PULSAR SUMMIT | Pulsar Summit Asia 2022

Building Modern Data Streaming Apps

Tim Spann, Developer Advocate

Tim Spann

Developer Advocate
StreamNative

FLiP(N) Stack = Flink, Pulsar and NiFi Stack
Streaming Systems & Data Architecture Expert

Experience

15+ years of experience with streaming technologies including Pulsar, Flink, Spark, NiFi, Big Data, Cloud, MXNet, IoT, Python and more.

Today, he helps to grow the Pulsar community sharing rich technical knowledge and experience at both global conferences and through individual conversations.

CLOUDERA · Pivotal · BARNES & NOBLE · EY · HORTONWORKS · airis.DATA · Hewlett Packard Enterprise

FLiP Stack Weekly

https://bit.ly/32dAJft

This week in Apache Flink, Apache Pulsar, Apache NiFi, Apache Spark and open source friends.

What Is this all about?

In my session, I will show you some best practices I have discovered over the last 7 years in building data streaming applications including IoT, CDC, Logs, and more.

In my modern approach, we utilize several Apache frameworks to maximize the best features of all. We often start with Apache NiFi as the orchestrator of streams flowing into Pulsar. From there we build streaming ETL with Spark, enhance events with Pulsar Functions for ML and enrichment. We build continuous queries against our topics with Flink SQL.

WHAT? WHY?

Data

```
'aircraft': [{'hex': 'ae6d7a',
'alt_baro': 25000, 'mlat': [],
'tisb': [], 'messages': 177,
'seen': 0.1, 'rssi': -22.7}
```

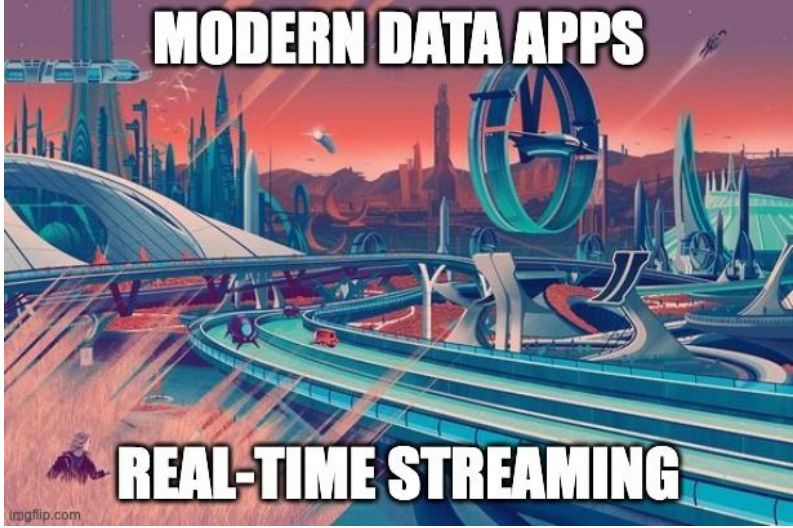
```
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet href="latest_ob.xsl" type="text/xsl"?>
        <current_observation><temp_c>20.0</temp_c>
        <pressure_string>1019.1 mb</pressure_string>
        <pressure_mb>1019.1</pressure_mb>
        <pressure_in>30.17</pressure_in>
        <dewpoint_string>26.6 F</dewpoint_string>
</current_observation>
```

```
{
  "title" : "BUS 76 - Oct 17, 2022 04:38:41 PM",
  "description" : "Bus route 76 is operating on a detour in Hasbrouck Heights. Terrace Avenue
is closed from Paterson Avenue to Madison Avenue for utility work.Buses will use Paterson",
  "link" : "https://www.njtransit.com/node/1540960",
  "guid" : "https://www.njtransit.com/node/1540960",
  "advisoryAlert" : 0,
  "pubDate" : "Oct 17, 2022 04:38:41 PM",
  "ts" : "1666039443600",
  "companyname" : "newjersey",
  "uuid" : "3b19dd32-db1e-4320-ba9a-f6bfd4a87bb9",
  "servicename" : "bus"
}
```

Information?

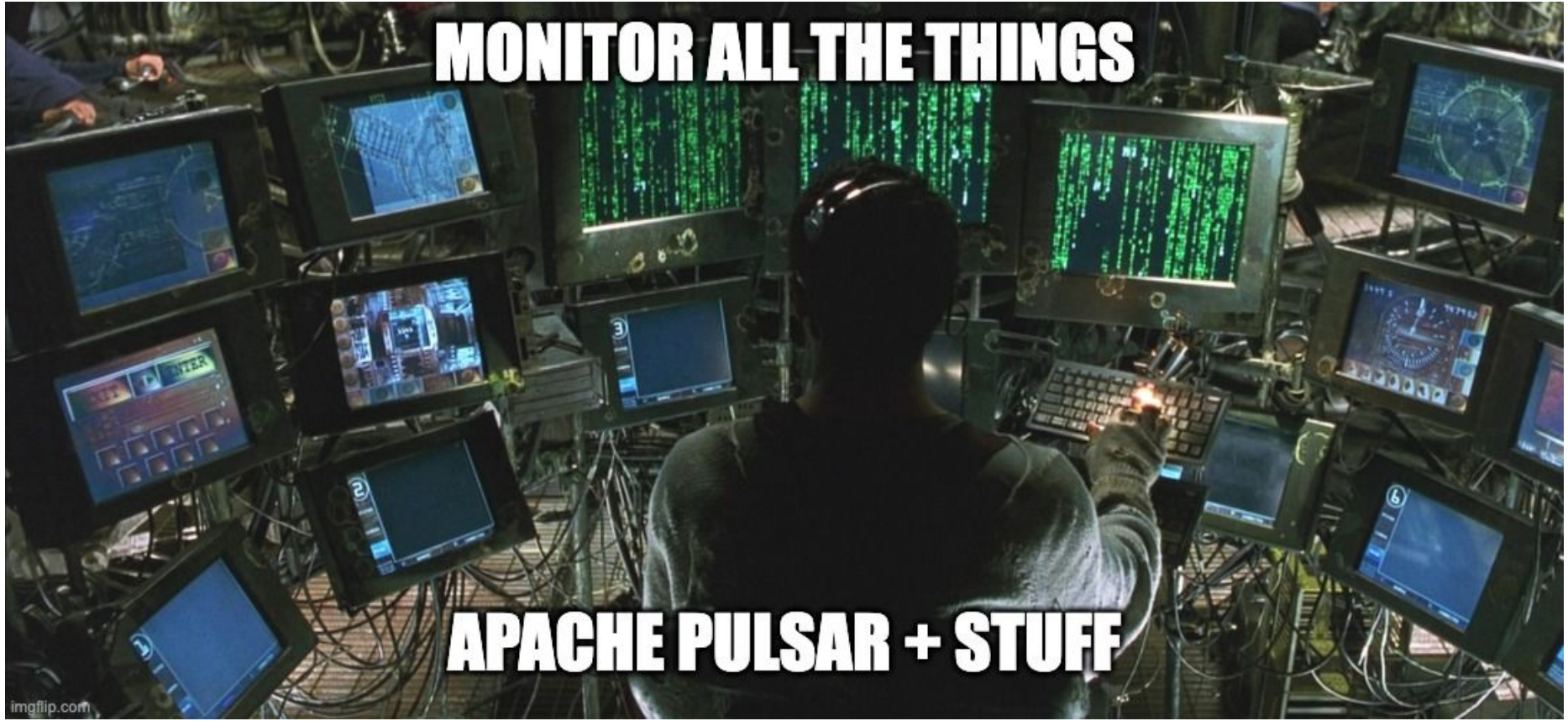
In Ancient Times

In Modern Times



Trains, Planes and Automobiles +++

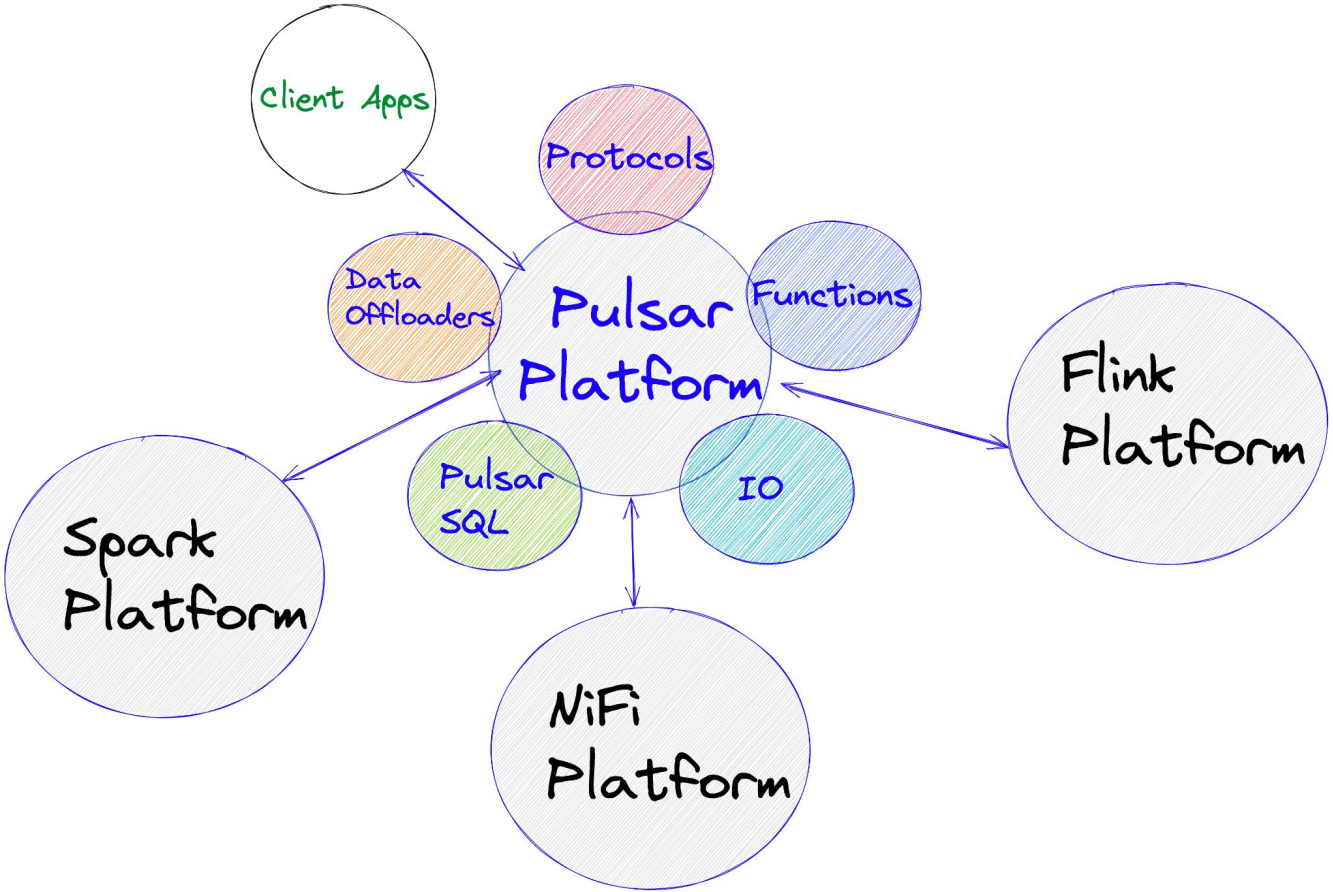
Information Needed	Data Feed(s)
Local weather conditions	• XML, JSON, RSS
Mass transit status & alerts	• XML, JSON, RSS
Regional highways & tunnels	• GeoRSS, XML, ProtoBuf, JSON
Local social media	• JSON
ADS-B Plane Data	• JSON
Local air quality	• JSON

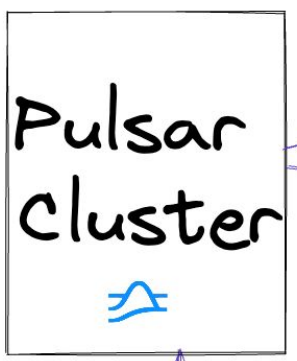

MONITOR ALL THE THINGS
APACHE PULSAR + STUFF
imgflip.com

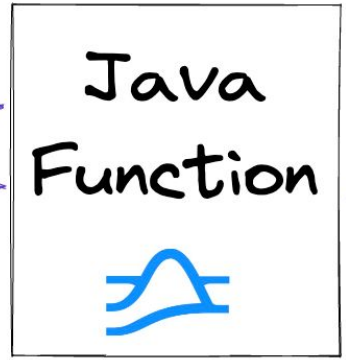
HOW?

What the FLiP is the FLiP Stack?

https://streamnative.io/blog/engineering/2022-04-14-what-the-flip-is-the-flip-stack/


Client Apps
Protocols
Data Offloaders
Pulsar Platform
Functions
Flink Platform
Spark Platform
Pulsar SQL
IO
NiFi Platform



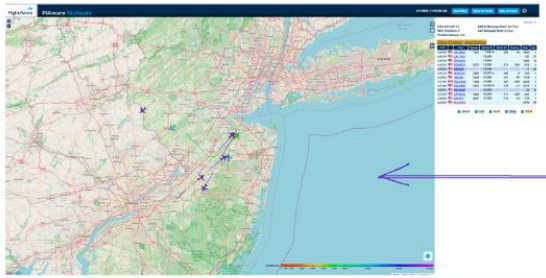


SQL

Topic aircraft

Topic ADSB-Raw

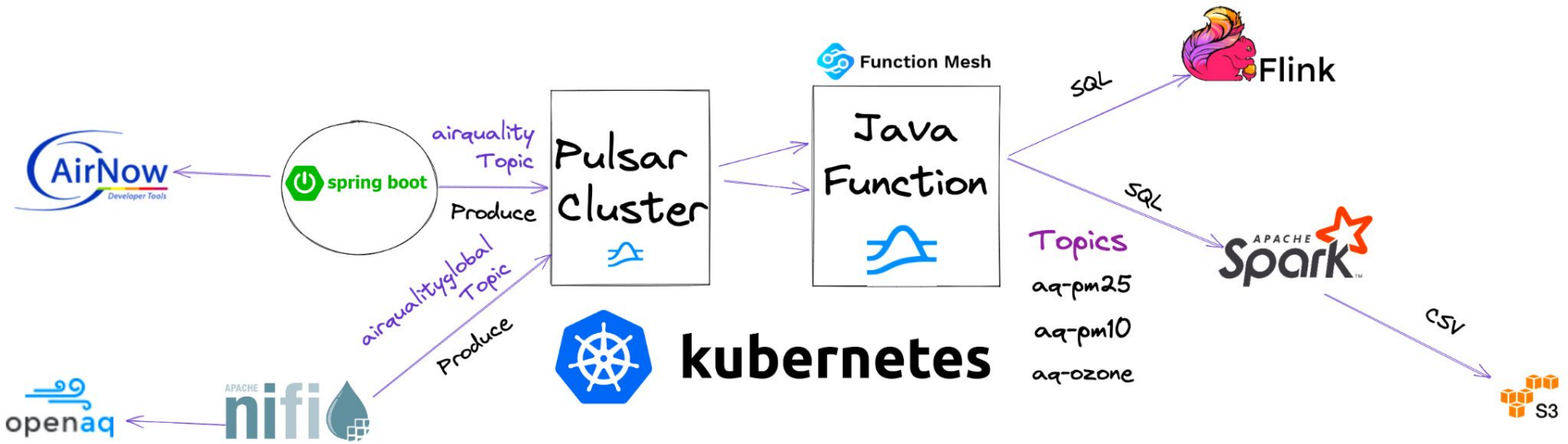
Produce

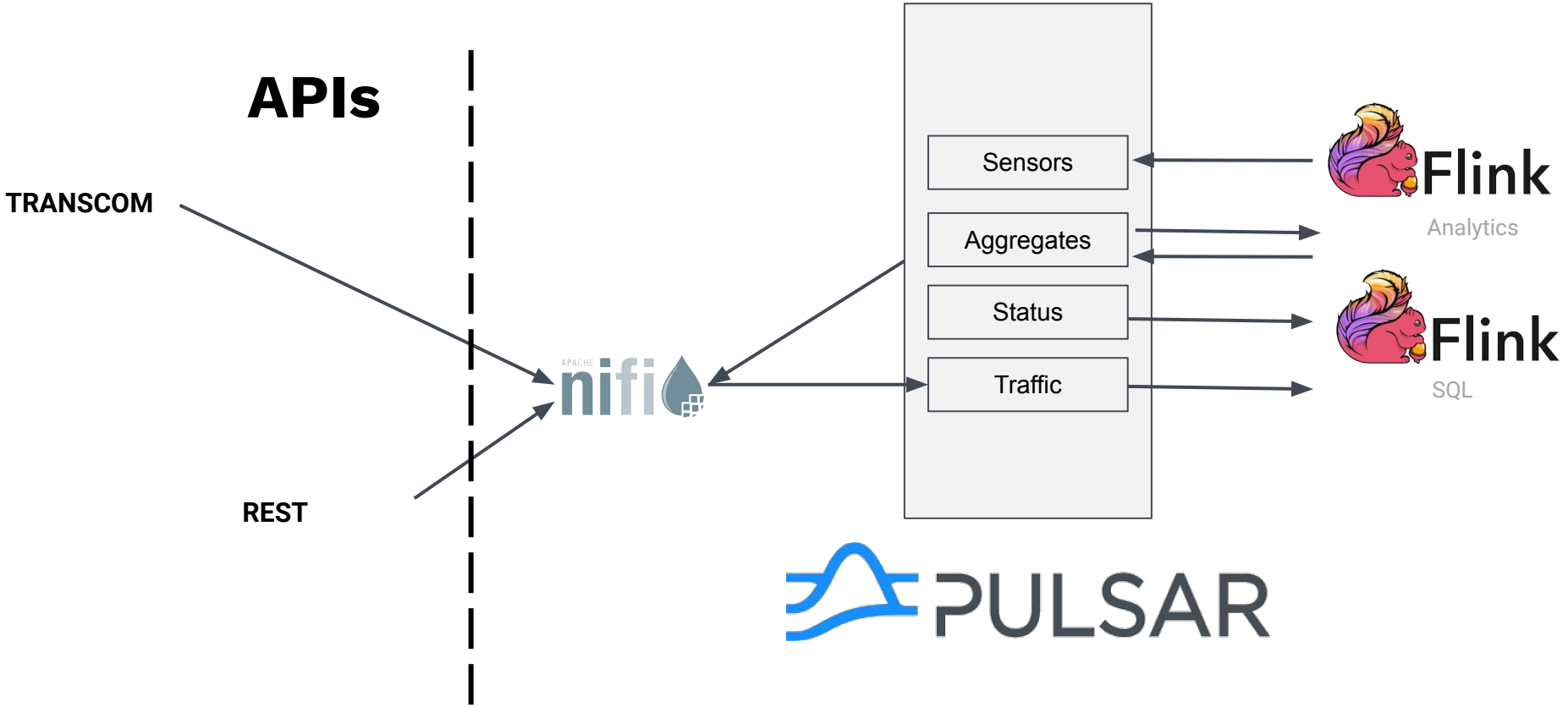


Apache Pulsar - Websockets - Data Tables - Aircraft (ADS-B)

Show 10 entries Search:

Altitude	Ground Speed	Mach	Hex	Flight #	Latitude	Longitude	PublishTime
7150	297.9	0.432	a43ea8	UAL759	40.320267	-74.745423	2022-09-12T16:50:28.385-04:00
7250	297.9	0.432	a43ea8	UAL759	40.369446	-74.746768	2022-09-12T16:50:27.376-04:00
7275	297.9	0.432	a43ea8	UAL759	40.368983	-74.745234	2022-09-12T16:50:26.355-04:00
7300	297.9	0.432	a43ea8	UAL759	40.367569	-74.747253	2022-09-12T16:50:25.355-04:00
7325	297.9	0.432	a43ea8	UAL759	40.365601	-74.75	2022-09-12T16:50:24.315-04:00
6125	284	0.424	a273f5	UAL2477	40.503706	-74.647917	2022-09-12T16:50:28.415-04:00
6125	284	0.424	a273f5	UAL2477	40.503706	-74.647917	2022-09-12T16:50:27.405-04:00
6125	284	0.424	a273f5	UAL2477	40.500916	-74.648082	2022-09-12T16:50:26.385-04:00
6125	284	0.424	a273f5	UAL2477	40.500916	-74.648082	2022-09-12T16:50:25.365-04:00
6125	284	0.424	a273f5	UAL2477	40.499759	-74.648978	2022-09-12T16:50:24.345-04:00
Altitude	Ground Speed	Mach	Hex	Flight #	Latitude	Longitude	PublishTime

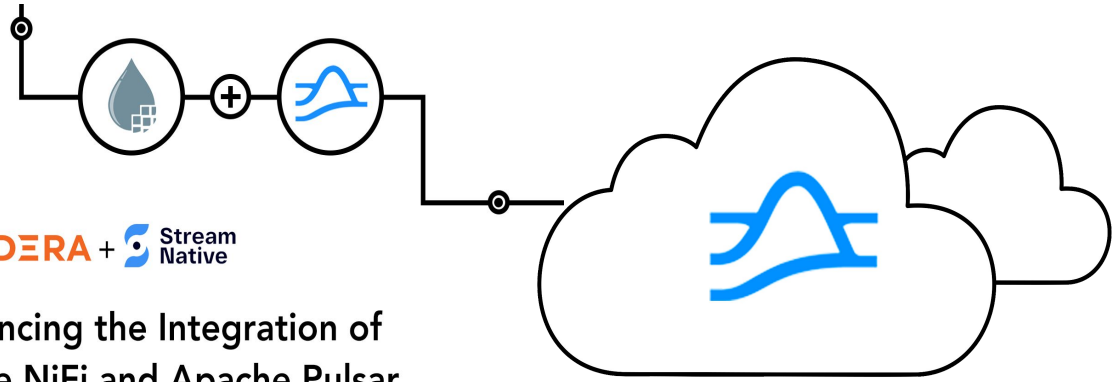

AirNow
Developer Tools
spring boot
airquality Topic
Produce
Pulsar Cluster
Function Mesh
Java Function
SQL
Flink
SQL
APACHE Spark
CSV
S3
Topics
aq-pm25
aq-pm10
aq-ozone
airqualityglobal Topic
Produce
kubernetes
APACHE nifi
openaq


APIs
TRANSCOM
REST
APACHE nifi
Sensors
Aggregates
Status
Traffic
Flink
Analytics
Flink
SQL
PULSAR

INGEST: Apache NiFi

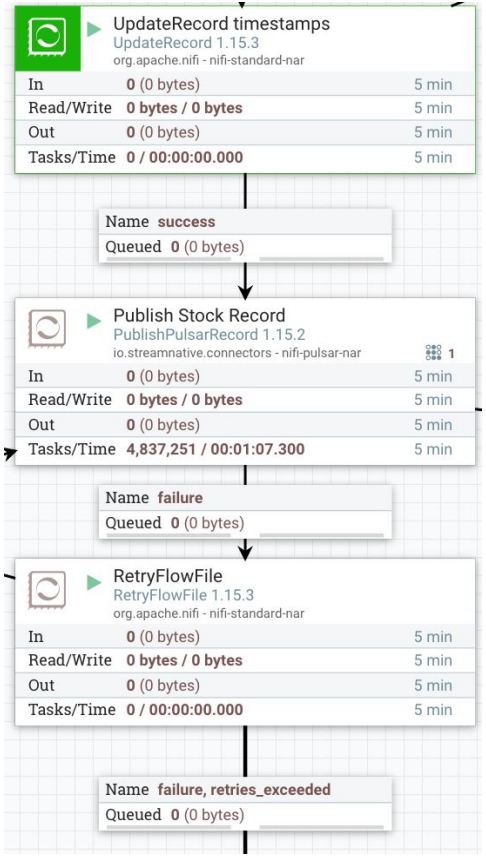

APACHE nifi
PULSAR
(Queuing + Streaming)
APACHE nifi

Apache NiFi Pulsar Connector



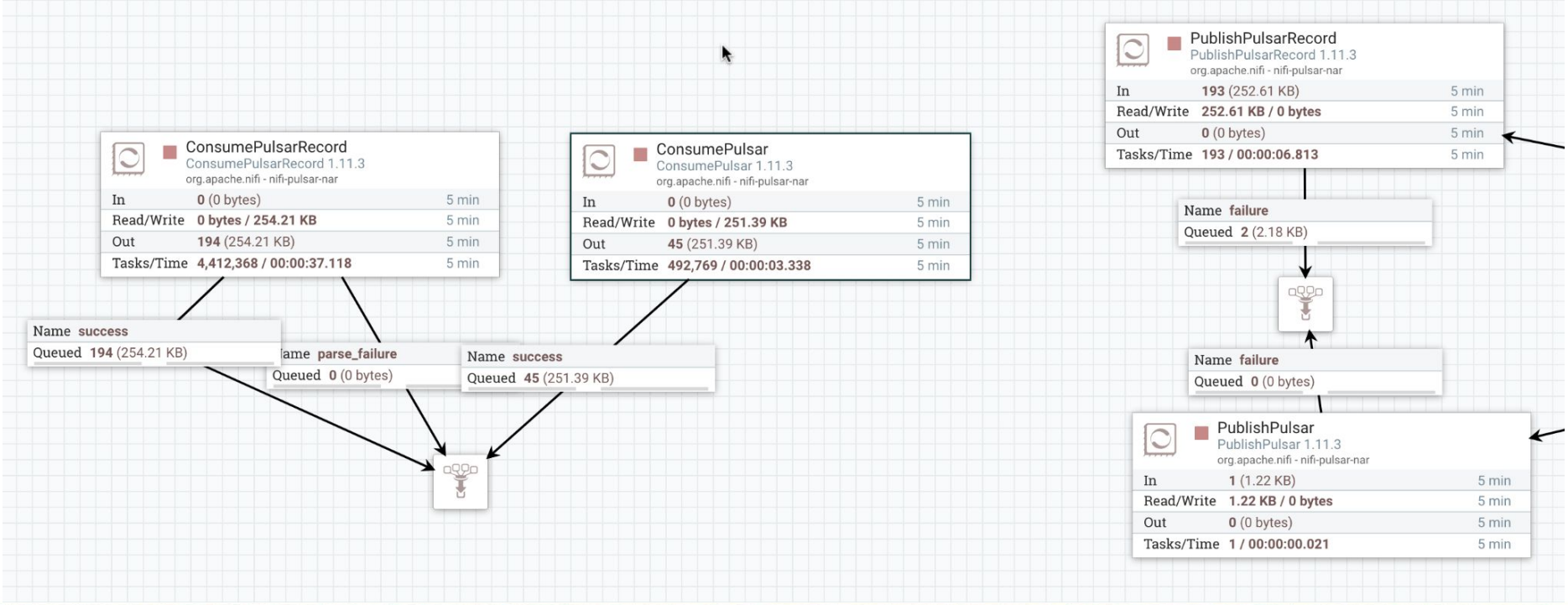
CLOUDERA + Stream Native

Announcing the Integration of Apache NiFi and Apache Pulsar



https://streamnative.io/apache-nifi-connector/

Apache NiFi <-> Apache Pulsar



Apache NiFi - Data Lineage / Provenance

NiFi Data Provenance

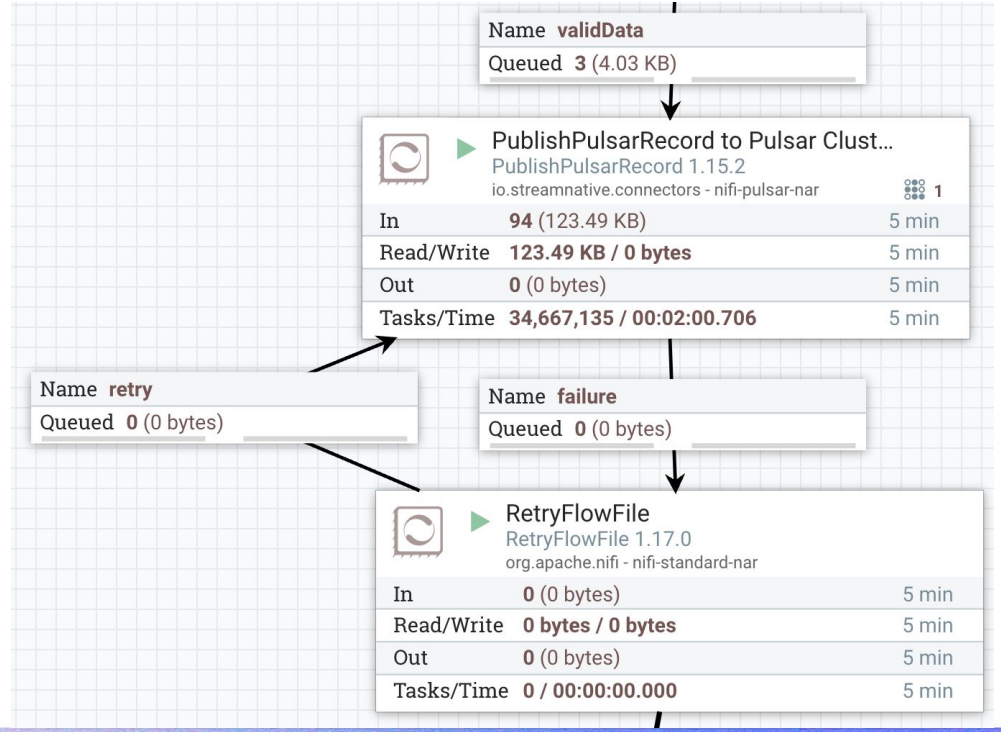
Displaying 1,000 of 1,000
Oldest event available: 09/06/2022 20:48:48 EDT

Showing 1,000 of 1,000+ events that match the specified query, please refine the search. Clear search

Filter | by component name

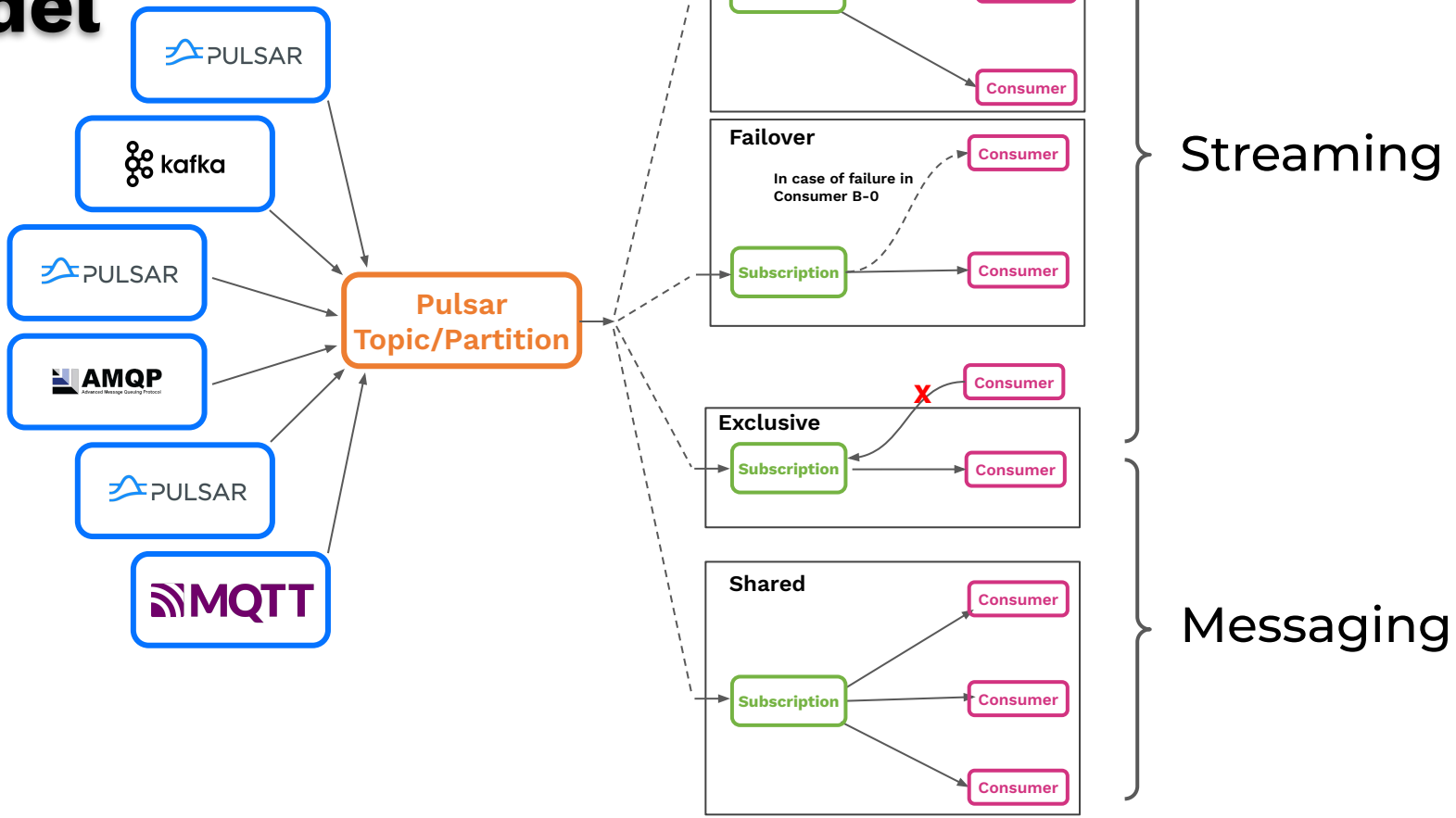
Date/Time	Type	FlowFile Uuid	Size	Component Name	Component Type
10/11/2022 20:21:04.021 EDT	DROP	c1b03f5b-7150-47e3-892e-86392e498c07	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.021 EDT	ATTRIBUTES_MODIFIED	c1b03f5b-7150-47e3-892e-86392e498c07	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.020 EDT	SEND	c1b03f5b-7150-47e3-892e-86392e498c07	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.013 EDT	DROP	34bf7692-9594-41bd-90b9-f3a6675729ea	1.32 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.013 EDT	ATTRIBUTES_MODIFIED	34bf7692-9594-41bd-90b9-f3a6675729ea	1.32 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.012 EDT	SEND	34bf7692-9594-41bd-90b9-f3a6675729ea	1.32 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.003 EDT	DROP	188db28b-cdf2-49b1-8ce9-9c928b032fa7	1.33 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.003 EDT	ATTRIBUTES_MODIFIED	188db28b-cdf2-49b1-8ce9-9c928b032fa7	1.33 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:04.002 EDT	SEND	188db28b-cdf2-49b1-8ce9-9c928b032fa7	1.33 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.992 EDT	DROP	c2110552-6f66-45a7-aab7-6b1918d104b2	1.31 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.992 EDT	ATTRIBUTES_MODIFIED	c2110552-6f66-45a7-aab7-6b1918d104b2	1.31 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.991 EDT	SEND	c2110552-6f66-45a7-aab7-6b1918d104b2	1.31 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.970 EDT	DROP	6b5d3d0d-a5c2-4401-b3d5-1c30fc4be76f	1.23 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.970 EDT	ATTRIBUTES_MODIFIED	6b5d3d0d-a5c2-4401-b3d5-1c30fc4be76f	1.23 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.969 EDT	SEND	6b5d3d0d-a5c2-4401-b3d5-1c30fc4be76f	1.23 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.961 EDT	DROP	eddaacb0-76d4-41bf-a95f-5c7f338d2b84	1.25 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.961 EDT	ATTRIBUTES_MODIFIED	eddaacb0-76d4-41bf-a95f-5c7f338d2b84	1.25 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.960 EDT	SEND	eddaacb0-76d4-41bf-a95f-5c7f338d2b84	1.25 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.943 EDT	ATTRIBUTES_MODIFIED	2de4f962-63ef-4ebb-88e7-bb61b516256a	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.942 EDT	DROP	2de4f962-63ef-4ebb-88e7-bb61b516256a	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.942 EDT	SEND	2de4f962-63ef-4ebb-88e7-bb61b516256a	1.4 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.933 EDT	DROP	27bb3538-851e-4940-b251-b8353f889417	1.39 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.933 EDT	ATTRIBUTES_MODIFIED	27bb3538-851e-4940-b251-b8353f889417	1.39 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord
10/11/2022 20:21:03.933 EDT	SEND	27bb3538-851e-4940-b251-b8353f889417	1.39 KB	PublishPulsarRecord to Pulsar Cluster	PublishPulsarRecord

Apache NiFi - Producing to Pulsar



TRANSIT:
Infinite Message Bus with Apache Pulsar

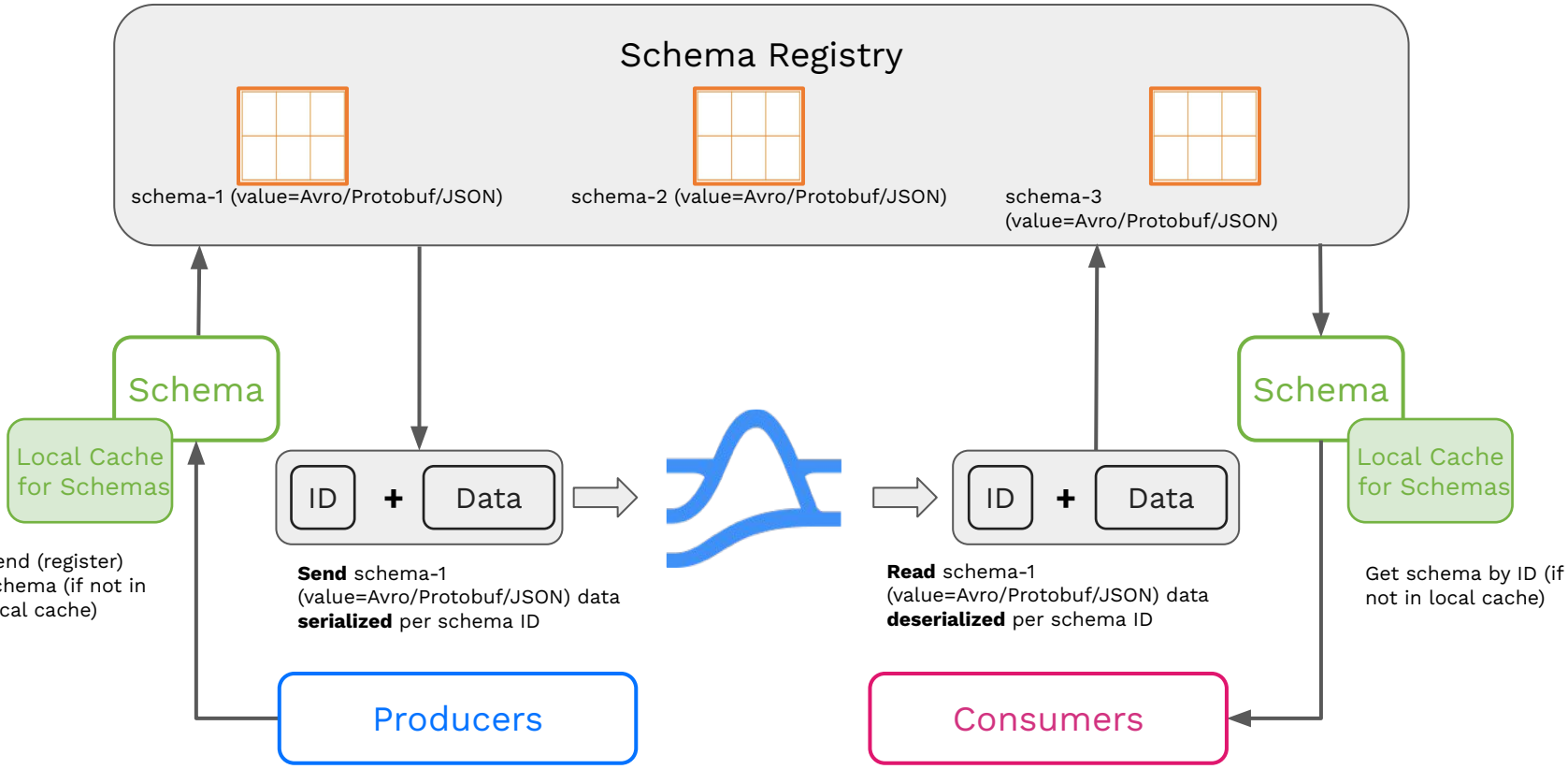
Unified Messaging Model



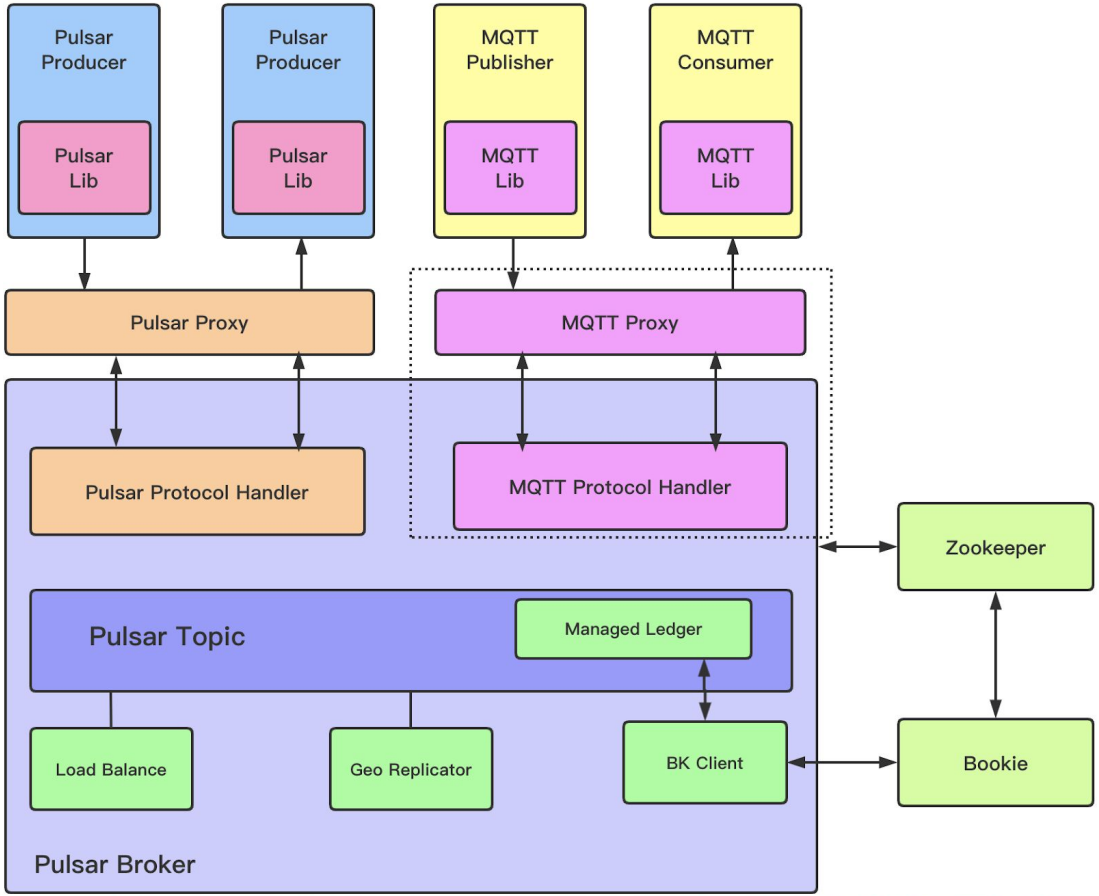
Apache Pulsar features

- Centralized cluster management and oversight.
- Cloud native with decoupled storage and compute layers.
- Geographic redundancy and high availability included.
- Built-in compatibility with your existing code and messaging infrastructure.
- Elastic horizontal and vertical scalability.
- Seamless and instant partitioning rebalancing with no downtime.
- Flexible subscription model supports a wide array of use cases.
- Compatible with the tools you use to store, analyze, and process data.

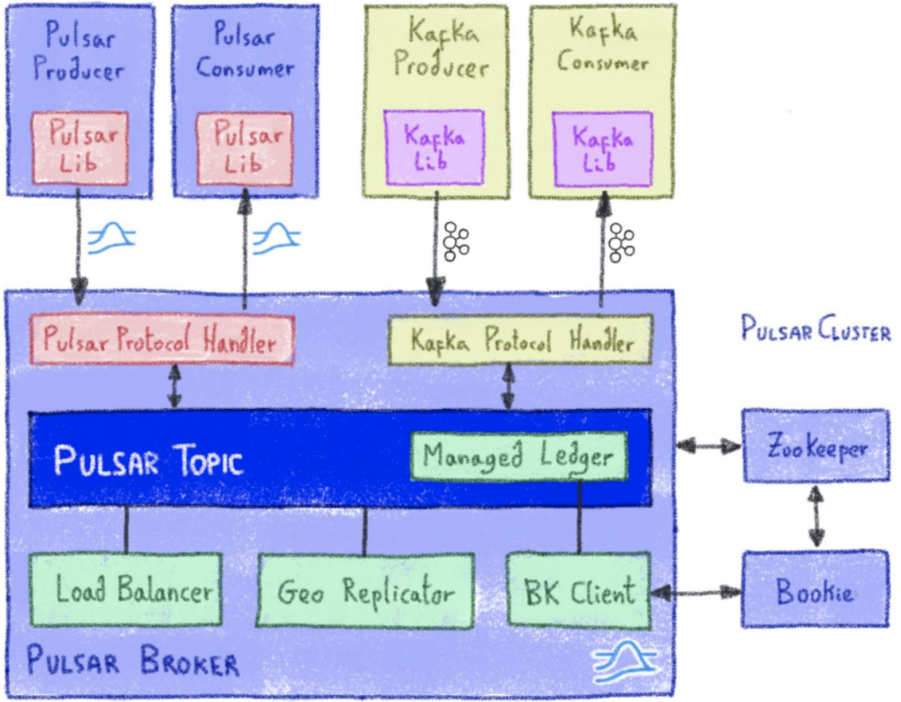
Schema Registry



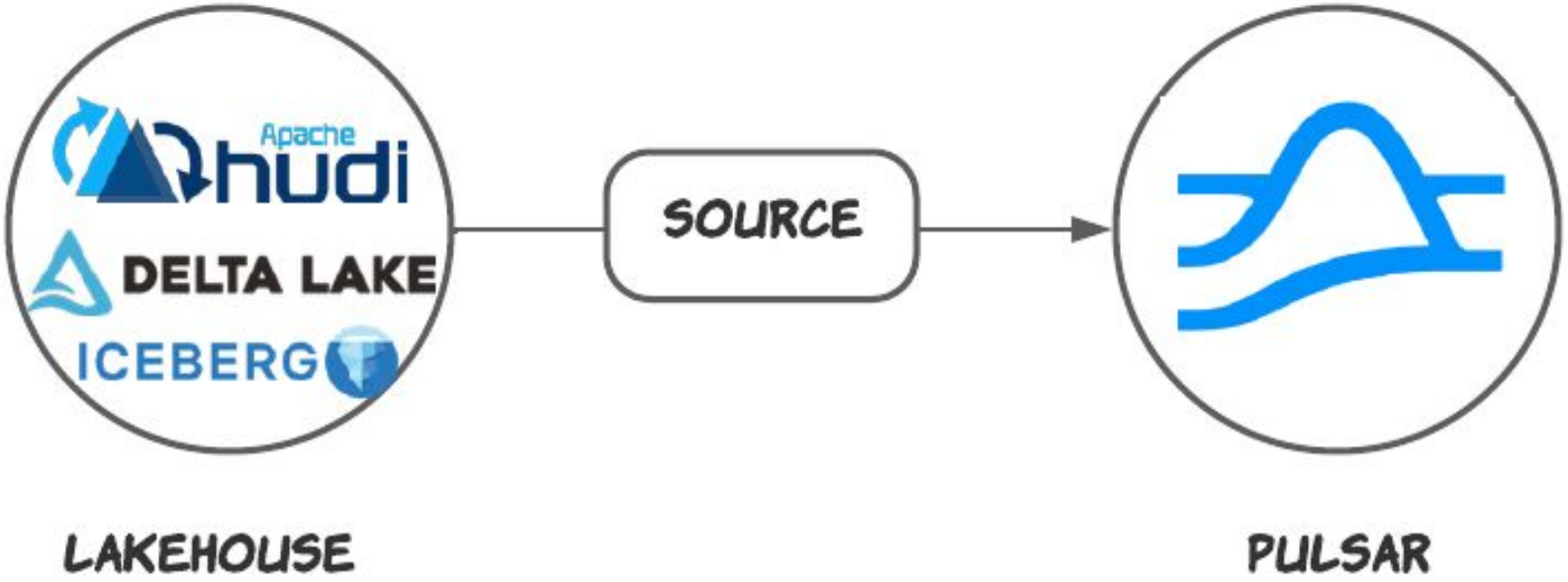
MQTT On Pulsar (MoP)



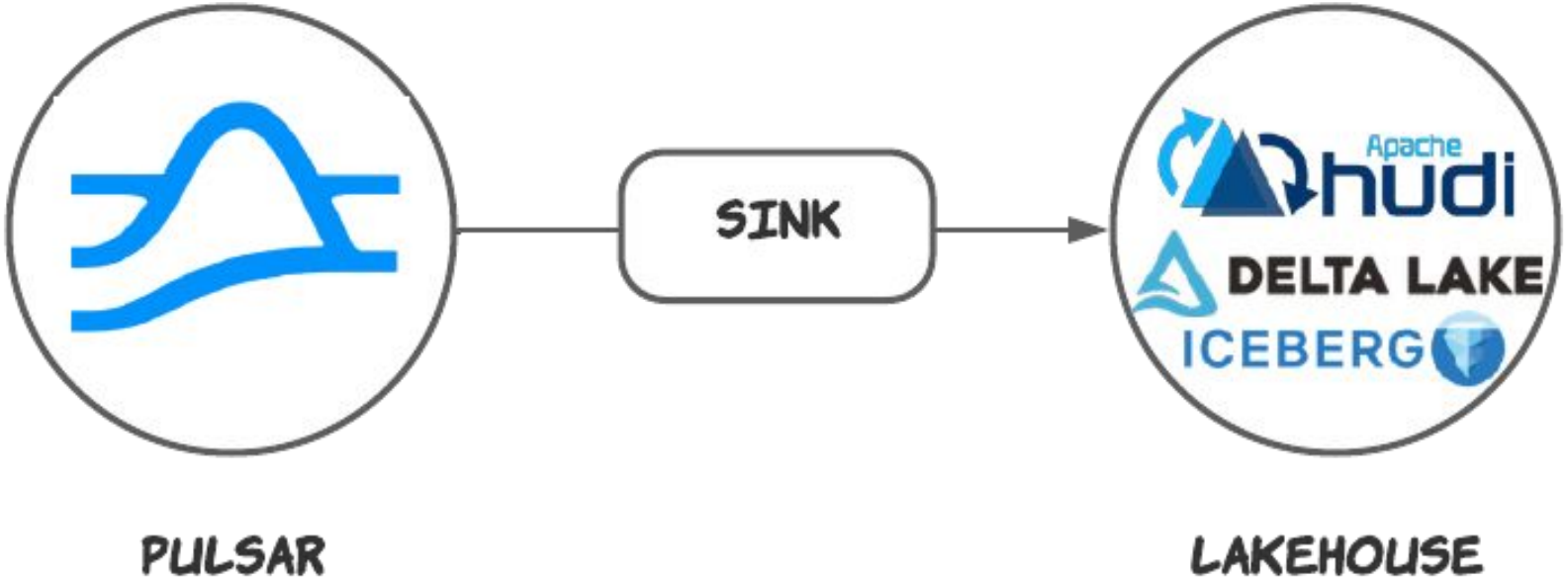
Kafka On Pulsar (KoP)



Use Pulsar to Stream from Lakehouses



Use Pulsar to Stream to Lakehouses



ETL:
Streaming with Apache Spark

Building Spark SQL View

```
val dfPulsar = spark.readStream.format("pulsar")
      .option("service.url", "pulsar://pulsar1:6650")
      .option("admin.url", "http://pulsar1:8080")
      .option("topic", "persistent://public/default/weather")
      .load()

dfPulsar.printSchema()

val pQuery = dfPulsar.selectExpr("*")
             .writeStream.format("console")
             .option("truncate", false)
             .start()
```

Example Spark Code

```
val dfPulsar = spark.readStream.format("pulsar")
        .option("service.url", "pulsar://pulsar1:6650")
        .option("admin.url", "http://pulsar1:8080")
        .option("topic", "persistent://public/default/airquality").load()
val pQuery = dfPulsar.selectExpr("*").writeStream.format("parquet")
.option("truncate", false).start()
```

ENRICH, ML & ROUTE: Pulsar Functions

Why Pulsar Functions for Microservices?

Desired Characteristic	Pulsar Functions...
Highly maintainable and testable	• Small pieces of code in Java, Python, or Go. • Easily maintained in source control repositories and tested with existing frameworks automatically.
Loosely coupled with other services	• Not directly linked to one another and communicate via messages.
Independently **deployable**	• Designed to be deployed independently
Can be developed by a small team	• Often developed by a single developer.
Inter-service **Communication**	• Support all message patterns using Pulsar as the underlying message bus.
Deployment & **Composition**	• Can run as individual threads, processes, or K8s pods. • Function Mesh allows you to deploy multiple Pulsar Functions as a single unit.

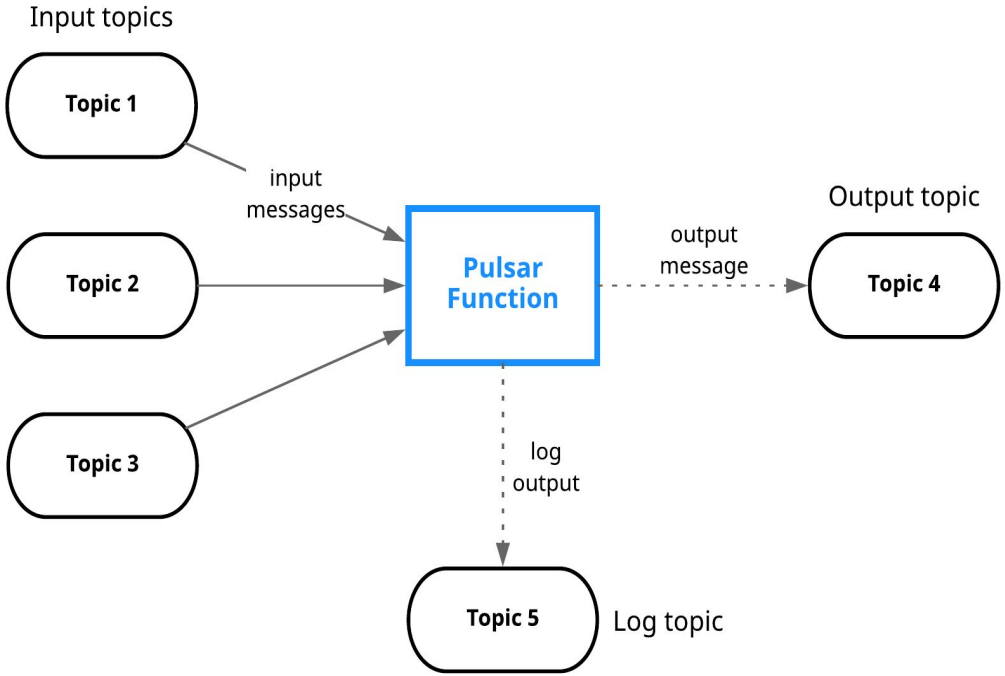

DZone
DZONE TREND REPORT
MARCH 2022
Enterprise AI
Machine Learning, Design Paradigms,
and Operational Impact

Pulsar Functions

- Route
- Enrich
- Convert
- Lookups
- Run Machine Learning

- Logging
- Auditing
- Parse
- Split
- Convert

Pulsar Functions



- Consume messages from one or more Pulsar topics.
- Apply user-supplied processing logic to each message.
- Publish the results of the computation to another topic.
- Support multiple programming languages (**Java**, Python, Go)
- Can leverage 3rd-party libraries to support the ***execution of ML models on the edge***.

Java Pulsar Function

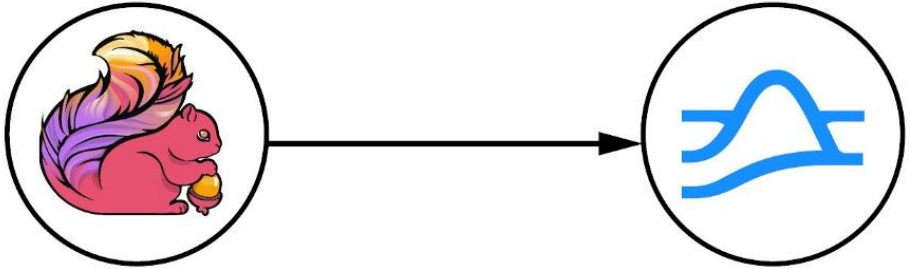
```java
/**
 * Example function that uses the built in publish function in the context
 * to publish to a desired topic based on config.
 */
public class PublishFunction implements Function<String, Void> {
    @Override
    public Void process(String input, Context context) {
        String publishTopic = (String) context.getUserConfigValueOrDefault("publish-topic", "publishtopic");
        String output = String.format("%s!", input);
        try {
            context.newOutputMessage(publishTopic, Schema.STRING).value(output).sendAsync();
        } catch (PulsarClientException e) {
            context.getLogger().error(e.toString());
        }
        return null;
    }
}
```

Python Pulsar Function

```
from pulsar import Function
import json
class Chat(Function):
    def __init__(self):
        pass

    def process(self, input, context):
        logger = context.get_logger()
        logger.info("Message Content: {0}".format(input))
        msg_id = context.get_message_id()
        row = { }
        row['id'] = str(msg_id)
        json_string = json.dumps(row)
        return json_string
```

CONTINUOUS ANALYTICS: Apache Flink SQL

Flink SQL -> Apache Pulsar



Announcing the Flink-Pulsar Sink Connector

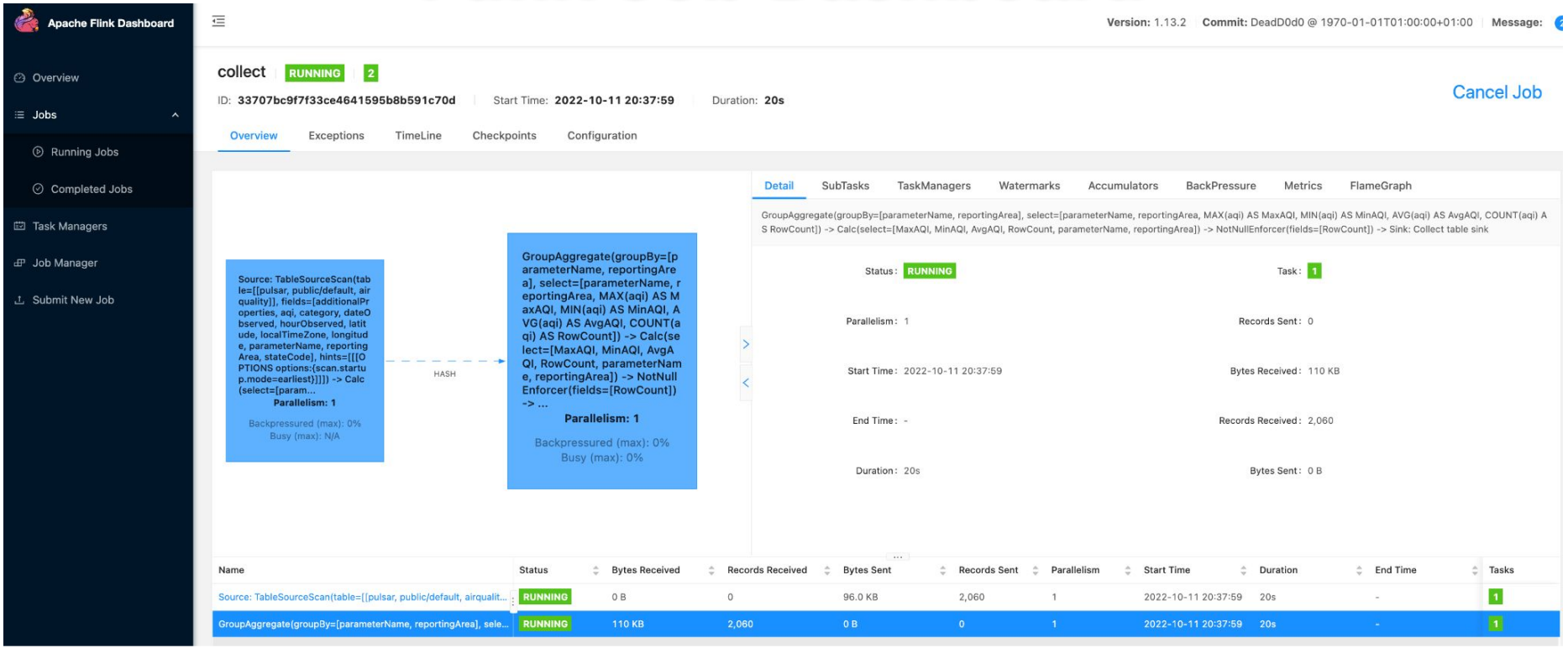
Flink Jobs

```
              STRING | true |      |        |            |
|            stateCode |
              STRING | true |      |        |            |
+--------------------+--------------------------------------------------------------------------------
--------------------+------+-----+--------+-----------+
11 rows in set

Flink SQL> select max(aqi) as MaxAQI, min(aqi) as MinAQI, avg(aqi) as AvgAQI, count(aqi) as RowCou
nt, parameterName, reportingArea from airquality group by parameterName, reportingArea;
>
```

Running Job List

Job Name	Start Time	Duration	End Time	Tasks	Status
collect	2022-10-11 20:34:59	3s	-	2 2	RUNNING

Flink Job Dashboard



APPS:
Java and Python Microservices

Flink SQL

```
select aqi, parameterName, dateObserved, hourObserved, latitude,
longitude, localTimeZone, stateCode, reportingArea from
airquality;

select max(aqi) as MaxAQI, parameterName, reportingArea from
airquality group by parameterName, reportingArea;

select max(aqi) as MaxAQI, min(aqi) as MinAQI, avg(aqi) as
AvgAQI, count(aqi) as RowCount, parameterName, reportingArea
from airquality group by parameterName, reportingArea;
```

```
2022-08-24 13:41:28.961  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: PULSAR 6790:7:-1:0
2022-08-24 13:41:28.967  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: Found existing Denver-Boulder
2022-08-24 13:41:28.971  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: Updated to ScyllaDB table true for
2022-08-24 13:41:28.977  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: MQTT
2022-08-24 13:41:28.980  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: AMQP
2022-08-24 13:41:28.983  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: KAFKA
2022-08-24 13:41:28.983  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: PM10=88 for CO Denver-Boulder
2022-08-24 13:41:29.001  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: PULSAR 6790:8:-1:0
2022-08-24 13:41:29.006  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: Found existing Denver-Boulder
2022-08-24 13:41:29.008  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: Updated to ScyllaDB table true for
2022-08-24 13:41:29.014  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: MQTT
2022-08-24 13:41:29.016  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
: AMQP
2022-08-24 13:41:29.018  INFO 18014 --- [   scheduling-1] d.datainmotion.airquality.AirQualityApp
```

Pulsar Spring Code

```
@Autowired
private PulsarTemplate<Observation> pulsarTemplate;

this.pulsarTemplate.setSchema(Schema.JSON(Observation.class));

MessageId msgid = pulsarTemplate.newMessage(observation)
     .withMessageCustomizer((mb) -> mb.key(uuidKey.toString()))
     .send();


@PulsarListener(subscriptionName = "aq-spring-reader", subscriptionType = "shared",
schemaType = SchemaType.JSON, topics = "persistent://public/default/aq-pm25")
void echoObservation(Observation message) {
  this.log.info("PM2.5 Message received: {}", message);
}
```

Pulsar Spring Configuration

```
spring:
   pulsar:
     client:
       service-url: pulsar+ssl://sn-academy.sndevadvocate.snio.cloud:6651
       auth-plugin-class-name: org.apache.pulsar.client.impl.auth.oauth2.AuthenticationOAuth2
       auth-params:
"{\"privateKey\":\"file:///tmp/tspann.json\",\"issuerUrl\":\"https://auth.streamnative.cloud/\",
\"audience\":\"urn:sn:pulsar:sn1:inst1\"}"
     producer:
       batching-enabled: false
       send-timeout-ms: 90000
       producer-name: airqualityjava
       topic-name: persistent://public/default/airquality
```

Pulsar Spring Reactive

```
ReactivePulsarClient reactivePulsarClient =
        AdaptedReactivePulsarClientFactory.create(pulsarClient);
ReactiveMessageSender<String> messageSender = reactivePulsarClient
            .messageSender(Schema.STRING)
            .topic(topicName)
            .maxInflight(100)
            .producerName("reactiveProducer")
            .sendTimeout(Duration.ofSeconds(60L))
            .accessMode(ProducerAccessMode.Shared)
            .build();
UUID uuidKey = UUID.randomUUID();
String message = messageBuilder();\
MessageSpecBuilder<String> messageSpecBuilder =
MessageSpec.builder(message).key(uuidKey.toString());
Mono<MessageId> messageId = messageSender.sendMessage(
                    Mono.just(messageSpecBuilder.build()));
```

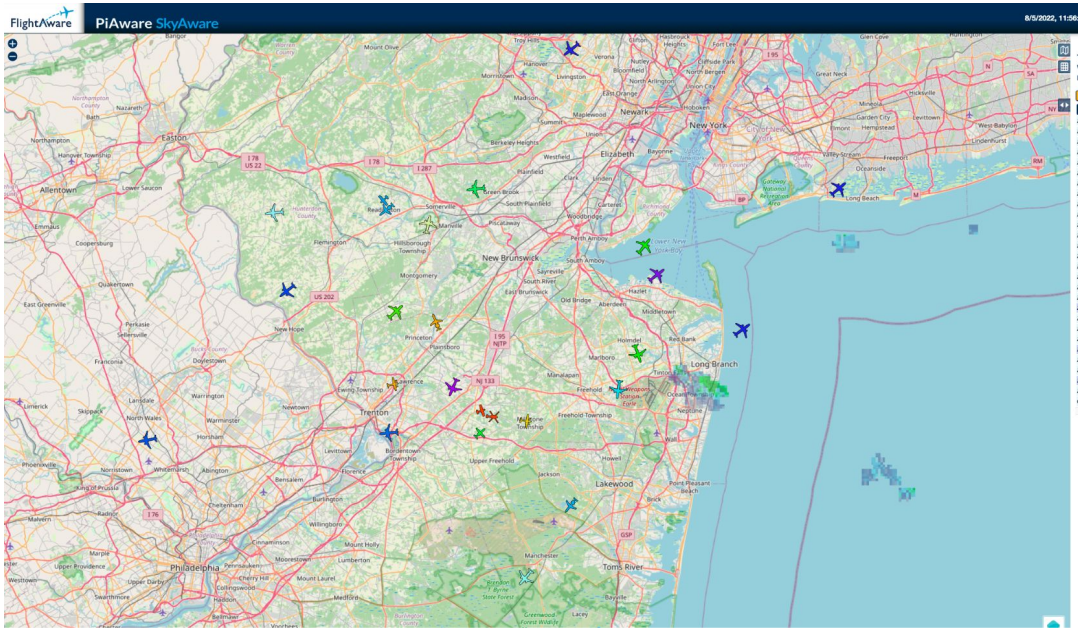
DASHBOARDS:
Enabled with SQL and WebSockets

```
<script>

$(document).ready(function() {
    var t = $('#example').DataTable();

var wsUri = "ws://pulsar1:8080/ws/v2/consumer/persistent/public/default/chatresult2/chatrreader?
subscriptionType=Shared&receiverQueueSize=500";
websocket = new WebSocket(wsUri);
websocket.onopen = function(evt) {
   console.log('open');
};
websocket.onerror = function(evt) {console.log('ERR', evt)};
websocket.onmessage = function(evt) {
    console.log(evt.data);
    var dataPoints = JSON.parse(evt.data);
    if ( dataPoints === undefined || dataPoints == null || dataPoints.payload === undefined ||
dataPoints.payload == null ) {
            return;
    }

</script>
```

```
var wsUri =
"ws://pulsar1:8080/ws/v2/consumer/persistent/public/default/pm25/a25-rdr?subscriptionType=Shared&receiverQueueSize=500";

websocket = new WebSocket(wsUri);

websocket.onopen = function(evt) {

console.log('open');

};

websocket.onerror = function(evt) {console.log('ERR', evt)};

websocket.onmessage = function(evt) {

var dataPoints = JSON.parse(evt.data);
```

```
if ( dataPoints === undefined || dataPoints == null ||
dataPoints.payload === undefined || dataPoints.payload == null
) { return; }

if (IsJsonString(atob(dataPoints.payload))) {

var pulsarMessage = JSON.parse(atob(dataPoints.payload));

if ( pulsarMessage === undefined || pulsarMessage == null ) {

return;

}

var publishTime = " ";
if ( !isEmpty(dataPoints.publishTime) ) {
  publishTime = dataPoints.publishTime;
```

PULSAR SUMMIT | Pulsar Summit Asia 2022



Apache Pulsar - Websockets - Data Tables - Aircraft (ADS-B)

Show 10 entries Search:

Altitude	Ground Speed	Mach	Hex	Flight #	Latitude	Longitude	PublishTime
2925	112		a2ce7f				2022-09-01T16:31:08.262-04:00
9200	319.3		abece7		40.315018	-73.845703	2022-09-01T16:31:08.245-04:00
36000	441.6		a48ac4	FFT1501	40.141762	-75.246048	2022-09-01T16:31:08.235-04:00
36900	435.4	0.736	a47303	EJA386	40.389513	-73.510083	2022-09-01T16:31:08.225-04:00
27225	413.3		a7de93		40.36093	-75.34274	2022-09-01T16:31:08.215-04:00
24000	440.6	0.756	a1373d	DAL2229	40.157636	-75.056284	2022-09-01T16:31:08.205-04:00
7475			a5a683		40.549622	-74.091736	2022-09-01T16:31:08.195-04:00
23400			a4c8de				2022-09-01T16:31:08.185-04:00
4300	144		a35585	N31386	40.308243	-74.563843	2022-09-01T16:31:08.175-04:00
6975	292.5		abc12a		40.50594	-74.250474	2022-09-01T16:31:08.165-04:00

Altitude	Ground Speed	Mach	Hex	Flight #	Latitude	Longitude	PublishTime

PULSAR SUMMIT | Pulsar Summit Asia 2022

SQL Query Result (Table)
Refresh: 1 s Page: Last of 42

MaxTemp	location
84.0	Milton, Whiting Field South, ~
76.0	MCAS New River, NC
71.0	Lakehurst Naval Air Station, ~
83.0	New Orleans, New Orleans Lake~
73.0	Marion Municipal Airport, IN
75.0	Annapolis, United States Nava~
83.0	Piney Island, Bt-11 Bombing R~
82.0	New Orleans Naval Air Station~
64.0	Fallon, Naval Air Station, NV
81.0	Corpus Christi, Corpus Christ~
70.0	Oceanside, Camp Pendleton, Ma~
78.0	Norfolk, Naval Air Station, VA
69.0	San Diego, Miramar MCAS/Mitsc~
85.0	Jacksonville Naval Air Statio~
84.0	El Centro, Naval Air Facility~
85.0	Pensacola Naval Air Station, ~
92.0	Key West Naval Air Station, FL
86.0	Kingsville, Naval Air Station~
66.0	China Lake, Naval Air Facilit~

StreamNative

User:

Question:

Contact Info:

Send to Pulsar

Sentiment	Publish Time	Message	By	--
Positive	2022-08-30T17:58:21.235-04:00	Is Pulsar the best streaming platform ever?	Tim Spann	Tim @sparkdev
Positive	2022-08-30T17:58:22.356-04:00	Is Pulsar the best streaming platform ever?	Tim Spann	Tim @sparkdev

Location	Obs Time	Weather	Temp	Windchill	Lat/Long	Pressure	Publish Time	[--]
Dayton, Cox Dayton International Airport, OH	Last Updated on Oct 11 2022, 6:56 pm EDT	Mostly Cloudy	68.0 F (20.0 C)		39.90611,-84.21861	1020.9 mb	2022-10-11T20:21:01.76-04:00	
Dublin, W H 'Bud' Barron Airport, GA	Last Updated on Oct 5 2022, 3:15 pm EDT	Fair	81.0 F (27.0 C)		32.56444,-82.985		2022-10-11T20:21:01.767-04:00	
Dubuque Regional Airport, IA	Last Updated on Oct 11 2022, 5:53 pm CDT	Fog/Mist	62.0 F (16.7 C)	60 F (16 C)	42.39837,-90.7091	1009.8 mb	2022-10-11T20:21:01.776-04:00	
Washington/Reagan National Airport, DC, VA	Last Updated on Oct 11 2022, 6:52 pm EDT	A Few Clouds	65.0 F (18.3 C)		38.84833,-77.03417	1025.6 mb	2022-10-11T20:21:01.781-04:00	
Chester Catawba Regional Airport, SC	Last Updated on Oct 11 2022, 6:35 pm EDT	Mostly Cloudy	66.0 F (19.0 C)		34.789,-81.196		2022-10-11T20:21:01.788-04:00	

PULSAR SUMMIT | Pulsar Summit Asia 2022

StreamNative

PM 2.5

Show 10 entries Search:

publishTime	Date	Hour	TZ	Area	State	Lat	Long	PM 2.5 AQI
2022-04-26T12:02:37.299-04:00	2022-04-26	10	EST	Atlanta	GA	33.65	-84.43	63
2022-04-26T12:02:47.304-04:00	2022-04-26	10	EST	Atlanta	GA	33.65	-84.43	63
2022-04-26T12:02:57.294-04:00	2022-04-26	10	EST	Atlanta	GA	33.65	-84.43	63
2022-04-26T12:03:07.295-04:00	2022-04-26	10	EST	Atlanta	GA	33.65	-84.43	63
2022-04-26T12:05:24.794-04:00	2022-04-26	10	EST	Atlanta	GA	33.65	-84.43	63
2022-04-26T12:05:25.853-04:00	Sat, 09 Apr 2022 12:45:34 GMT	-1	EST	MiraculousMess::US		42.038597	-87.69545	0
2022-04-26T12:05:25.886-04:00	Sat, 09 Apr 2022 12:45:34 GMT	-1	EST	Fruitdale::US		39.778137	-105.11771	190
2022-04-26T12:05:25.949-04:00	Sat, 09 Apr 2022 12:45:34 GMT	-1	EST	Willow Glen West::US		37.304996	-121.91792	730
2022-04-26T12:05:26.024-04:00	Sat, 09 Apr 2022 12:45:34 GMT	-1	EST	Greengate school exterior::US		38.68429	-121.79665	340
2022-04-26T12:05:26.087-04:00	Sat, 09 Apr 2022 12:45:34 GMT	-1	EST	Golden Bear::US		37.87195	-122.27165	819

Showing 1 to 10 of 751 entries Previous 1 2 3 4 5 … 76 Next

PM 10

Show 10 entries Search:

publishTime	Date	Hour	TZ	Area	State	Lat	Long	PM 10 AQI
2022-10-06T12:31:49.978-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Field Rose::US		40.6398	-111.82251	70
2022-10-06T12:31:49.5-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Ashling::US		39.345142	-120.26349	140
2022-10-06T12:31:49.089-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Fall River Mills::US		41.007404	-121.440155	120
2022-10-06T12:31:48.825-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	LRAPA-Elmira HS::US		44.073837	-123.35885	310
2022-10-06T12:31:48.549-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Monte Vista Ave::US		37.93318	-122.534706	100
2022-10-06T12:31:48.28-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	aeris qualitas::US		34.42251	-119.72129	390
2022-10-06T12:31:47.797-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Monroe Municipal Campus::US		47.85146	-121.981575	880
2022-10-06T12:31:47.511-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Crown Road::US		37.94726	-122.57016	10
2022-10-06T12:31:46.967-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Isola way::US		37.935814	-122.094215	570
2022-10-06T12:31:46.707-04:00	Wed, 07 Sep 2022 00:48:57 GMT	-1	EST	Gym outside::US		48.056057	-122.27148	1520

Showing 1 to 10 of 1,000 entries Previous 1 2 3 4 5 … 100 Next

StreamNative

PM 2.5

Show 10 entries

Search:

publishTime	Date	Hour	TZ	Area	State	Lat	Long	PM 2.5 AQI
2022-06-30T15:13:52.041-04:00	2022-06-30	11	PST	Seattle-Bellevue-Kent Valley	WA	47.562	-122.3405	33
2022-06-30T15:13:52.062-04:00	2022-06-30	14	EST	Wake County	NC	35.878	-78.787	41
2022-06-30T15:13:59.101-04:00	2022-06-30	13	CST	Birmingham	AL	33.56	-86.75	44
2022-06-30T15:14:09.108-04:00	2022-06-30	11	PST	Redwood City	CA	37.48	-122.22	13
2022-06-30T15:14:19.104-04:00	2022-06-30	14	EST	Central	NJ	40.401	-74.325	52
2022-06-30T15:14:29.106-04:00	2022-06-30	13	CST	Houston-Galveston-Brazoria	TX	29.751	-95.351	16
2022-06-30T15:14:39.128-04:00	2022-06-30	14	EST	New York City Region	NY	40.8419	-73.8359	38
publishTime	**Date**	**Hour**	**TZ**	**Area**	**State**	**Lat**	**Long**	**PM 2.5 AQI**

Showing 1 to 7 of 7 entries

Previous 1 Next

CODE:
NiFi Flows, SQL, Python, Java

```
df.select("uuid","humidity","co2","cputempf","datetimestamp","ts").show(5,100)
+----------------------+--------+------+--------+-----------------------------------+----------+
|                  uuid|humidity|   co2|cputempf|                      datetimestamp|        ts|
+----------------------+--------+------+--------+-----------------------------------+----------+
|thrml_xlh_20220715135558|   36.56|1127.0|     106|2022-07-15 13:56:01.958974+00:00|1657893362|
|thrml_cuv_20220715135602|   36.69|1127.0|     107|2022-07-15 13:56:06.737358+00:00|1657893367|
|thrml_ebc_20220715135607|   36.71|1126.0|     108|2022-07-15 13:56:11.513152+00:00|1657893372|
|thrml_dug_20220715135612|   36.64|1126.0|     106|2022-07-15 13:56:16.292526+00:00|1657893377|
|thrml_hnj_20220715135617|   36.71|1125.0|     107|2022-07-15 13:56:21.069232+00:00|1657893382|
+----------------------+--------+------+--------+-----------------------------------+----------+
only showing top 5 rows
```

```
+--------+------+--------------------+--------+----------+--------------------+
|humidity|   co2|       datetimestamp|cputempf|        ts|                uuid|
+--------+------+--------------------+--------+----------+--------------------+
|   36.56|1127.0|2022-07-15 13:56: ...|     106|1657893362|thrml_xlh_2022071 ...|
|   36.69|1127.0|2022-07-15 13:56: ...|     107|1657893367|thrml_cuv_2022071 ...|
+--------+------+--------------------+--------+----------+--------------------+
```

Github

- https://github.com/tspannhw/FLiP-Py-ADS-B
- https://github.com/tspannhw/pulsar-adsb-function
- https://github.com/tspannhw/pulsar-weather-function
- https://github.com/tspannhw/spring-pulsar-airquality
- https://github.com/tspannhw/pulsar-airquality-function
- https://github.com/tspannhw/FLiP-Current22-LetsMonitorAllTheThings
- https://github.com/tspannhw/FLiP-Transit
- https://github.com/tspannhw/SmartTransit
- https://github.com/tspannhw/airquality-kafka-consumer

Content

- https://medium.com/@tspann/tracking-aircraft-in-real-time-with-open-source-554124125011
- https://medium.com/@tspann/parsing-weather-feeds-to-add-to-real-time-streams-ec5ecc2849fb
- https://medium.com/@tspann/using-the-new-spring-boot-apache-pulsar-integration-8a38447dce7b
- https://www.datainmotion.dev/2021/01/flank-real-time-transit-information-for.html
- https://www.clouddataops.dev/transit-monitoring
- https://www.datainmotion.dev/2020/07/ingesting-all-weather-data-with-apache.html

PULSAR SUMMIT | Thanks

https://www.linkedin.com/in/timothyspann

https://github.com/tspannhw

@PassDev

https://streamnative.io/pulsar-python/