

EVENT DRIVEN STREAMING ANALYTICS

Demonstration on Architecture of IoT

Lei Xu 2016.09

I.DEMO SCENARIO

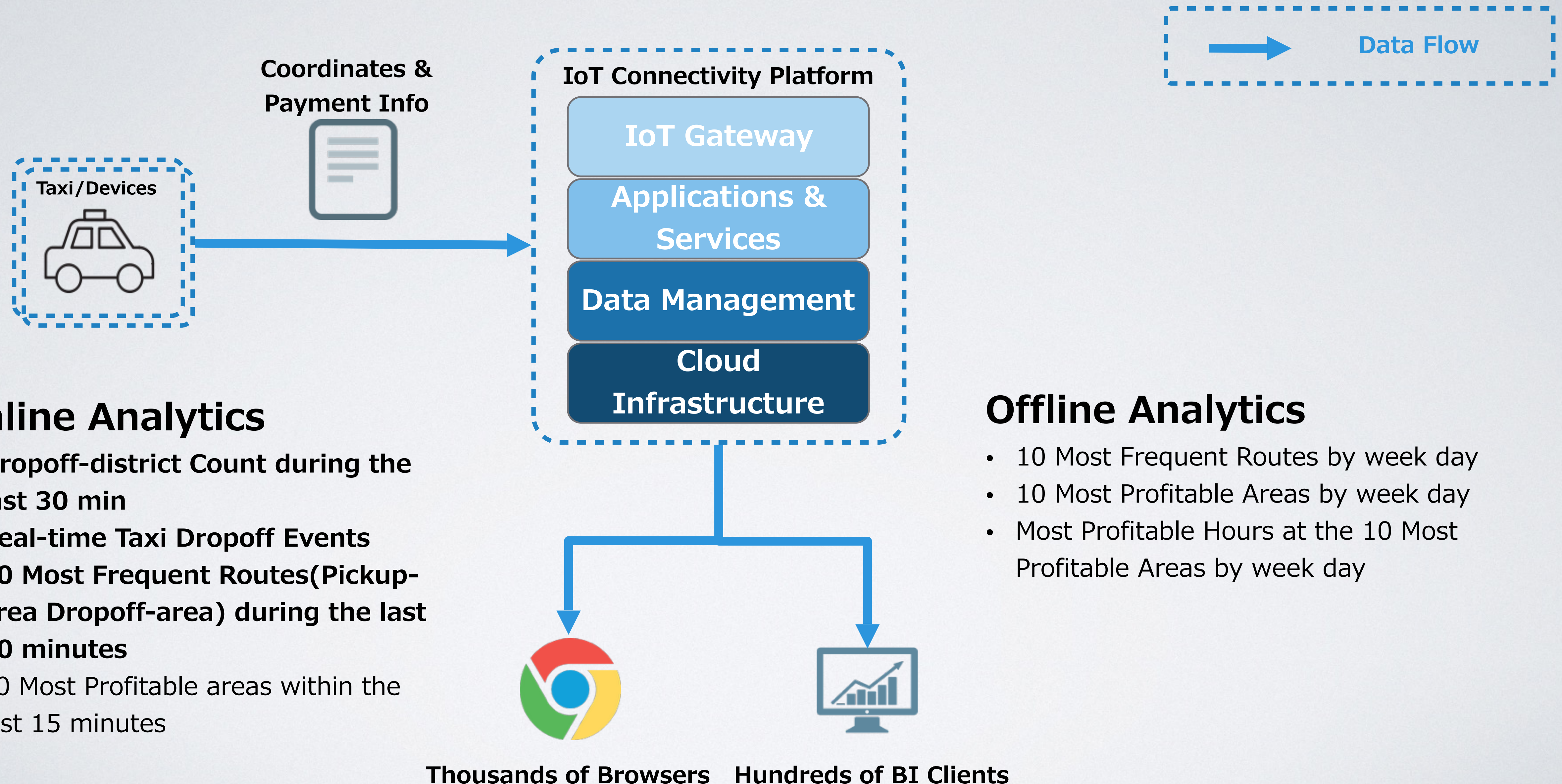
Devices generate transactions after payment

Coordinates & Payment Info



Fields	Descriptions	Sample Data
medallion	an md5sum of the identifier of the taxi - vehicle bound	22D70BF00EEB0ADC83BA8177BB861991
hack_license	an md5sum of the identifier for the taxi license	3FF2709163DE7036FCAA4E5A3324E4BF
pickup_datetime	time when the passenger(s) were picked up	2016-08-11 00:02:00
dropoff_datetime	time when the passenger(s) were dropped off	2016-08-11 00:05:00
trip_time_in_secs	duration of the trip	300
trip_distance	trip distance in miles	5400
pickup_longitude	longitude coordinate of the pickup location	139.706348
pickup_latitude	latitude coordinate of the pickup location	35.576938
dropoff_longitude	longitude coordinate of the drop-off location	139.76401
dropoff_latitude	latitude coordinate of the drop-off location	35.6651362
payment_type	the payment method - credit card or cash	CSH
fare_amount	fare amount in currency	2000
surcharge	surcharge in currency	0
mta_tax	tax in currency	160
tip_amount	tip in currency	0
tolls_amount	bridge and tunnel tolls in currency	0
total_amount	total paid amount in currency	2160

Transaction being sent to IoT Platform for Analytics



IoT Platform Components

Apps & Services

Rule Engine & Process Mgmt.

Complex Event Processing

Business & Data MicroServices

IoT Gateway

Device Management

Authentication & Authorization

Gateways & Interfaces

Data Management

Multi-Channel Data Integration

Structured & Unstructured Data

Analytics & Machine Learning

Visualize Data

Reporting, BI & User Apps

History Data & Real-time Data

Multiple Devices & Platforms

Cloud Infrastructure

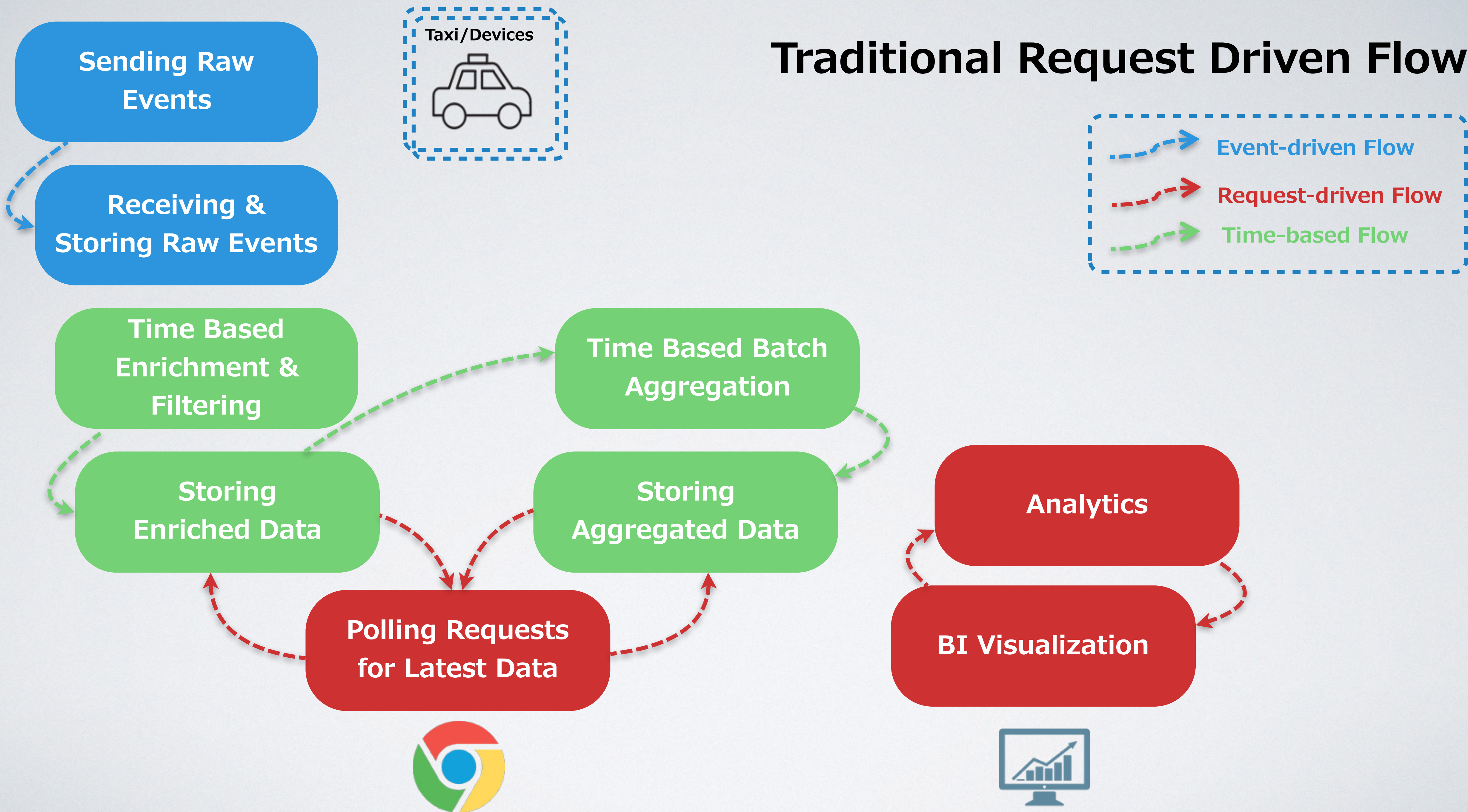
Cloud Native Platform

Scalability & High Availability

Security & Governance

2.REQUEST DRIVEN FLOW

Traditional Request Driven Flow



DRAWBACKS & CHALLENGES

- Polling + Request/Response Style => Overhead
- Time Based Batch Process => High Latency
- Architecture - Monolith or SOA will work well with HA?
- How to scale & how to maintain system performance with lots of user requests?
- How to scale & how to maintain system performance with huge amount of data?
- Portable for both Cloud and On-Premise environments?

REQUEST DRIVEN & EVENT DRIVEN

	Request Driven	Event Driven
How is an action being taken?	As a response to a specific request	Triggered by the fact of a specific situation
When is an action being taken?	When the request is being processed	Determined by the context of the situation
What happens when the event/request occurs?	A response is produced	The event can be ignored, increment the state, trigger an internal derive event, or trigger a solution

<http://epthinking.blogspot.jp/2012/12/more-on-request-driven-vs-event-driven.html>

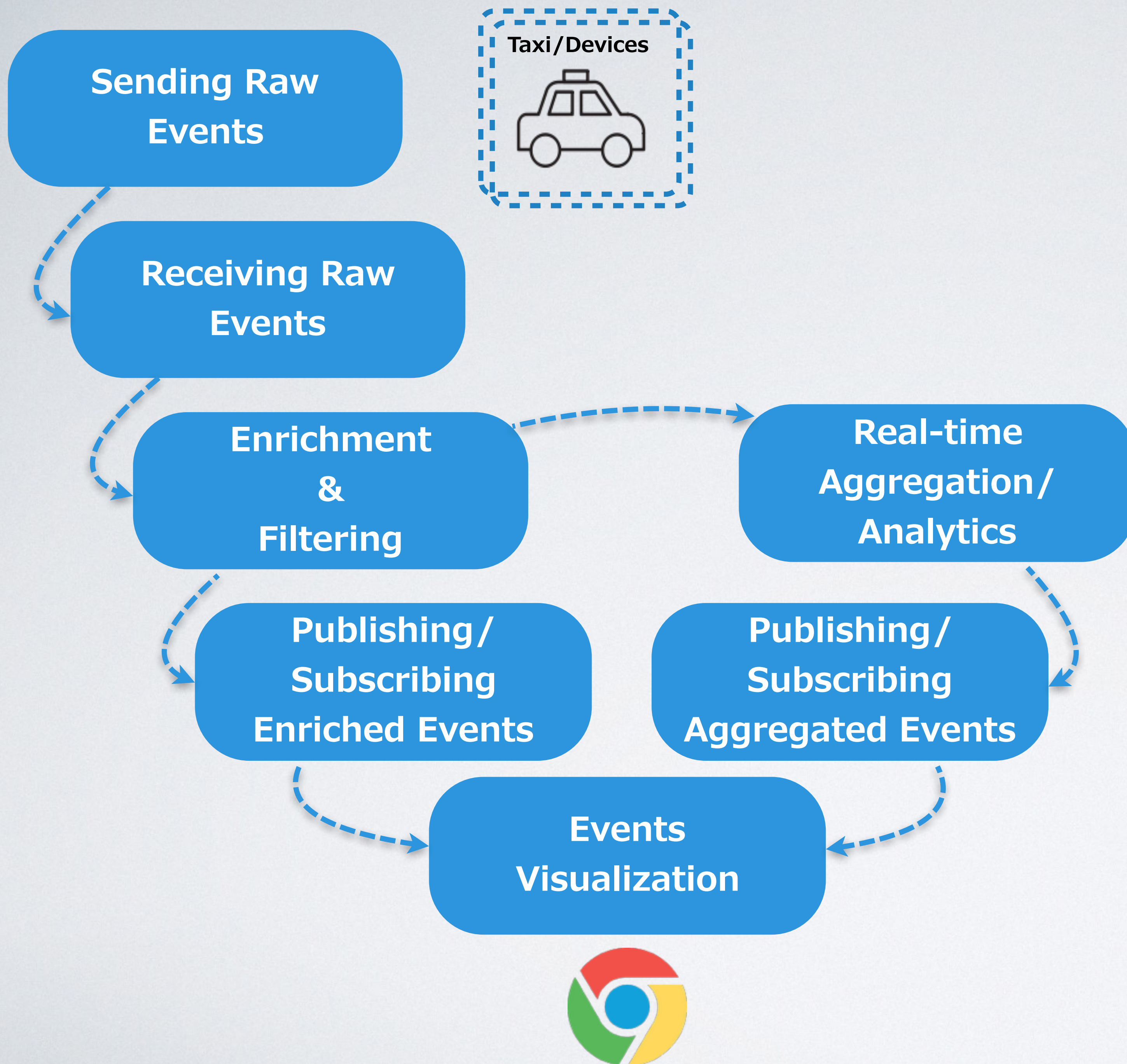
3.EVENT DRIVEN FLOW

Online Analytics

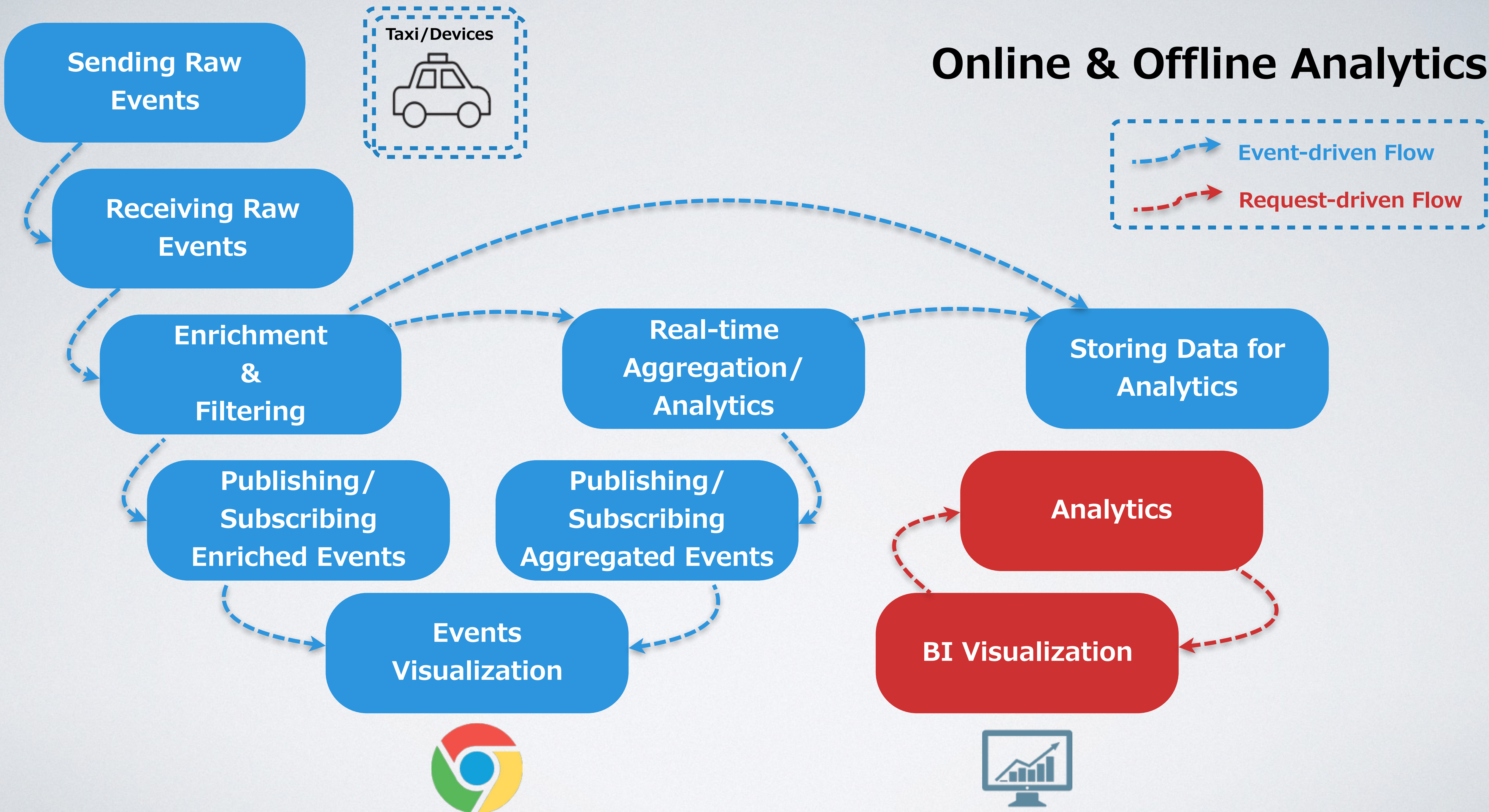


Expectations

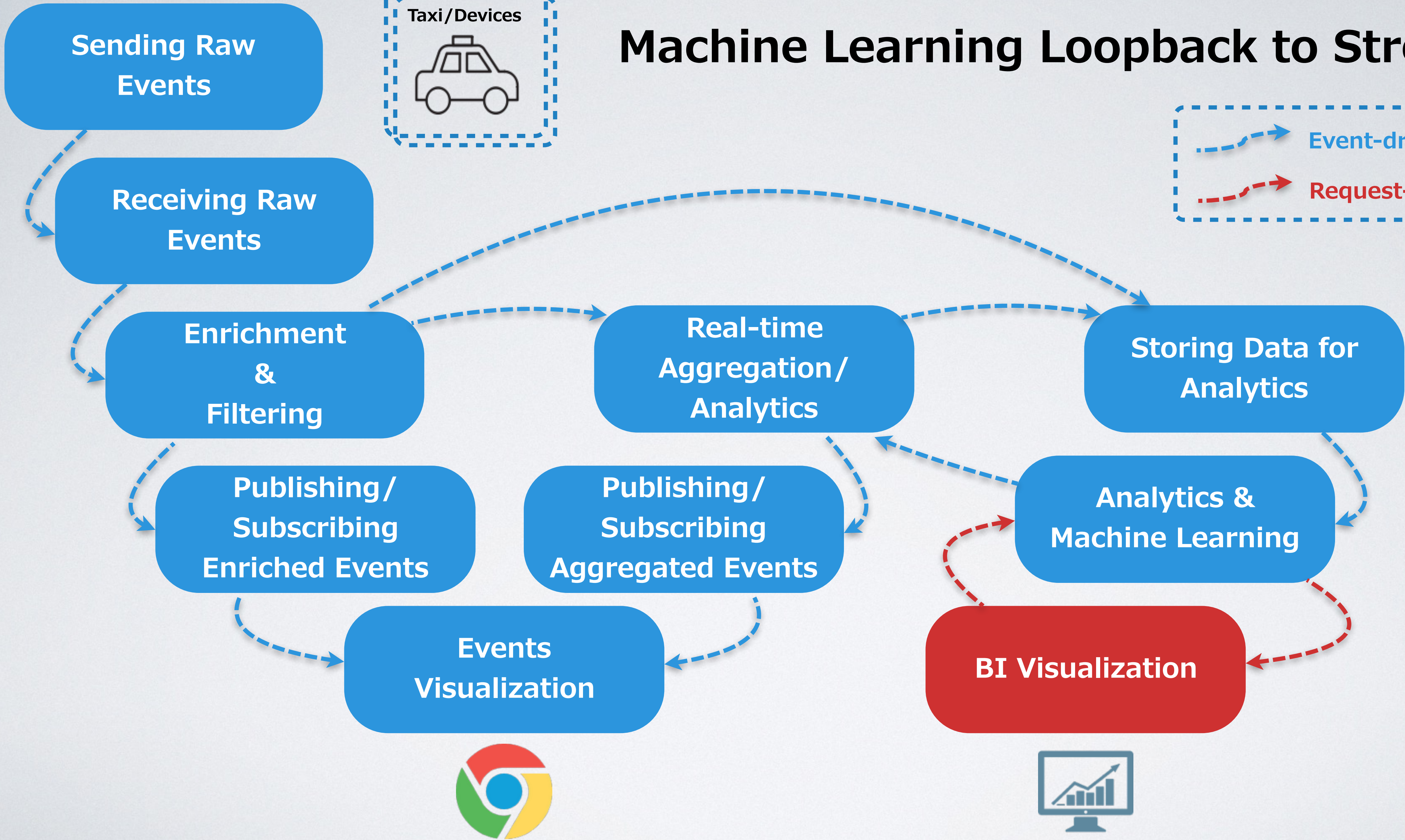
- Event Driven
- Low Latency
- Scalability
- High Availability
- Portability
- Light Weight



Online & Offline Analytics

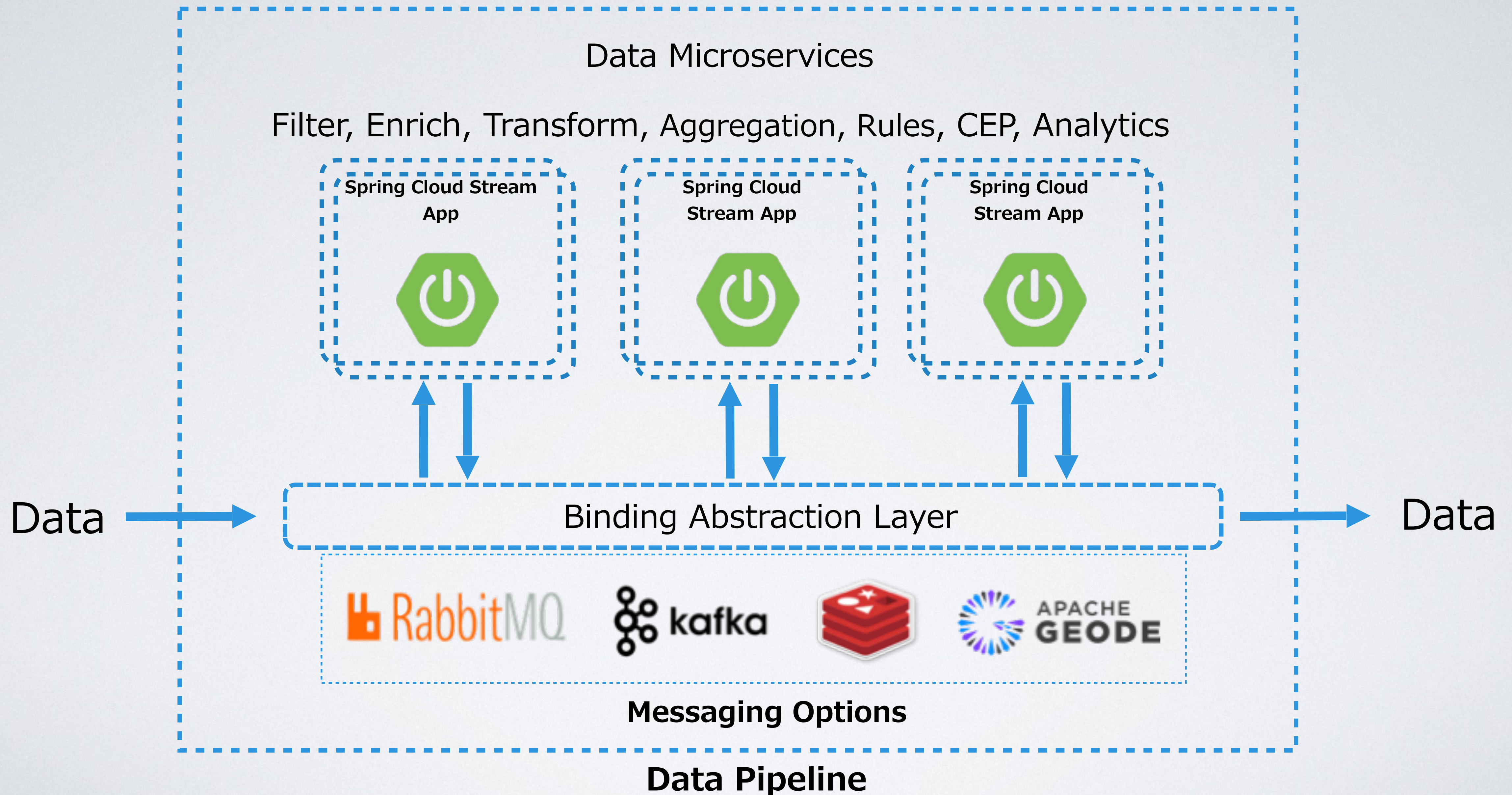


Machine Learning Loopback to Streaming

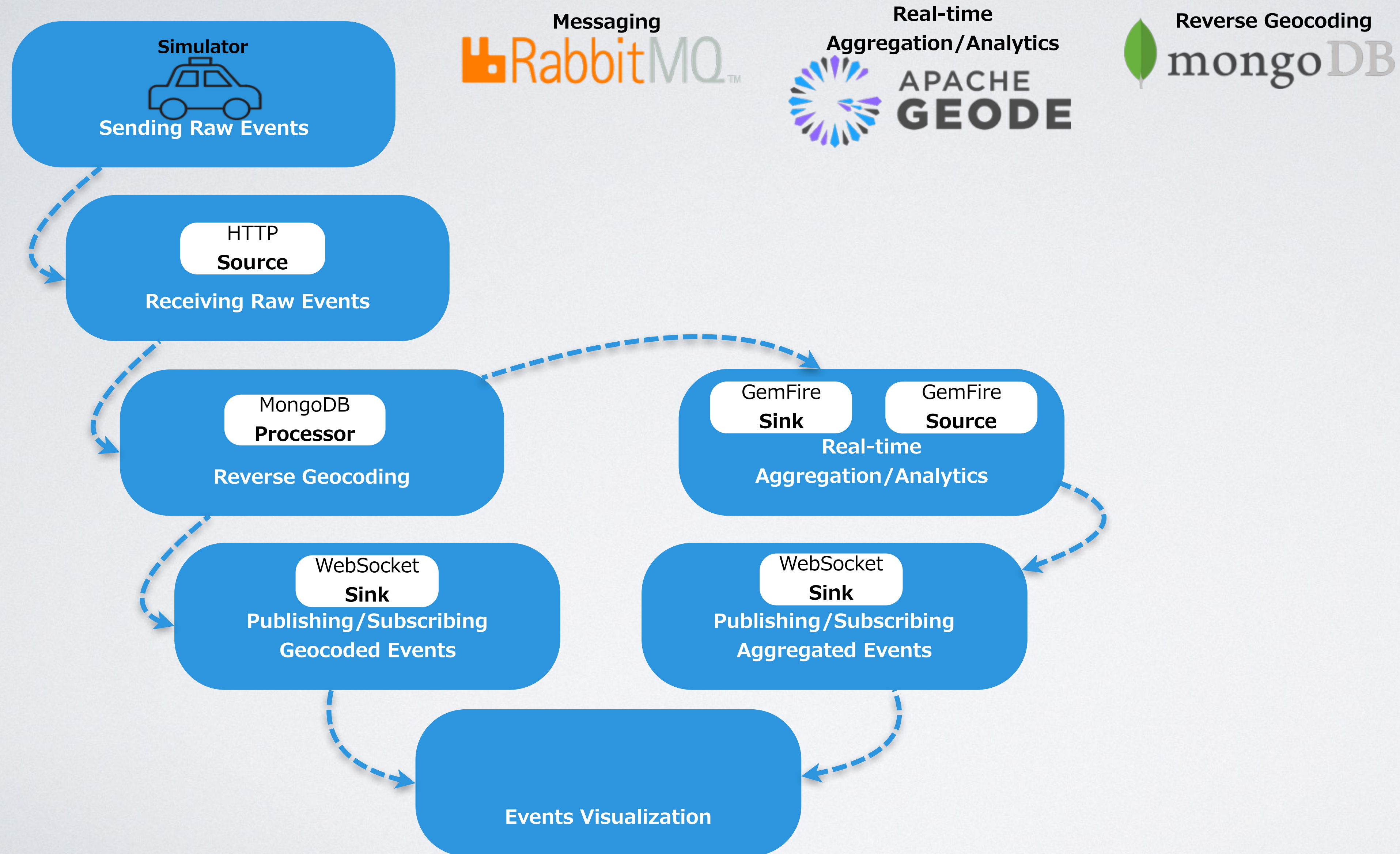


4.EVENT DRIVEN ARCHITECTURE

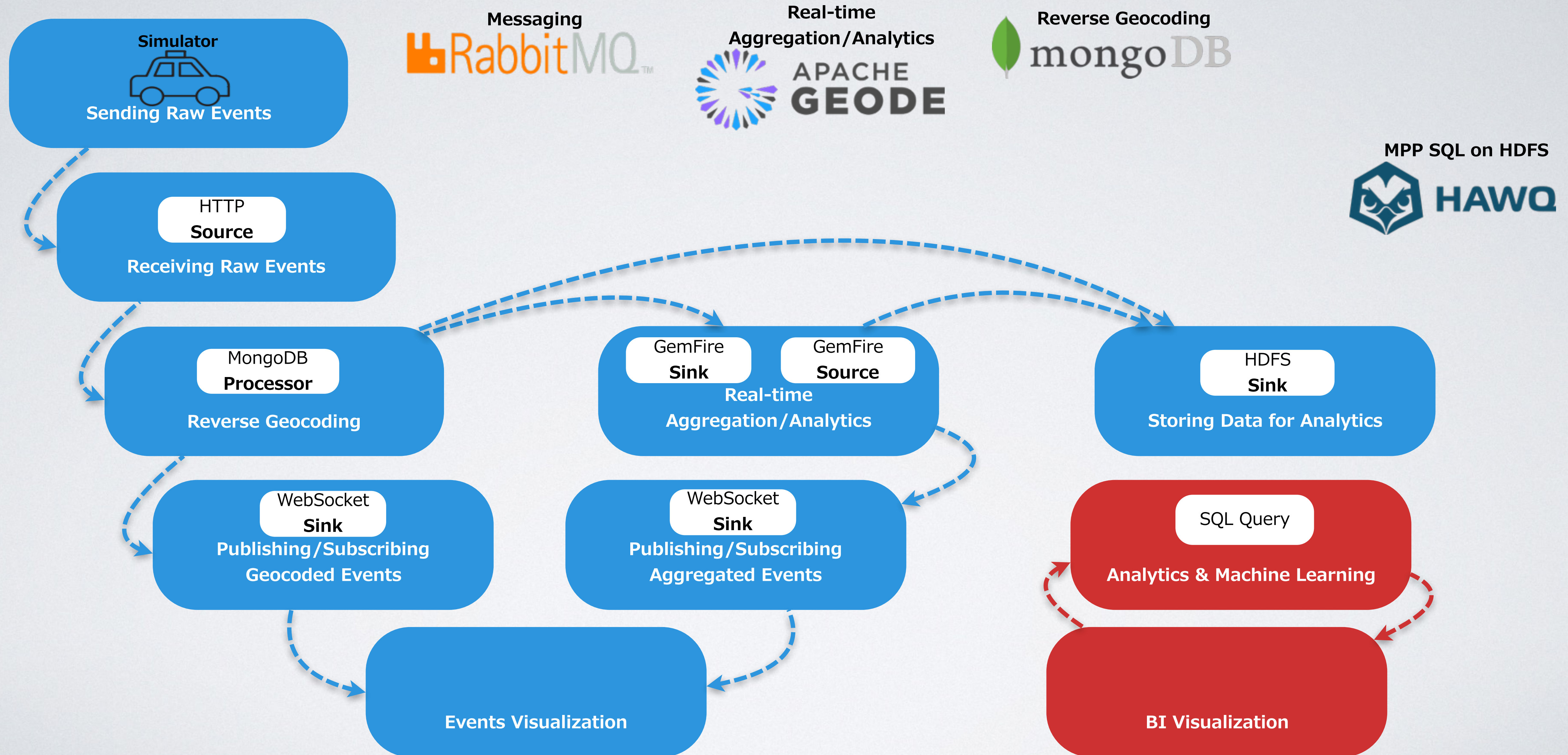
Spring Cloud Stream/Data Flow as Event Driven Architecture



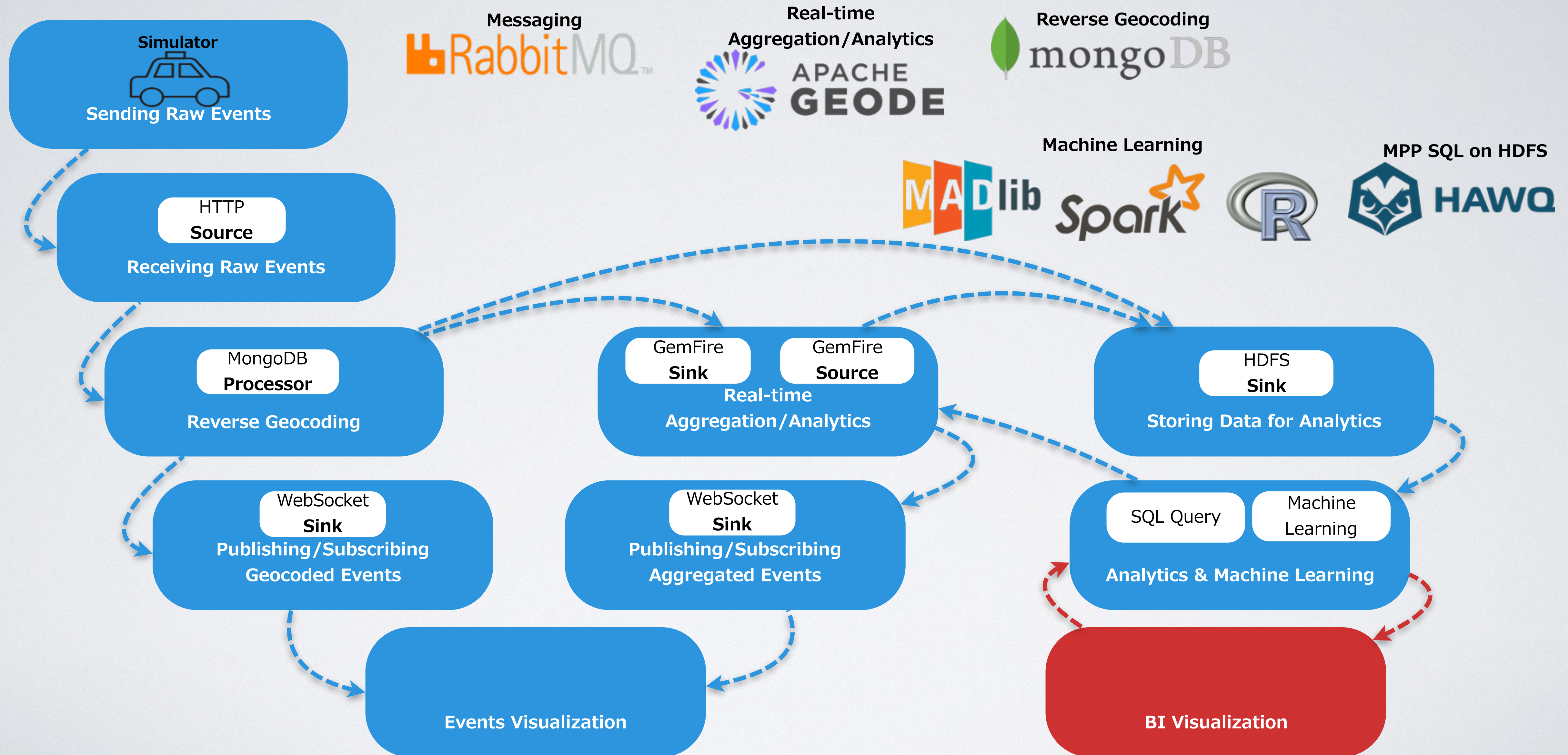
Apps & Services – Online Analytics



Apps & Services – Online & Offline Analytics

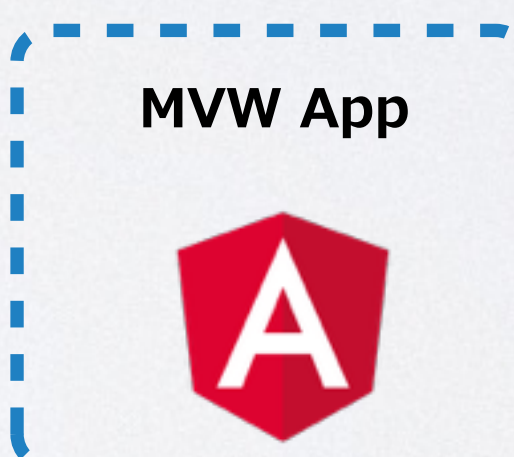
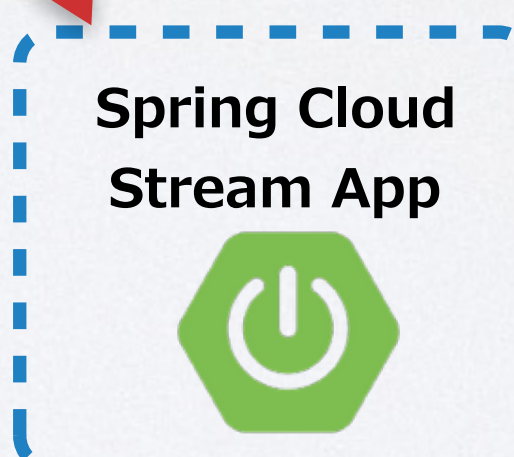
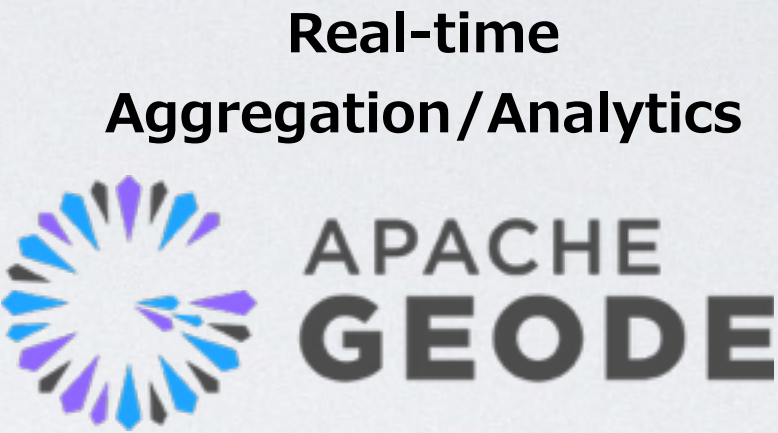
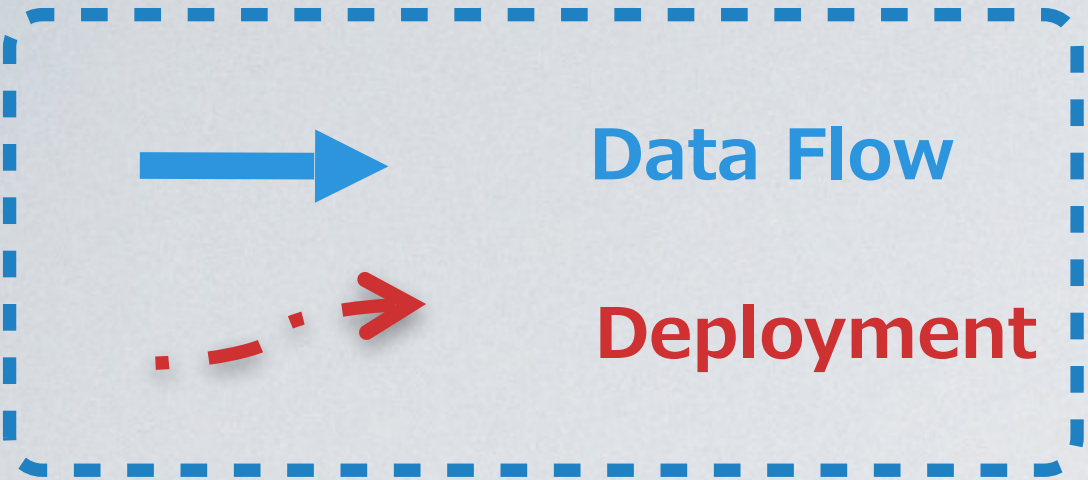


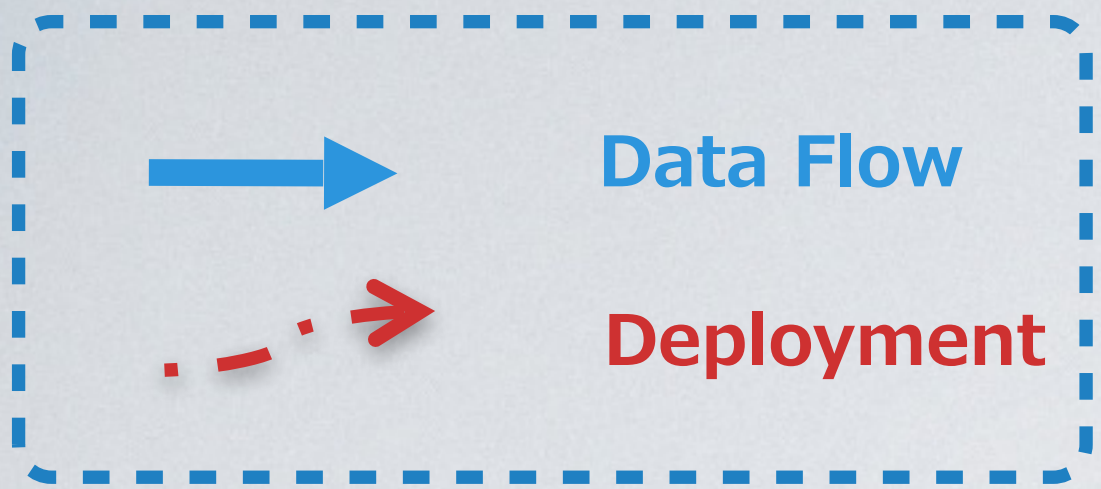
Apps & Services – Machine Learning Loopback to Streaming



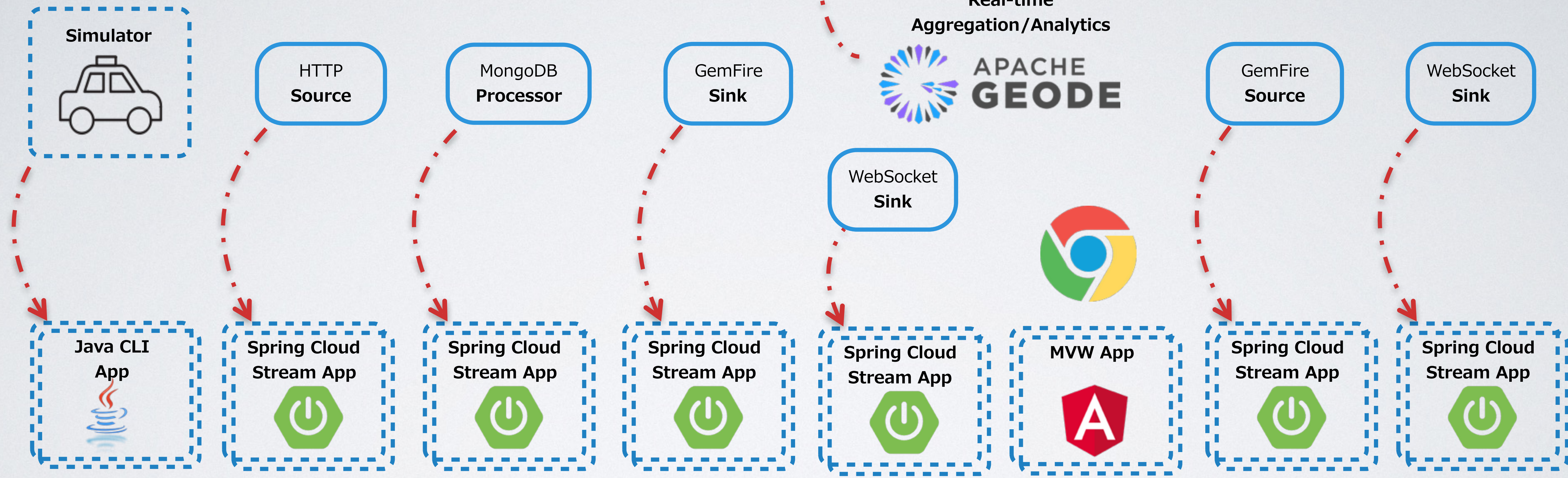
5.ONLINE ANALYTICS LOCAL DEPLOYMENT

Deployment

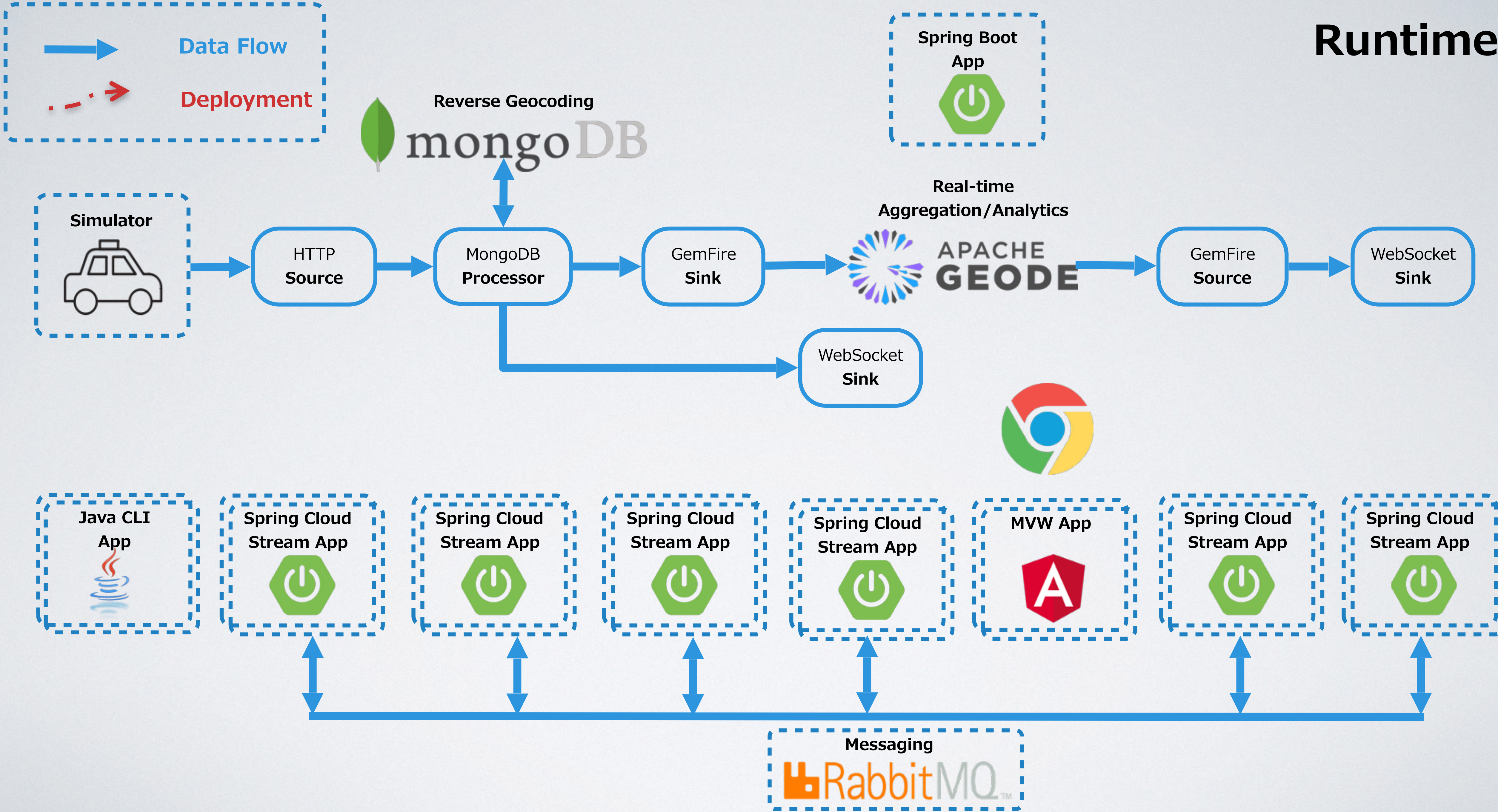




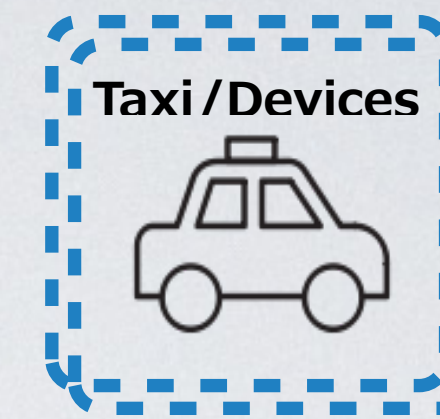
Scale Out When Necessary



Runtime



6. PRODUCTION LEVEL DEPLOYMENT

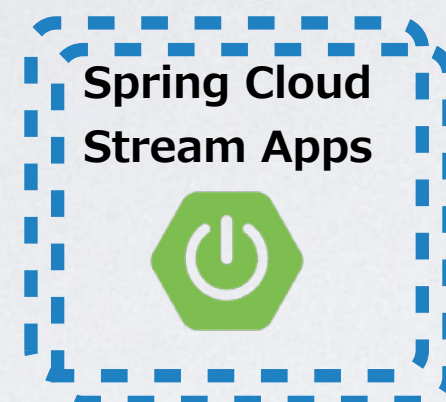


Containerized Applications

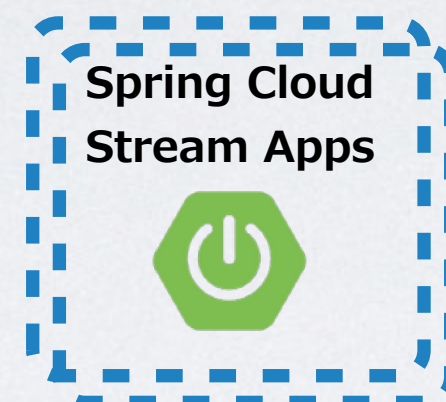
Business MicroServices



Data MicroServices



.....



Front-end Services

Broker
by RabbitMQ

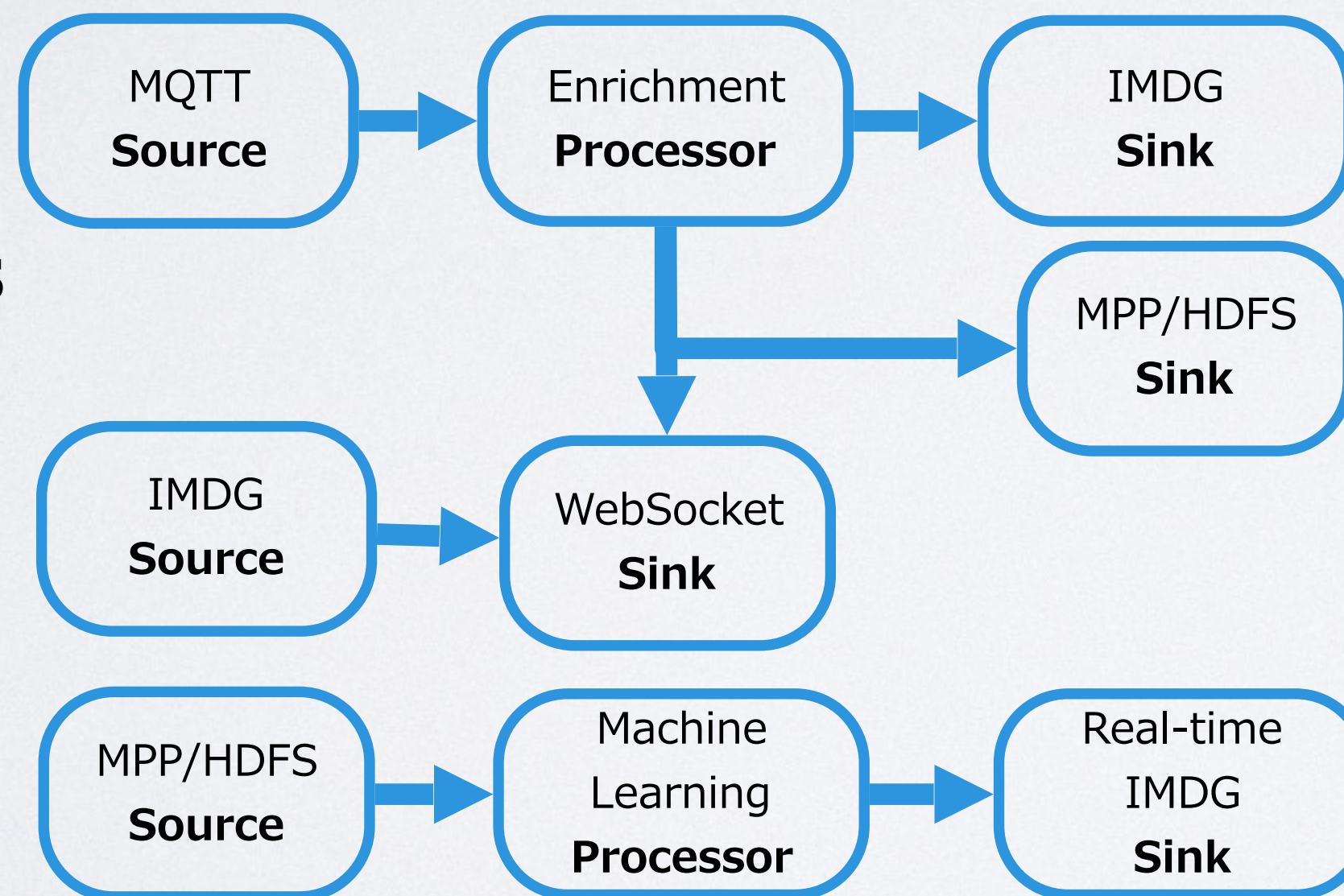


WebSocket
by RabbitMQ



Data Pipelines

Spring Cloud Data Flow



Back-end Services

Geode



RabbitMQ



MongoDB



HAWQ



- Event Driven
Spring Cloud Stream
- Low Latency
Messaging by RabbitMQ
Real-time by GemFire
- Scalability
Apps Scale-Out by Containers
Governance by Data Flow
Services Scale-out
- High Availability
Platform HA by Cloud Native
Platform or PaaS
- Portability
Apps & Services portable to
other Cloud Native Platforms
- Light Weight
Loosely Coupled MicroServices

CLOUD **FOUNDRY**

or



docker

or



or



etc.

7.DEPLOYMENT & RUNTIME SCREENSHOTS

Apps













This section lists all the available applications and provides the control to register/unregister them (if applicable).

All Applications

+ Register Application(s)

Unregister Application(s)

Quick filter

<input type="checkbox"/>	Name	Type	URI	Actions	
<input type="checkbox"/>	gemfire-source-rabbit	source	file:///temp/gemfire-source-rabbit-1.0.2.RELEASE.jar		
<input type="checkbox"/>	http-source-rabbit	source	file:///temp/http-source-rabbit-1.0.2.RELEASE.jar		
<input type="checkbox"/>	geocoding-processor-rabbit	processor	file:///temp/spring-cloud-stream-processor-geocoding-reverse-0.0.1-SNAPSHOT.jar		
<input type="checkbox"/>	gemfire-sink-rabbit	sink	file:///temp/gemfire-sink-rabbit-1.0.2.RELEASE.jar		
<input type="checkbox"/>	multi-topic-stomp-sink-rabbit	sink	file:///temp/spring-cloud-stream-sink-stomp-multi-topic-0.0.1-SNAPSHOT.jar		
<input type="checkbox"/>	stomp-sink-rabbit	sink	file:///temp/spring-cloud-stream-sink-stomp-0.0.1-SNAPSHOT.jar		

Streams

Create a stream using text based input or the visual editor.

[Definitions](#)
[Create Stream](#)
[Create Stream](#)
[Clear](#)
[Layout](#)

 Zoom: 159 %
☒ Auto Link

☐ Grid

```
1 http2multi-stomp=http-source-rabbit --port=9000 | geocoding-processor-rabbit --database=geojson --collection=blocks | multi-
  topic-stomp-sink-rabbit --withsockjs=true --topic-path=/dropoffDistrictCode --port=9210
```

▼ source

☐ gemfire-so...

☐ http-sourc...

▼ processor

☐ λ geocoding-...

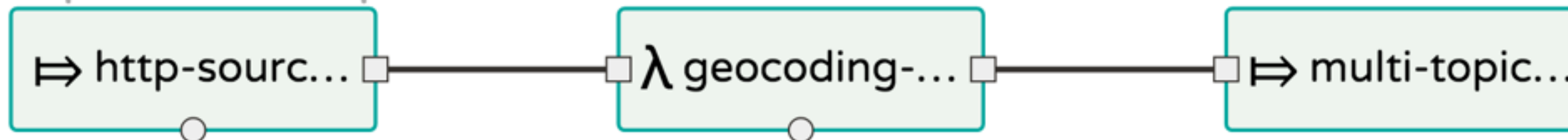
▼ sink

☐ gemfire-si...

☐ multi-topic...

☐ stomp-sin...

http2multi-stomp



Streams

This section lists all the stream definitions and provides the ability to deploy/undeploy or destroy streams.

Definitions [Create Stream](#)

Quick filter

Name	Definition	Status ?	Actions		
gem2stomp_top_district	gemfire-source-rabbit --cache-event-expression=newValue --region-name=RegDropoffDistrictTop stomp-sink-rabbit --withsockjs=true --topic=topdropoff --port=9410	deployed	<div>Undeploy</div>	<div>Deploy</div>	<div>Destroy</div>
gem2stomp_top_route	gemfire-source-rabbit --cache-event-expression=newValue --region-name=RegRouteTopTen stomp-sink-rabbit --withsockjs=true --topic=toproute --port=9400	deployed	<div>Undeploy</div>	<div>Deploy</div>	<div>Destroy</div>
http2gem	:http2multi-stomp.geocoding-processor-rabbit > gemfire-sink-rabbit --json=true --key-expression=payload.getField('uuid') --region-name=RegRaw	deployed	<div>Undeploy</div>	<div>Deploy</div>	<div>Destroy</div>
http2multi-stomp	http-source-rabbit --port=9000 geocoding-processor-rabbit --database=geojson --collection=blocks multi-topic-stomp-sink-rabbit --withsockjs=true --topic-path=/dropoffDistrictCode --port=9210	deployed	<div>Undeploy</div>	<div>Deploy</div>	<div>Destroy</div>

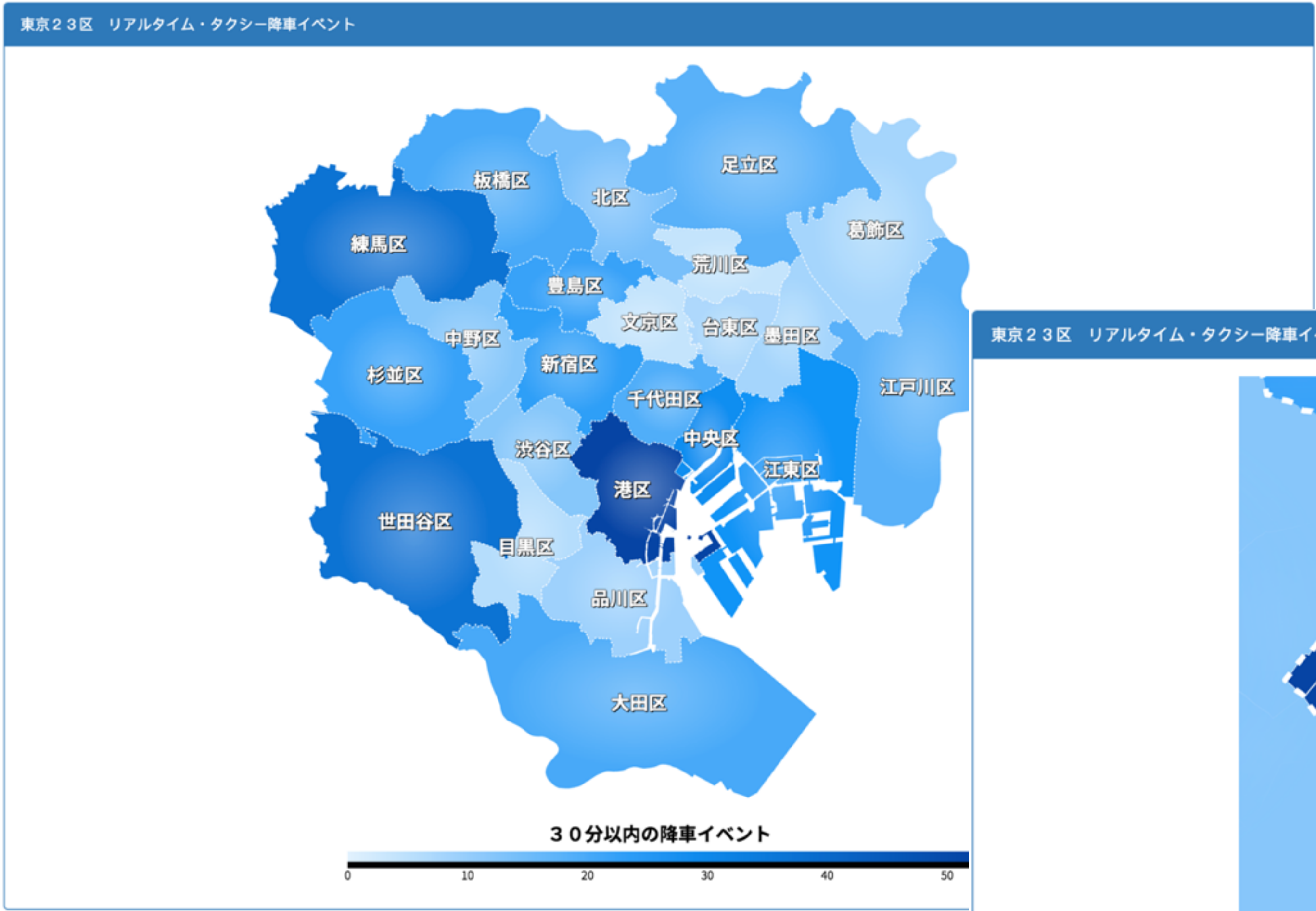
Cluster view

This section shows the Spring Cloud Data Flow cluster view with the list of all running apps.

Runtime Apps

Quick filter

App Id	State	# of Instances
gem2stomp_top_district.gemfire-source-rabbit	deployed	1
gem2stomp_top_district.stomp-sink-rabbit	deployed	1
gem2stomp_top_route.gemfire-source-rabbit	deployed	1
gem2stomp_top_route.stomp-sink-rabbit	deployed	1
http2gem.gemfire-sink-rabbit	deployed	1
http2multi-stomp.geocoding-processor-rabbit	deployed	1
http2multi-stomp.http-source-rabbit	deployed	1
http2multi-stomp.multi-topic-stomp-sink-rabbit	deployed	1



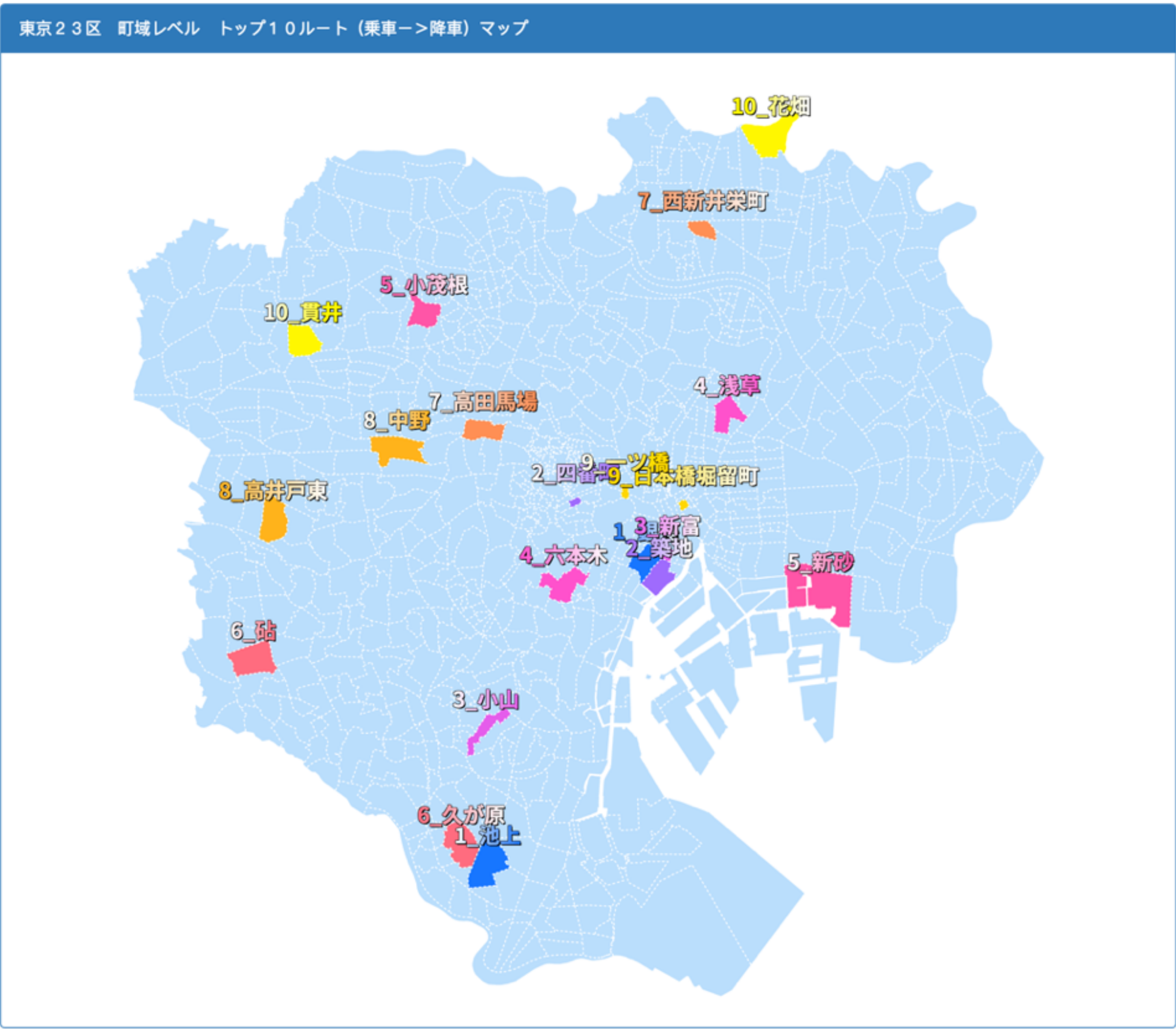
• Dropoff-district Count

• Real-time Taxi Dropoff Events



東京23区 町域レベル トップ10ルート（乗車→降車）リスト			
Rank	From	To	Count
1	中央区 銀座	大田区 池上	8
2	中央区 築地	千代田区 四番町	8
3	中央区 新富	品川区 小山	8
4	港区 六本木	台東区 浅草	8
5	板橋区 小茂根	江東区 新砂	8
6	大田区 久が原	世田谷区 砧	8
7	足立区 西新井栄町	新宿区 高田馬場	8
8	杉並区 高井戸東	中野区 中野	8
9	中央区 日本橋堀留町	千代田区 一ツ橋	8
10	足立区 花畑	練馬区 貫井	8

- 10 most frequent routes



- 10 most frequent routes

8. LOCAL DEPLOYMENT DEMO VIDEO