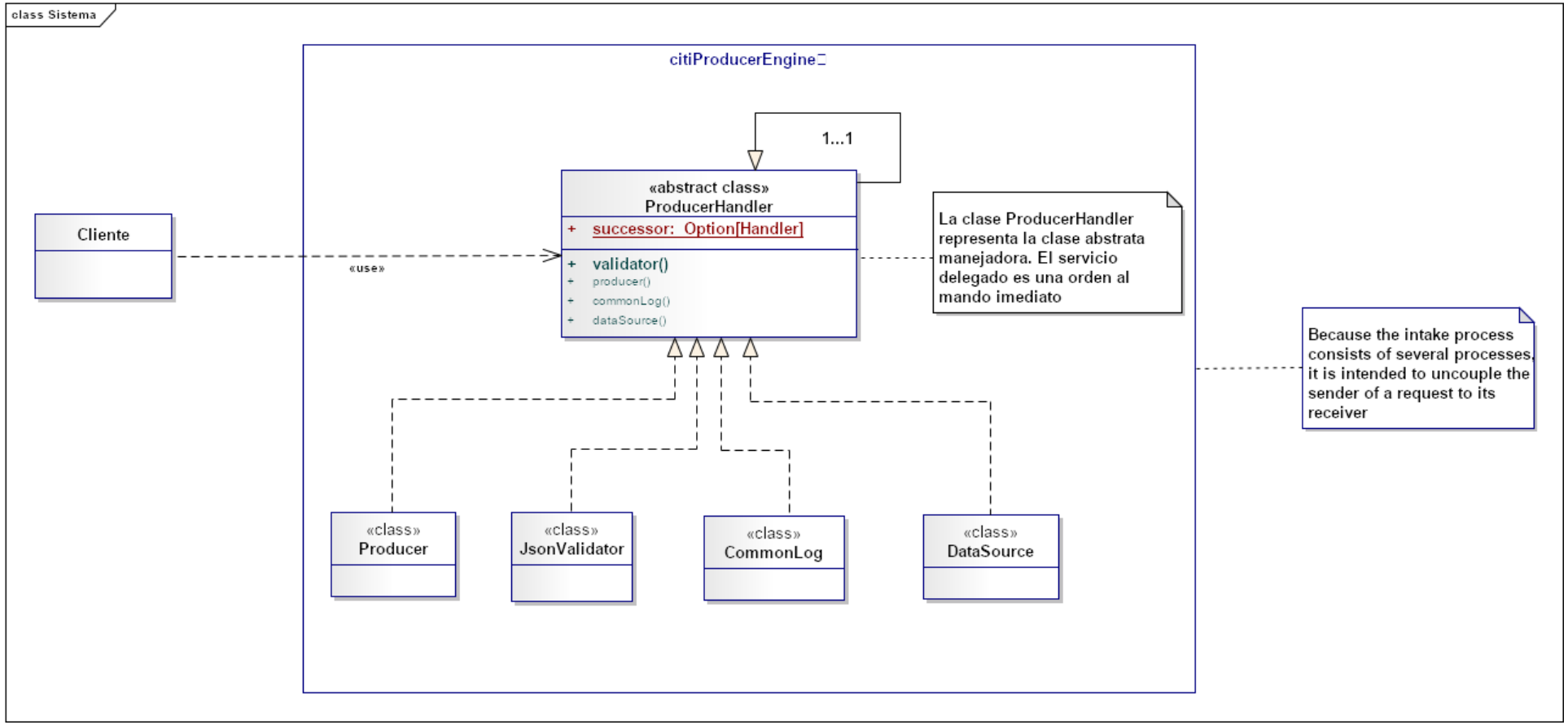
Even Hub

Architecture Internal Structure Of cmpCitiDataSource



Even Hub

Architecture Internal Structure Of citiProducerEngine



Big Data

Global Architecture