

CLOUDERA

Building Modern Data Streaming Apps with NiFi, Flink and Kafka

Tim Spann
Principal Developer Advocate

8-June-2023

June 5-8, 2023 · Data platforms · Data engineering · Data management

BUDAPEST DATA FORUM

Co-hosted with the Budapest ML Forum



IoT

CLUSTERED & IOT
REAL-TIME MONITORING
OF YOUR HEART'S HEALTH.
THANKS DATA.

CLUSTERED



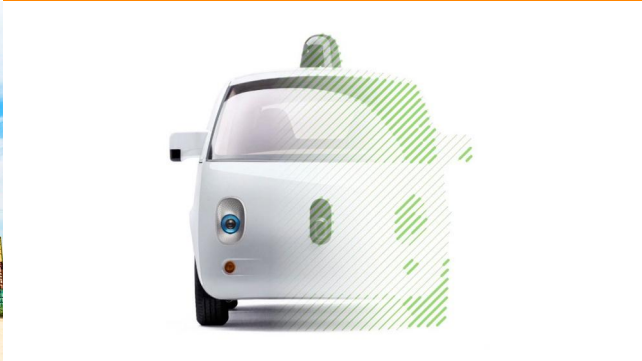
CLUSTERED

EDGE 2AI

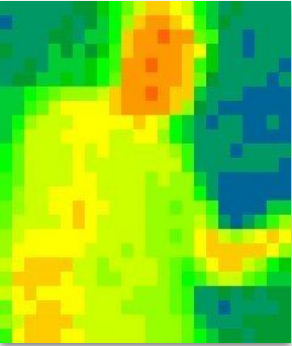
CLUSTERED

ENTERPRISE DATA CLOUD

CLUSTERED



FLiPN-FLaNK Stack



Tim Spann

@PaasDev // Blog: www.datainmotion.dev

Principal Developer Advocate.

Princeton Future of Data Meetup.

ex-Pivotal, ex-Hortonworks, ex-StreamNative, ex-PwC

<https://github.com/tspannhw/EverythingApacheNiFi>

<https://medium.com/@tspann>

Apache NiFi x Apache Kafka x Apache Flink x Java



FLiP Stack Weekly



<https://bit.ly/32dAJft>



This week in Apache NiFi, Apache Flink, Apache Pulsar, Apache Spark, Apache Iceberg, Python, Java and Open Source friends.

Future of Data - Princeton + Virtual



FUTURE
OF DATA

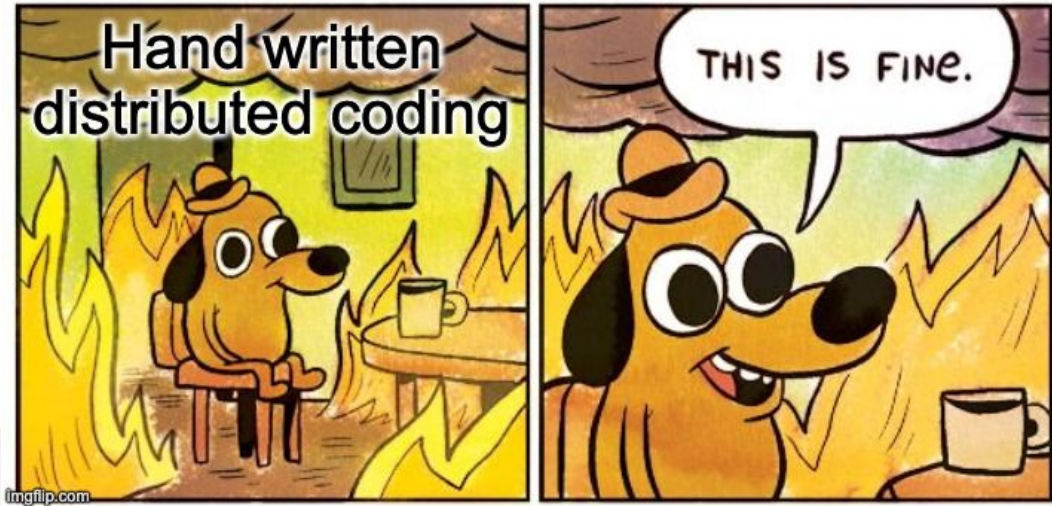
<https://www.meetup.com/futureofdata-princeton/>

From Big Data to AI to Streaming to Containers to Cloud to Analytics to Cloud Storage to Fast Data to Machine Learning to Microservices to ...



@PaasDev

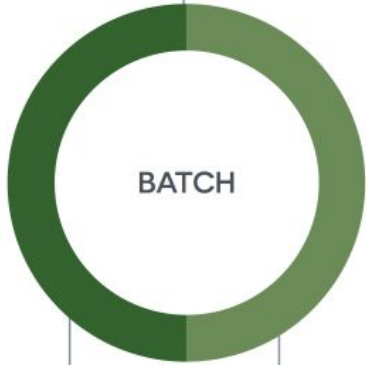
STREAMING



What is Real-Time?

> 1 HOUR

high throughput



BATCH

adhoc queries

monthly active users relevance for ads

10 MS - 1 SEC

approximate



REAL TIME

ad impressions count
hash tag trends

< 500 MS

latency sensitive



OLTP

deterministic workflows

fanout Tweets
search for Tweets

< 1 MS

low latency

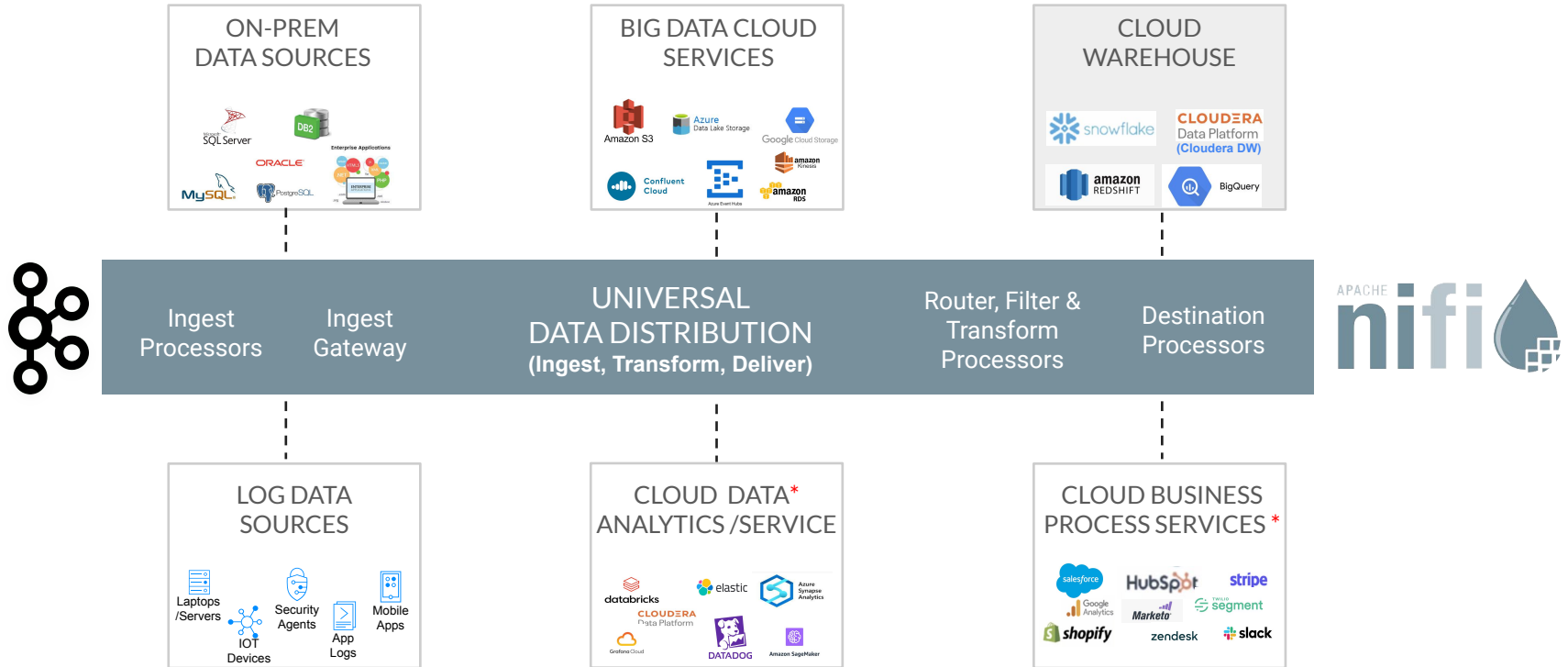


REAL
REAL TIME

Financial
Trading

Streaming From ... To ...

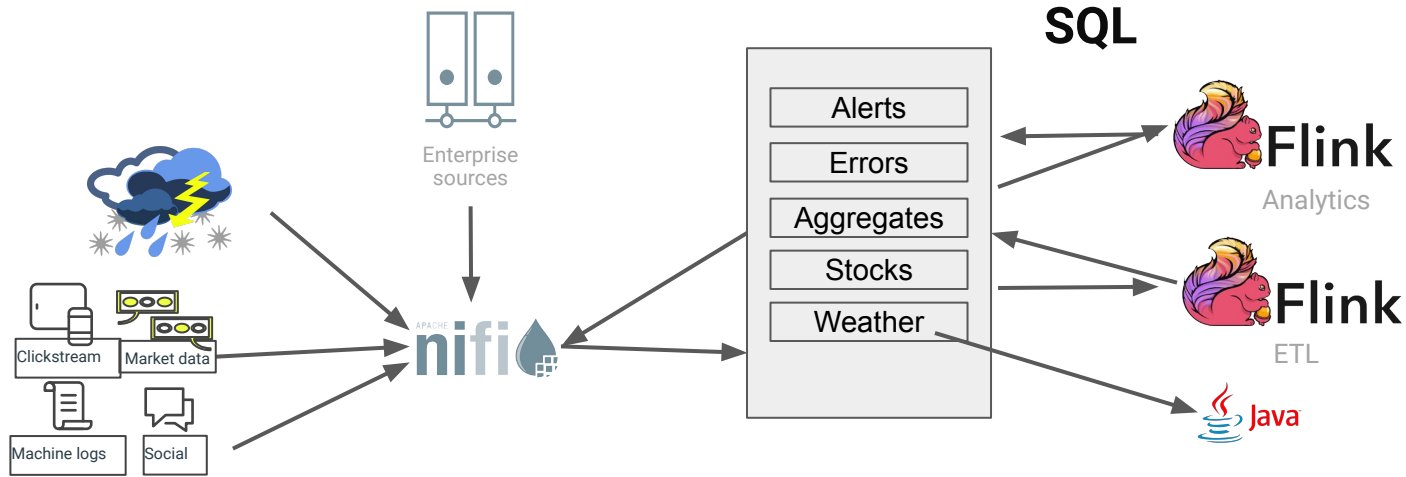
Data distribution as a first class citizen



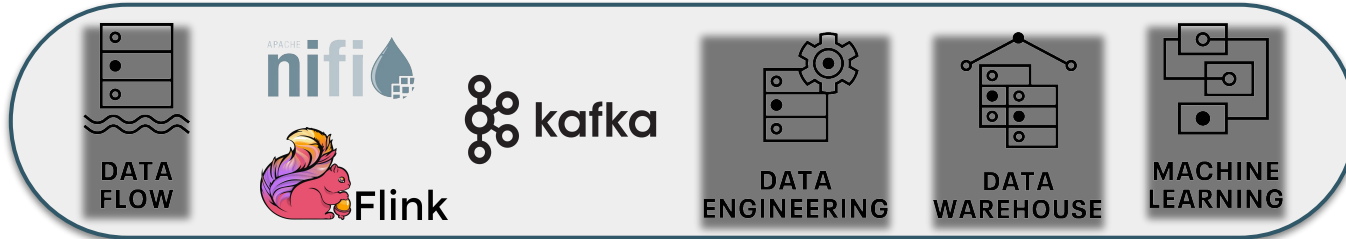
BUILDING REAL-TIME REQUIRES A TEAM



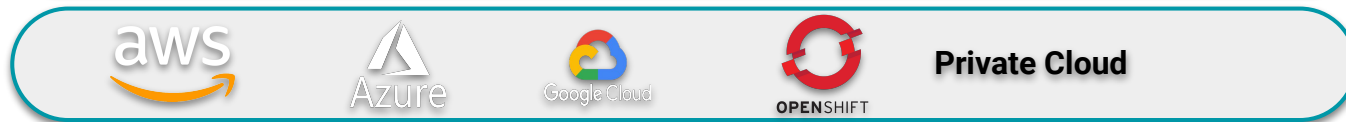
End to End Streaming Pipeline Example



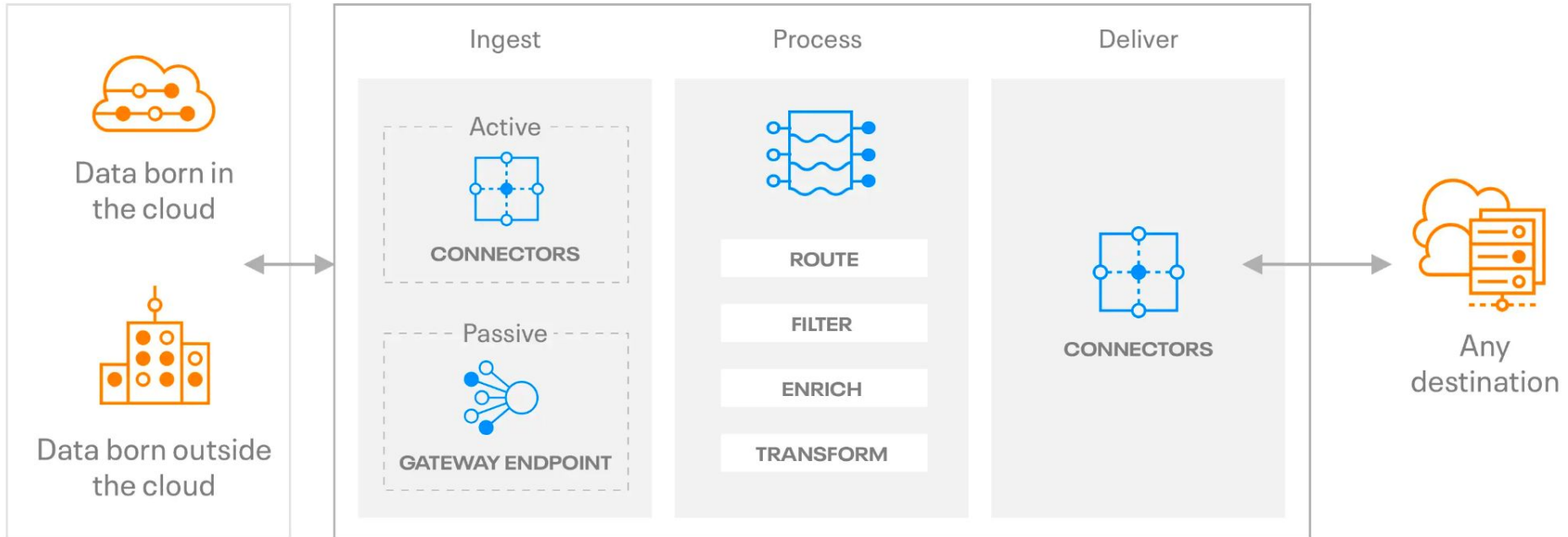
CDP: AN OPEN DATA LAKEHOUSE



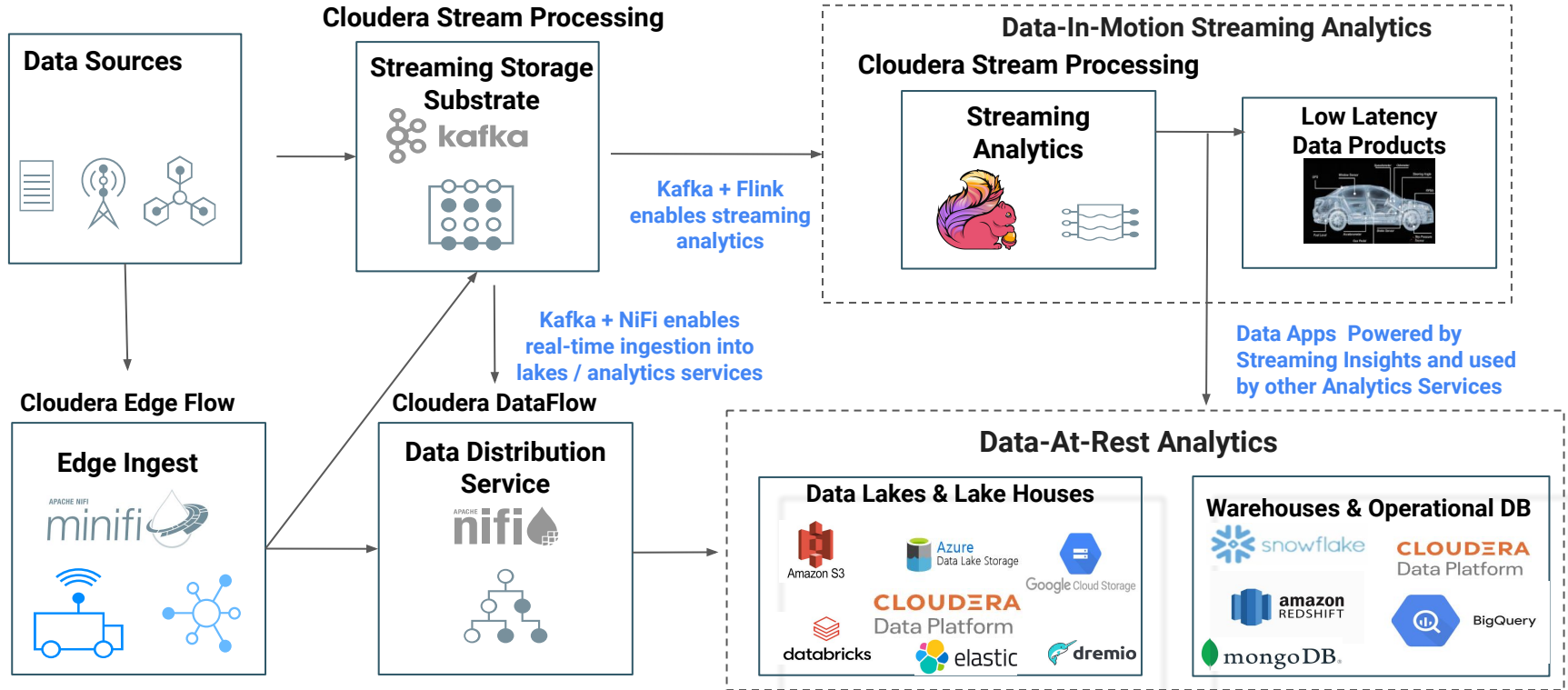
ICEBERG



DATAFLOW FOR THE PUBLIC CLOUD

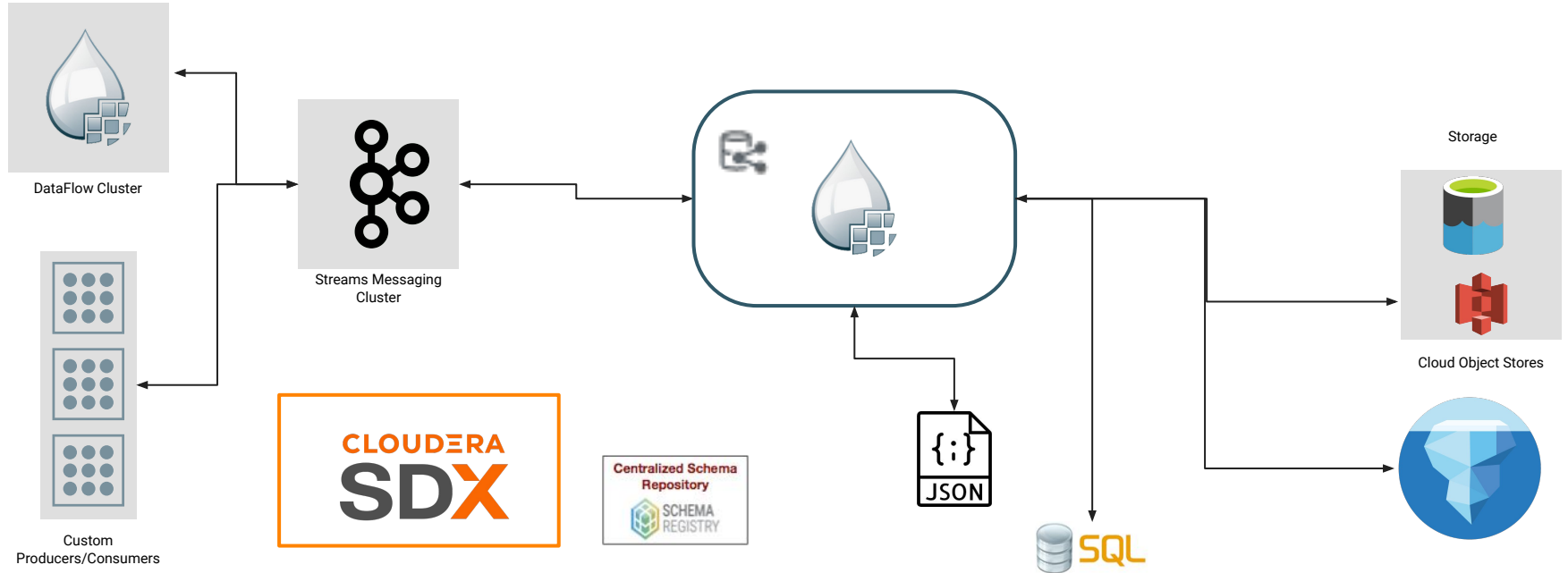


Analytics-in-Stream

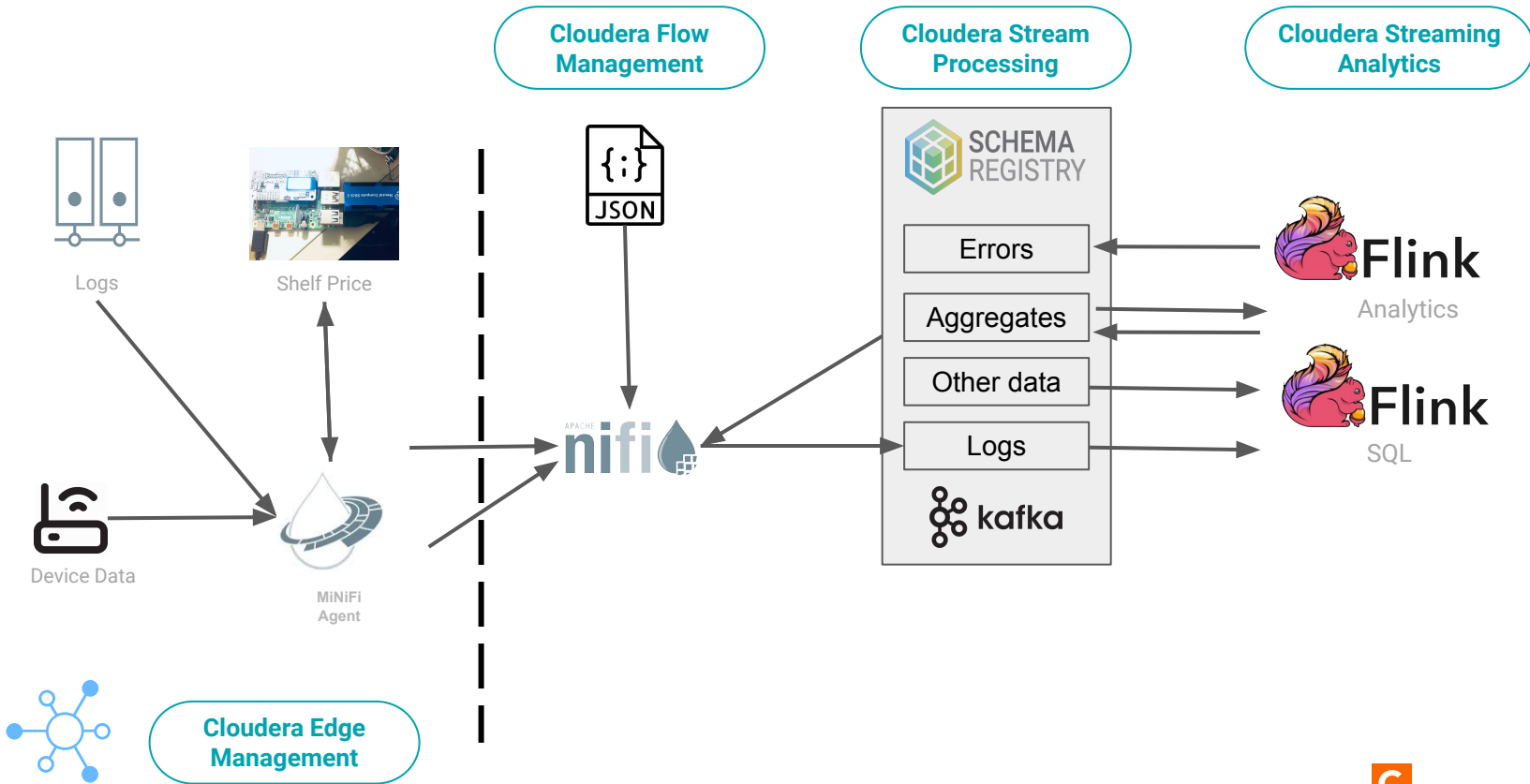


Retail Use Case: Ingest Retail Goods Prices

Codeless Data Movement



Pricing Pipeline



APACHE KAFKA

The background of the slide features three parallel diagonal stripes in shades of gray, running from the bottom-left towards the top-right. The stripes are of varying widths and are set against a white background.

What is Apache Kafka?

Distributed: horizontally scalable

Partitioned: the data is split-up and distributed across the brokers

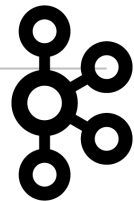
Replicated: allows for automatic failover

Unique: Kafka does not track the consumption of messages (the consumers do)

Fast: designed from the ground up with a focus on performance and throughput

Kafka was built at LinkedIn in 2011

Open sourced as an Apache project



Yes, Franz, It's Kafka

Let's do a metamorphosis on your data. Don't fear changing data.

You don't need to be a brilliant writer to stream data.



Franz Kafka was a German-speaking Bohemian novelist and short-story writer, widely regarded as one of the major figures of 20th-century literature. His work fuses elements of realism and the fantastic.

[Wikipedia](#)



What is Can You Do With Apache Kafka?

Web site activity: track page views, searches, etc. in real time

Events & log aggregation: particularly in distributed systems where messages come from multiple sources

Monitoring and metrics: aggregate statistics from distributed applications and build a dashboard application

Stream processing: process raw data, clean it up, and forward it on to another topic or messaging system

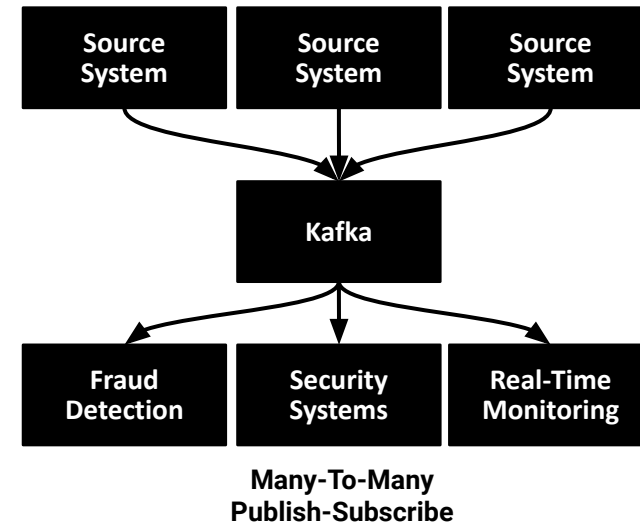
Real-time data ingestion: fast processing of a very large volume of messages

Kafka Terms

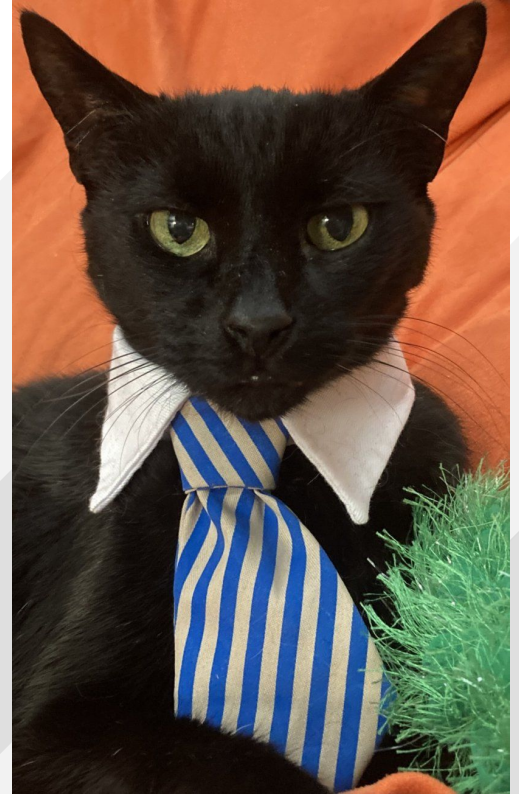
- Kafka is a publish/subscribe messaging system comprised of the following components:
 - **Topic:** a message feed
 - **Producer:** a process that publishes messages to a topic
 - **Consumer:** a process that subscribes to a topic and processes its messages
 - **Broker:** a server in a Kafka cluster

- Highly reliable distributed messaging system
- Decouple applications, enables many-to-many patterns
- Publish-Subscribe semantics
- Horizontal scalability
- Efficient implementation to operate at speed with big data volumes
- Organized by topic to support several use cases

EVENTS



APACHE FLINK



Flink SQL



- Streaming Analytics
- Continuous SQL
- Continuous ETL
- Complex Event Processing
- Standard SQL Powered by Apache Calcite

The screenshot displays the Apache Flink Dashboard interface. On the left is a dark sidebar with navigation options: Overview, Jobs, Running Jobs, Completed Jobs, Task Managers, Job Manager, and Submit New Job. The main area shows a job named 'xendochial_royce' in a 'RUNNING' state, started on 2021-04-07 at 10:08:37 with a duration of 3h 6m 21s. Below the job name is a job graph with a source node 'kafkaSource: weath...' and a sink node 'Writeback Process -> Sink: WEATHER_HTS_SINK'. The sink node is highlighted with a blue box and shows 'Parallelism: 1'. To the right of the graph is a detailed view of the sink task, showing its status as 'RUNNING', parallelism of 1, and various metrics like records sent and received. At the bottom, a table lists the job's components:

Name	Status	Bytes Received	Records Received	Bytes Sent	Records Sent	Parallelism	Start Time	Tasks
Source: kafkaSource: weath... -> Kafka TS assigner -> SourceConnectorTabletDe...	RUNNING	0 B	0	8.33 MB	8,000	1	2021-04-07	1
Source: ConnectSource	RUNNING	0 B	0	0 B	0	1	2021-04-07	1
Writeback Process -> Sink: WEATHER_HTS_SINK	RUNNING	8.64 MB	8,000	0 B	0	1	2021-04-07	1

<https://www.datainmotion.dev/2021/04/cloudera-sql-stream-builder-ssb-updated.html>

Flink SQL

Key Takeaway: Rich SQL grammar with advanced time and aggregation tools

```
-- specify Kafka partition key on output
SELECT foo AS _eventKey FROM sensors

-- use event time timestamp from kafka
-- exactly once compatible
SELECT eventTimestamp FROM sensors

-- nested structures access
SELECT foo.'bar' FROM table; -- must quote nested
column

-- timestamps
SELECT * FROM payments
WHERE eventTimestamp > CURRENT_TIMESTAMP-interval
'10' second;

-- unnest
SELECT b.*, u.*
FROM bgp_avro b,
UNNEST(b.path) AS u(pathitem)
```

```
-- aggregations and windows
SELECT card,
MAX(amount) as theamount,
TUMBLE_END(eventTimestamp, interval '5' minute) as
ts
FROM payments
WHERE lat IS NOT NULL
AND lon IS NOT NULL
GROUP BY card,
TUMBLE(eventTimestamp, interval '5' minute)
HAVING COUNT(*) > 4 -- >4==fraud

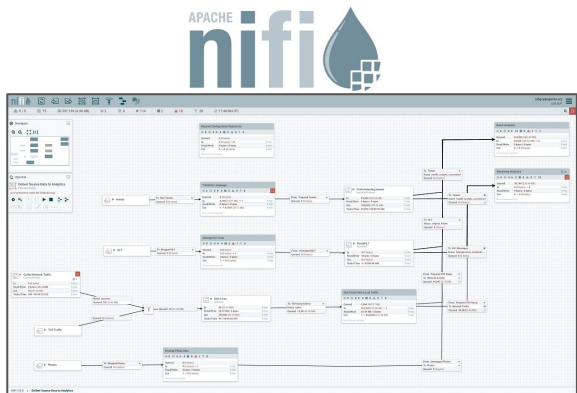
-- try to do this ksql!
SELECT us_west.user_score+ap_south.user_score
FROM kafka_in_zone_us_west us_west
FULL OUTER JOIN kafka_in_zone_ap_south ap_south
ON us_west.user_id = ap_south.user_id;
```

DATAFLOW APACHE NIFI

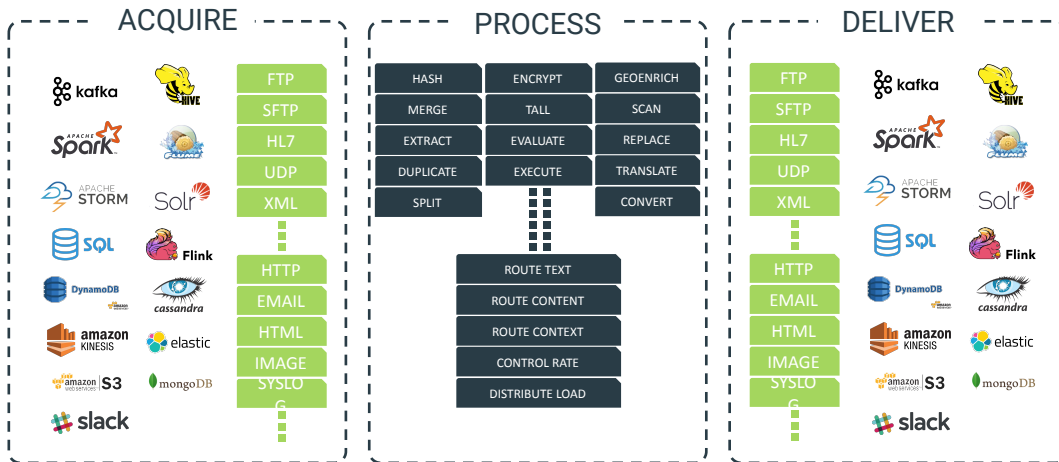
The background of the slide features three broad, parallel diagonal stripes in shades of light gray, extending from the bottom-left towards the top-right.

Apache NiFi

Enable easy ingestion, routing, management and delivery of any data anywhere (Edge, cloud, data center) to any downstream system with built in end-to-end security and provenance



Advanced tooling to industrialize flow development
(Flow Development Life Cycle)



- Over 300 Prebuilt Processors
- Easy to build your own
- Parse, Enrich & Apply Schema
- Filter, Split, Merger & Route
- Throttle & Backpressure

- Guaranteed Delivery
- Full data provenance from acquisition to delivery
- Diverse, Non-Traditional Sources
- Eco-system integration

Provenance

Displaying 13 of 104
Oldest event available: 11/15/2016 13:34:50 EST Showing the most recent events.

ConsumeKafka by component name 🔍

Date/Time	Type	FlowFile UUID	Size	Component Name	Component Type	
11/15/2016 13:35:03.8...	RECEIVE	379fc4f6-60e0-4151-9743-28...	44 bytes	ConsumeKafka	ConsumeKafka	🔗 →
11/15/2016 13:35:02.7...	RECEIVE	78f8c38b-89fc-4d00-a8d8-51...	44 bytes	ConsumeKafka	ConsumeKafka	🔗 →
11/15/2016 13:35:01.6...	RECEIVE	2bcd5124-bb78-489f-ad8a-7...	44 bytes	ConsumeKafka	ConsumeKafka	🔗 →

- Tracks data at each point as it flows through the system
- Records, indexes, and makes events available for display
- Handles fan-in/fan-out, i.e. merging and splitting data
- View attributes and content at given points in time

Provenance Event

DETAILS ATTRIBUTES CONTENT

Attribute Values

filename
328717796819631
No value previously set

kafka.offset
44815
No value previously set

kafka.partition
6
No value previously set

kafka.topic
nifi-testing
No value previously set

path
/
No value previously set

uuid
2057112-2852-4c6f-8096-10520731-05

Extensibility

- Built from the ground up with extensions in mind
- Service-loader pattern for...
 - Processors
 - Controller Services
 - Reporting Tasks
 - Prioritizers
- Extensions packaged as NiFi Archives (NARs)
 - Deploy NiFi lib directory and restart
 - Same model as standard components

The screenshot displays the IntelliJ IDEA IDE interface. On the left, a project tree shows the structure of the 'nifi-mxnetinference-processors' project, including source code, resources, and a target directory. The main editor shows the 'InferenceProcessorTest' class with the following code:

```
public class InferenceProcessorTest {  
    private TestRunner testRunner;  
    @Before  
    public void init() {  
        testRunner = TestRunners.newTestRunner(InferenceProcessor.class);  
    }  
    private String pathOfResourceURL r = this.getResourceURL(uri = r.toURL());  
    return Paths.get(uri);  
}  
    private void runAndAssertValidDataUpdateAttribute(InferenceProcessor processor, List<MockFlowFile> sourceFiles, List<MockFlowFile> sinkFiles) {  
        testRunner.run();  
        testRunner.assertValidDataUpdateAttribute(processor, sourceFiles, sinkFiles);  
    }  
    for (MockFlowFile mockFlowFile : sourceFiles) {  
        testRunner.assertValidDataUpdateAttribute(processor, mockFlowFile, sinkFiles);  
    }  
}
```

On the right, a flow diagram illustrates the execution of a test. It consists of four nodes connected by arrows:

- LinkProcessor**: LinkProcessor (LinkProcessor). In: 0 (0 bytes), 5 min; Read/Write: 0 bytes / 31.45 KB, 5 min; Out: 2 (31.45 KB), 5 min; Tasks/Time: 2 / 00:00:04.808, 5 min.
- UpdateAttribute**: UpdateAttribute (UpdateAttribute). In: 2 (31.45 KB), 5 min; Read/Write: 0 bytes / 0 bytes, 5 min; Out: 2 (31.45 KB), 5 min; Tasks/Time: 2 / 00:00:00.005, 5 min.
- PutHDFS**: PutHDFS (PutHDFS). In: 2 (31.45 KB), 5 min; Read/Write: 31.45 KB / 0 bytes, 5 min; Out: 0 (0 bytes), 5 min; Tasks/Time: 2 / 00:00:00.603, 5 min.

Each node also shows a 'Name success' status and 'Queued 0 (0 bytes)'.

Parquet Reader/Writers

- Native Record Processors for Apache Parquet Files!
- CSV <-> Parquet
- XML <-> Parquet
- AVRO <-> Parquet
- JSON <-> Parquet
- More...

Property	Value
Record Reader	JsonTreeReader
Record Writer	ParquetRecordSetWriter
Merge Strategy	Bin-Packing Algorithm
Correlation Attribute Name	No value set
Attribute Strategy	Keep Only Common Attributes
Minimum Number of Records	
Maximum Number of Records	
Minimum Bin Size	
Maximum Bin Size	
Max Bin Age	
Maximum Number of Bins	

Add Controller Service

Requires Controller Service
RecordReaderFactory 1.10.0.2.0.0.0-35 from org.apache.nifi - nifi-standard-services-api-nar

Compatible Controller Services

ParquetReader 1.10.0.2.0.0-35

Controller Service Name

ParquetReader

Bundle

org.apache.nifi - nifi-parquet-nar

Tags

reader, record, parse, row, parquet

Description

CANCEL CREATE

ReadyFlow Gallery

- Cloudera provided flow definitions
- Cover most common data flow use cases
- Optimized to work with CDP sources/destinations
- Can be deployed and adjusted as needed

The screenshot displays the Cloudera ReadyFlow Gallery interface. On the left is a dark sidebar with navigation options: Dashboard, Catalog, ReadyFlow Gallery (selected), and Environments. The main content area is titled 'ReadyFlow Gallery' and features a search bar. Below the search bar, there are six data flow definitions arranged in a 3x2 grid. Each definition includes a flow diagram, a title, a description, and a 'View Added Flow Definition' link. The definitions are:

- Kafka filter to Kafka** (Version 1): Consumes JSON, CSV or Avro events from Kafka, filters them before writing them back to Kafka as JSON, CSV or Avro.
- Kafka to Cloudera Operational Database** (Version 1): Consumes JSON, CSV or Avro events from Kafka and ingests them into Cloudera Operational Database (COD).
- Kafka to Kafka** (Version 1): Consumes events from Kafka and writes them to another Kafka topic.
- Kafka to Kudu** (Version 1): Consumes JSON, CSV or Avro events from Kafka and ingests them into Kudu.
- Kafka to S3 Avro** (Version 1): Consumes JSON, CSV or Avro events from Kafka and writes Avro files to S3.
- S3 to S3 Avro** (Version 1): Consumes JSON, CSV or Avro files from source S3 location and writes Avro files to a destination S3 location.

At the bottom of the sidebar, there is a 'Help' icon, a cookie notice for 'alpha_intcookieuser cookie', and the version number '1.0.1-b578'.

Flow Catalog

- Central repository for flow definitions
- Import existing NiFi flows
- Manage flow definitions
- Initiate flow deployments

Flow Catalog

Search by name

Import Flow Definition

REFRESHED 25 seconds ago

Name ↑	Type	Versions	Last Updated	
cc_fraud_template_int101run	Custom Flow Definition	2	a day ago	>
cc_fraud_template_int101run2	Custom Flow Definition	1	9 days ago	>
JSON_Kafka_To_Avro_S3	Custom Flow Definition	2	a day ago	>
Kafka filter to Kafka	ReadyFlow	1	2 days ago	>
Kafka to Cloudera Operational Database	ReadyFlow	1	2 days ago	>
Kafka to S3 Avro	ReadyFlow	1	14 hours ago	>
nifi_flows	Custom Flow Definition	1	2 months ago	>
Weather Data Flow	Custom Flow Definition	1	a day ago	>
Weather_Data	Custom Flow Definition	1	15 days ago	>
Weather_JSON_Kafka_To_Avro_S3	Custom Flow Definition	1	21 days ago	>

Items per page: 10 1 - 10 of 10 < < > >

1.0.1-b570

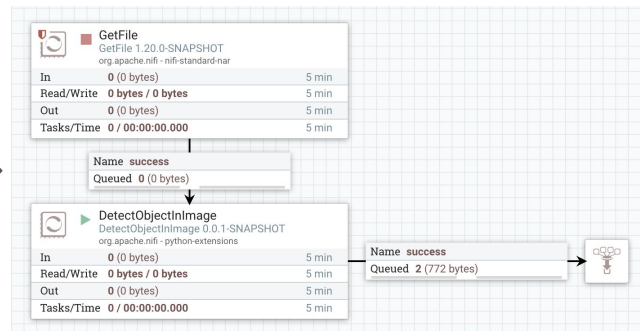
Apache NiFi with Python Custom Processors

Python as a 1st class citizen

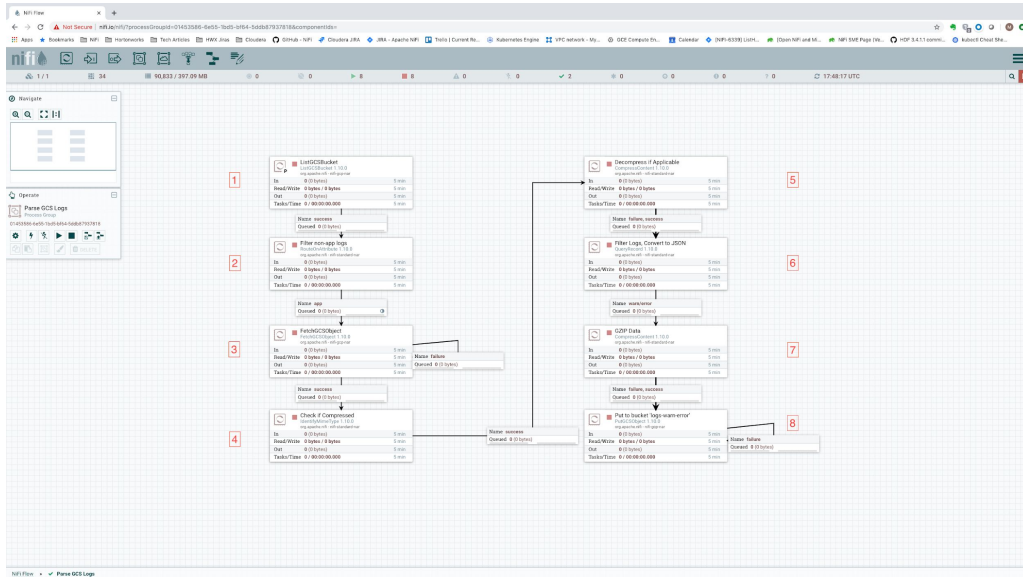
```
import cv2
import numpy as np
import json
from nifiapi.properties import PropertyDescriptor
from nifiapi.properties import ResourceDefinition
from nifiapi.flowfiletransform import FlowFileTransformResult

SCALE_FACTOR = 0.00392
NMS_THRESHOLD = 0.4 # non-maximum suppression threshold
CONFIDENCE_THRESHOLD = 0.5

class DetectObjectInImage:
    class Java:
        implements = ['org.apache.nifi.python.processor.FlowFileTransform']
    class ProcessorDetails:
        version = '0.0.1-SNAPSHOT'
        dependencies = ['numpy >= 1.23.5', 'opencv-python >= 4.6']
    def __init__(self, jvm=None, **kwargs):
        self.jvm = jvm
        # Build Property Descriptors
        self.model_file = PropertyDescriptor(
            name = 'Model File',
            description = 'The binary file containing the trained Deep Neural Network weights. Supports Caffe (*.caffemodel), TensorFlow (*.pb), Torch (*.t7, *.net), Darknet (*.weights), ' +
            'DLDT (*.bin), and ONNX (*.onnx)',
            required = True,
            resource_definition = ResourceDefinition(allow_file = True)
        )
        self.config_file = PropertyDescriptor(
            name = 'Network Config File',
            description = 'The text file containing the Network configuration. Supports Caffe (*.prototxt), TensorFlow (*.pbtxt), Darknet (*.cfg), and ONNX (*.xml)',
            required = False,
            resource_definition = ResourceDefinition(allow_file = True)
        )
        self.class_name_file = PropertyDescriptor(
            name = 'Class Names File',
            description = 'A text file containing the names of the classes that may be detected by the model. Expected format is one class name per line, new-line terminated.',
            required = True,
            resource_definition = ResourceDefinition(allow_file = True)
        )
        self.descriptors = [self.model_file, self.config_file, self.class_name_file]
    def getPropertyDescriptors(self):
        return self.descriptors
    def onScheduled(self, context):
        # read class names from text file
        class_name_file = context.getProperty(self.class_name_file.name).getValue()
        if class_name_file is None:
```



Processing one million events per second with NiFi



Nodes	Data rate/sec	Events/sec	Data rate/day	Events/day
1	192.5 MB	946,000	16.6 TB	81.7 Billion
5	881 MB	4.97 Million	76 TB	429.4 Billion
25	5.8 GB	26 Million	501 TB	2.25 Trillion
100	22 GB	90 Million	1.9 PB	7.8 Trillion
150	32.6 GB	141.3 Million	2.75 PB	12.2 Trillion

SOURCES AND SINKS



APACHE ICEBERG

A Flexible, Performant & Scalable Table Format

- Donated by **Netflix** to the Apache Foundation in 2018
- Flexibility
 - Hidden partitioning
 - Full schema evolution
- Data Warehouse Operations
 - Atomic Consistent Isolated Durable (ACID) Transactions
 - Time travel and rollback
- Supports best in class SQL performance
 - High performance at Petabyte scale





AMQP



AWS Lambda



Airtable



Amazon
API Gateway
Amazon API Gateway



Amazon CloudWatch



Amazon DynamoDB



Amazon Kinesis Data Firehose



Amazon Kinesis
Data Streams
Amazon Kinesis Data Streams



Amazon SQS



Amazon Simple Notification Services
(SNS)



Amazon S3

Amazon Simple Storage Service (S3)



Apache Accumulo



Apache Cassandra



Apache HBase



Apache Hive



Apache Iceberg



Apache Ignite



Apache Kafka



Apache Kudu



Apache Solr

FREE LEARNING ENVIRONMENT

CSP Community Edition



- Kafka, KConnect, SMM, SR, Flink, and SSB in Docker
- Runs in Docker
- Try new features quickly
- Develop applications locally
- Docker compose file of CSP to run from command line w/o any dependencies, including Flink, SQL Stream Builder, Kafka, Kafka Connect, Streams Messaging Manager and Schema Registry
 - `$>docker compose up`
- Licensed under the Cloudera Community License
- **Unsupported**
- Community Group Hub for CSP
- Find it on docs.cloudera.com under Applications



CSP Community Edition

A readily available, dockerized deployment of Apache Kafka and Apache Flink that allows you to test the features and capabilities of Cloudera Stream Processing.

[Learn More](#)

<https://www.cloudera.com/downloads/cdf/csp-community-edition.html>

Open Source Edition



- Apache NiFi in Docker
- Runs in Docker
- Try new features quickly
- Develop applications locally
- Docker NiFi
 - `docker run --name nifi -p 8443:8443 -d -e SINGLE_USER_CREDENTIALS_USERNAME=admin -e SINGLE_USER_CREDENTIALS_PASSWORD=ctsBtRBKHRax69EqUghvvgEvjnaLjFEB apache/nifi:latest`
 - Licensed under the ASF License
 - **Unsupported**

<https://hub.docker.com/r/apache/nifi>

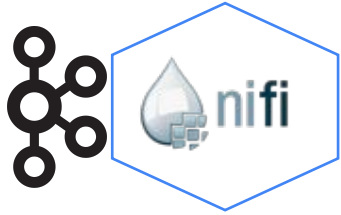
DEMO AND CODE

The background of the slide features three parallel diagonal stripes in shades of light gray, extending from the bottom-left towards the top-right.

Collect: Bring Together

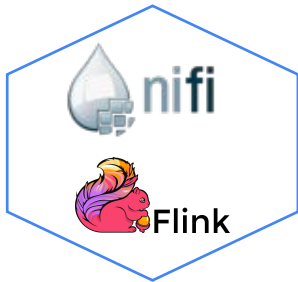


Aggregate all data from sensors, drones, logs, geo-location devices, images from cameras, results from running predictions on pre-trained models.



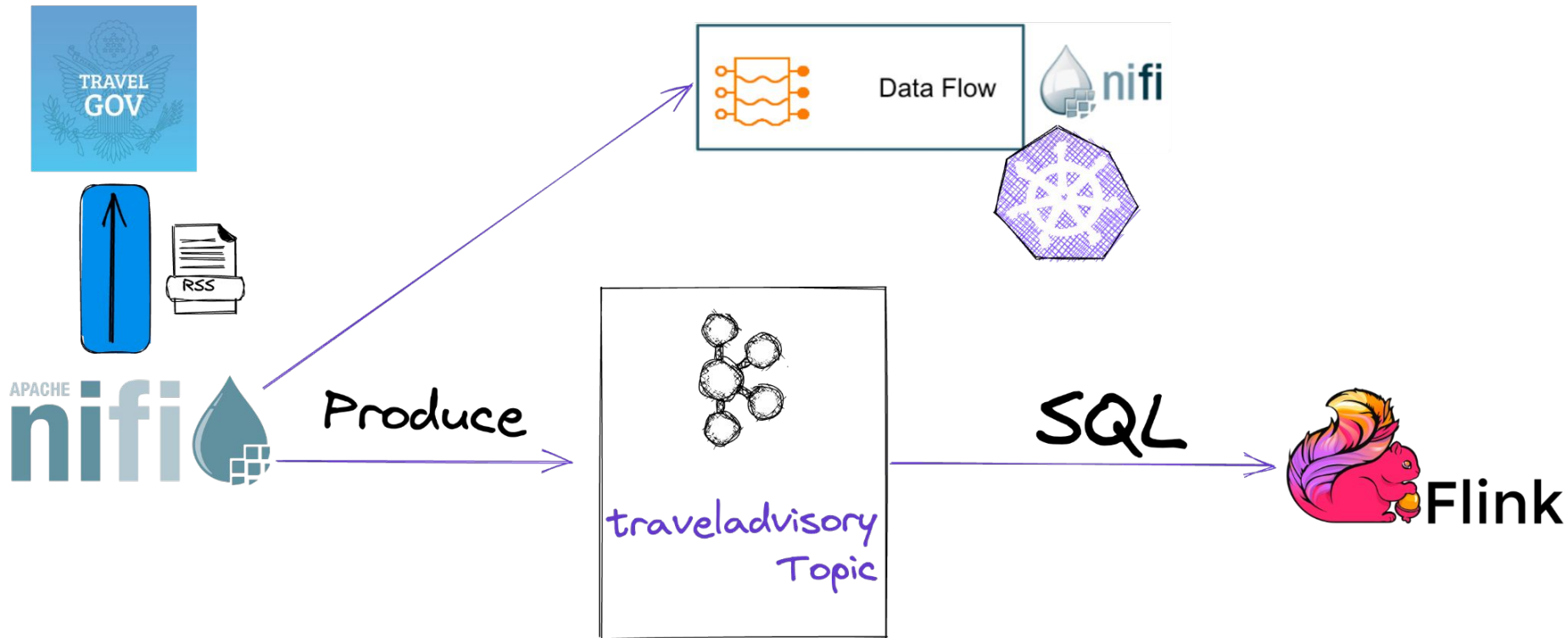
Conduct: Mediate the Data Flow

Mediate point-to-point and bi-directional data flows, distribute, delivering data reliably to Apache Iceberg, S3, Snowflake, Slack and Email.



Curate: Gain Insights

Orchestrate, parse, merge, aggregate, filter, join, transform, fork, query, sort, dissect, store, enrich with weather, location, sentiment analysis, image analysis, object detection, image recognition and more with Apache Tika, Apache OpenNLP and Machine Learning.



CLouDERA

METRICS ASSIGNMENT DATA EXPLORER CONFIGS LATENCY

ISOLATION LEVEL: read_uncommitted

DESERIALIZER: Keys: String Values: Avro

VALUE SCHEMA NAME:

traveladvisory

VALUE SCHEMA VERSIONS:

1

[Show schema text](#)

Offset	Timestamp	Key	Value
162	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Afghanistan - Level 4: Do Not Travel\", \"pubdate\": \"Thu, 20 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/traveladvisories/t show more
163	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Cura\u00e7ao - Level 1: Exercise Normal Precautions\", \"pubdate\": \"Tue, 04 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travelad show more
164	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Cura\u00e7ao - Level 1: Exercise Normal Precautions\", \"pubdate\": \"Tue, 04 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travelad show more
165	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Cura\u00e7ao - Level 1: Exercise Normal Precautions\", \"pubdate\": \"Tue, 04 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travelad show more
166	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Azerbaijan - Level 2: Exercise Increased Caution\", \"pubdate\": \"Tue, 15 Nov 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travel show more
167	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Azerbaijan - Level 2: Exercise Increased Caution\", \"pubdate\": \"Tue, 15 Nov 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travel show more
168	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Cameroon - Level 2: Exercise Increased Caution\", \"pubdate\": \"Wed, 12 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travelad show more
169	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Colombia - Level 3: Reconsider Travel\", \"pubdate\": \"Wed, 04 Jan 2023\", \"link\": \"http://travel.state.gov/content/travel/en/traveladvisories/ show more
170	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Colombia - Level 3: Reconsider Travel\", \"pubdate\": \"Wed, 04 Jan 2023\", \"link\": \"http://travel.state.gov/content/travel/en/traveladvisories/ show more
171	Fri, Mar 31 2023, 10:09:34	null	{\"title\": \"Georgia - Level 1: Exercise Normal Precautions\", \"pubdate\": \"Tue, 04 Oct 2022\", \"link\": \"http://travel.state.gov/content/travel/en/travelad show more



RUNNING



Flink Dashboard

<> Templates

Editor



Materialized View



Job Settings

Job Actions

```

1 select title, domain, category, link, pubdate, ts, uuid, advisoryId
2 FROM
3 `srl`.`default_database`.traveladvisory
4

```


<input type="checkbox"/>	title	domain	category	link	pubdate	ts	uuid
<input type="checkbox"/>	Bhutan - Level 1: Exercise Normal Precautions	BT,advisory	Level 1: Exercise Normal ...	http://travel.state.gov/co...	Wed, 05 Oct 2022	1680277517680	00152509-ed06-4000-93...
<input type="checkbox"/>	China - Level 3: Reconsider Travel	CH,advisory,MC,HK	CH	http://travel.state.gov/co...	Fri, 10 Mar 2023	1680277517682	79e7912a-5d40-4afb-96...
<input type="checkbox"/>	China - Level 3: Reconsider Travel	CH,advisory,MC,HK	HK	http://travel.state.gov/co...	Fri, 10 Mar 2023	1680277517682	528c584a-e2cc-4119-ac...
<input type="checkbox"/>	Tajikistan - Level 2: Exercise Increased Caution	TJ,advisory	Level 2: Exercise Increas...	http://travel.state.gov/co...	Wed, 05 Oct 2022	1680277517683	24fef95e-42a9-4011-9f3...
<input type="checkbox"/>	Zambia - Level 1: Exercise Normal Precautions	ZA,advisory	advisory	http://travel.state.gov/co...	Tue, 28 Mar 2023	1680277517684	a4e8106e-5f55-4ef9-a5e...
<input type="checkbox"/>	Taiwan - Level 1: Exercise Normal Precautions	TW,advisory	advisory	http://travel.state.gov/co...	Mon, 24 Oct 2022	1680277517688	ed3bad9e-96a0-42ca-a6...
<input type="checkbox"/>	Chad - Level 3: Reconsider Travel	CD,advisory	Level 3: Reconsider Travel	http://travel.state.gov/co...	Tue, 04 Oct 2022	1680277517690	1ac6673c-dd29-4186-b8...



Materialized View

Configuration

Primary Key ⓘ

uuid

Enable MV ⓘ

Retention (Seconds) ⓘ

Recreate on Job Start ⓘ

Ignore NULLs ⓘ

Min Row Retention Count ⓘ

10000

API Key ⓘ

traveladvisory1

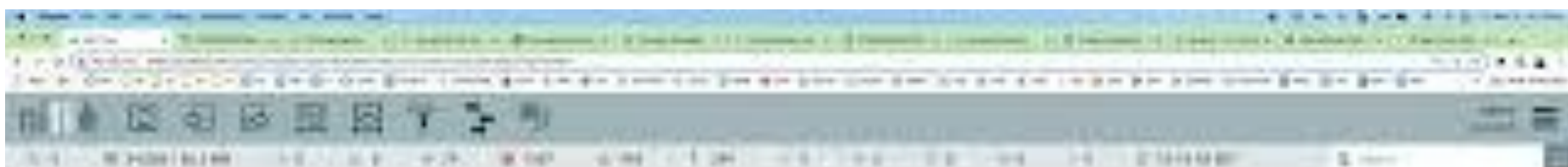


Queries

[+ Add New Query](#)

```
/api/v1/query/5201/travel?key=66ba91a9-507f-422c-bbb4-86250a9f7bb1&limit=100
```





Weather Data For USA

Weather Data	
1 2 3 4 5 6 7 8 9 10 11 12	
Temp:	80.0000
W:	80.0000
Humidity:	80.0000
W:	80.0000

Unit: Fahrenheit	0
Timezone:	0
Refresh Interval:	0

Weather History and Comments	
1 2 3 4 5 6 7 8 9 10 11 12	
Temp:	80.0000
W:	80.0000
Humidity:	80.0000
W:	80.0000

Live Transit Feeds

Transit Feeds	
1 2 3 4 5 6 7 8 9 10 11 12	
Temp:	80.0000
W:	80.0000
Humidity:	80.0000
W:	80.0000

Unit: Fahrenheit	0
Timezone:	0
Refresh Interval:	0

Transit History and Comments	
1 2 3 4 5 6 7 8 9 10 11 12	
Temp:	80.0000
W:	80.0000
Humidity:	80.0000
W:	80.0000

Coastal Data Feeds

Coastal Data	
1 2 3 4 5 6 7 8 9 10 11 12	
Temp:	80.0000
W:	80.0000
Humidity:	80.0000
W:	80.0000

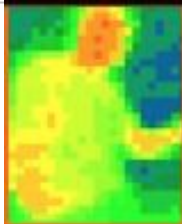
RESOURCES AND WRAP-UP

Streaming Resources

- <https://dzone.com/articles/real-time-stream-processing-with-hazelcast-and-streamnative>
- <https://flipstackweekly.com/>
- <https://www.datainmotion.dev/>
- <https://www.flankstack.dev/>
- <https://github.com/tspannhw>
- <https://medium.com/@tspann>
- <https://medium.com/@tspann/predictions-for-streaming-in-2023-ad4d7395d714>
- https://www.apachecon.com/acna2022/slides/04_Spann_Tim_Citizen_Streaming_Engineer.pdf

Resources



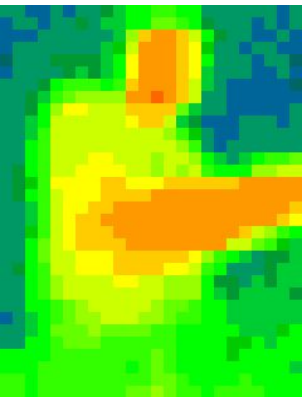


FLANK STACK



ALL YOUR DATAFLOWS ARE BELONG TO US

imgflip.com



THANK YOU 

