Dealer Event Catalog

An Event-Driven System Using Log-Based Architecture

[Prerequisites 2](#_Toc12453117)

[Installation 2](#_Toc12453118)

[Confluent Platform 3](#_Toc12453119)

[Elasticsearch 3](#_Toc12453120)

[Kibana 3](#_Toc12453121)

[NiFi 3](#_Toc12453122)

[Postgres 3](#_Toc12453123)

[Removal 3](#_Toc12453124)

[Basic Usage 4](#_Toc12453125)

[Kafka REST Proxy 4](#_Toc12453126)

[Producing JSON messages 4](#_Toc12453127)

[Consuming JSON Messages 4](#_Toc12453128)

[Inspecting Topics 5](#_Toc12453129)

[API Docs 6](#_Toc12453130)

[Additional Content Types 6](#_Toc12453131)

[Kafka 6](#_Toc12453132)

[Producing Messages 6](#_Toc12453133)

[Creating a Topic 6](#_Toc12453134)

[Producing Messages 6](#_Toc12453135)

[Consuming Messages 7](#_Toc12453136)

[Dealer Event Catalog 7](#_Toc12453137)

[Overview 7](#_Toc12453138)

[Sample Flow 8](#_Toc12453139)

[Deployment 9](#_Toc12453140)

[Elasticsearch 9](#_Toc12453141)

[Kibana 9](#_Toc12453142)

[NiFI 9](#_Toc12453143)

[Kafka Connect 10](#_Toc12453144)

[Environment Specific Changes 10](#_Toc12453145)

[Image Repos 10](#_Toc12453146)

[Security Context 10](#_Toc12453147)

[Chart Image Names 11](#_Toc12453148)

[Network Policy 11](#_Toc12453149)

[Images 11](#_Toc12453150)

[Control Center 11](#_Toc12453151)

[Kafka 11](#_Toc12453152)

[Kafka Connect 12](#_Toc12453153)

[Kafka REST 12](#_Toc12453154)

[KSQL Server 12](#_Toc12453155)

[Schema Registry 12](#_Toc12453156)

[Zookeeper 12](#_Toc12453157)

[Postgres 12](#_Toc12453158)

[NiFi 12](#_Toc12453159)

[Elasticsearch 13](#_Toc12453160)

[Kibana 13](#_Toc12453161)

[Future tasks 13](#_Toc12453162)

# Prerequisites

Kubernetes cluster

* Minikube for local development - [installation instructions](https://docs.confluent.io/current/installation/installing_cp/cp-helm-charts/docs/index.html#create-local-minikube)
* Or [ng-garage-nafta-qa.app.corpintra.net](http://ng-garage-nafta-qa.app.corpintra.net/) for non-local development

Helm installed on Kubernetes cluster - [installation instructions](https://helm.sh/docs/using_helm/#quickstart-guide).

**Important**

Helm must be initialized in a namespace where user has appropriate permissions, in this example namespace kafka is used:

helm init --tiller-namespace=kafka

# Installation

Execute the following commands to deploy Kafka, NiFi, Elasticsearch and related components.

**Important**

Docker images must be pulled from official repos, re-tagged and pushed to reg-dhc-americas.app.corpintra.net prior to being usable in Kubernetes deployments. See [Image Repos](#_Image_Repos) section for more details.

## Confluent Platform

helm install helm-charts/confluent \

--name confluent \

--tiller-namespace=kafka \

--namespace=kafka \

--set \

cp-kafka-rest.external.enabled=true,\

cp-kafka-rest.external.type=ClusterIP,\

cp-kafka-rest.external.externalTrafficPolicy=null,\

cp-zookeeper.persistence.enabled=false,\

cp-kafka.persistence.enabled=false,\

cp-kafka-connect.persistence.enabled=false,\

cp-ksql-server.persistence.enabled=false,\

cp-schema-registry.persistence.enabled=false

## Elasticsearch

helm install helm-charts/elasticsearch \

--name elasticsearch \

--tiller-namespace=kafka \

--namespace=kafka

## Kibana

helm install helm-charts/kibana \

--name kibana \

--tiller-namespace=kafka \

--namespace=kafka

## NiFi

helm install helm-charts/nifi \

--name nifi \

--tiller-namespace=kafka \

--namespace=kafka

## Postgres

helm install helm-charts/postgres \

--name postgres \

--tiller-namespace=kafka \

--namespace=kafka

# Removal

helm del --purge confluent --tiller-namespace=kafka

helm del --purge elasticsearch --tiller-namespace=kafka

helm del --purge kibana --tiller-namespace=kafka

helm del --purge nifi --tiller-namespace=kafka

helm del --purge postgres --tiller-namespace=kafka

# Basic Usage

Once Helm charts have been successfully installed, Kafka can be accessed by clients external and internal to the Kubernetes cluster.

## Kafka REST Proxy

Due to limitations of current ingress and DNS, Kafka is unable to accept native connections from clients external to the Kubernetes cluster. Kafka REST Proxy provides an HTTP-based interface for those clients, as well as clients without native Kafka support.

### Producing JSON messages

1. Produce a message using JSON with the value {"foo": "bar"} to jsontest topic

curl -X POST -L -k \

-H "Content-Type: application/vnd.kafka.json.v2+json" \

-H "Accept: application/vnd.kafka.v2+json" \

--data '{"records":[{"value":{"foo":"bar"}}]}' \

"https://ng-garage-nafta-qa.app.corpintra.net/topics/jsontest"

Expected response:

{"offsets":[{"partition":0,"offset":0,"error\_code":null,"error":null}],"key\_schema\_id":null,"value\_schema\_id":null}

### Consuming JSON Messages

1. Create a consumer for JSON data, starting at the beginning of the topic's log

curl -X POST \

-H "Content-Type: application/vnd.kafka.v2+json" \

--data '{"name": "my\_consumer\_instance", "format": "json", "auto.offset.reset": "earliest"}' \

https://ng-garage-nafta-qa.app.corpintra.net/consumers/my\_json\_consumer

Expected response:

{"instance\_id":"my\_consumer\_instance", "base\_uri":"http://localhost:8082/consumers/my\_json\_consumer/instances/my\_consumer\_instance"}

2. Subscribe to jsontest topic

curl -X POST \

-H "Content-Type: application/vnd.kafka.v2+json" \

--data '{"topics":["jsontest"]}' \

https://ng-garage-nafta-qa.app.corpintra.net/consumers/my\_json\_consumer/instances/my\_consumer\_instance/subscription

Expected response:

<empty>

3. Consume data using the base URL in the first response

curl -X GET \

-H "Accept: application/vnd.kafka.json.v2+json" \ https://ng-garage-nafta-qa.app.corpintra.net/consumers/my\_json\_consumer/instances/my\_consumer\_instance/records

Expected response:

[{"key":null,"value":{"foo":"bar"},"partition":0,"offset":0,"topic":"jsontest"}]

4. Close the consumer with a DELETE to make it leave the group and clean up its resources

curl -X DELETE \

-H "Content-Type: application/vnd.kafka.v2+json" \

https://ng-garage-nafta-qa.app.corpintra.net/consumers/my\_json\_consumer/instances/my\_consumer\_instance

Expected response:

<empty>

### Inspecting Topics

1. Get a list of topics

curl "https://ng-garage-nafta-qa.app.corpintra.net/topics"

Expected response:

["\_\_consumer\_offsets","\_schemas","avrotest","binarytest","jsontest"]

2. Get info about one topic

curl "https://ng-garage-nafta-qa.app.corpintra.net/topics/jsontest"

Expected response:

{"name":"avrotest","configs":{},"partitions":[{"partition":0,"leader":0,"replicas":[{"broker":0,"leader":true,"in\_sync":true}]}]}

3. Get info about a topic's partitions

curl "https://ng-garage-nafta-qa.app.corpintra.net/topics/jsontest/partitions"

Expected response:

[{"partition":0,"leader":0,"replicas":[{"broker":0,"leader":true,"in\_sync":true}]}]

### API Docs

Kafka REST Proxy exposes Kafka’s functionality via a REST API, for the full list of available commands visit <https://docs.confluent.io/current/kafka-rest/api.html>

### Additional Content Types

In addition to JSON messages, Kafka REST Proxy supports Avro and Binary formats. For more information visit:

Content Types Overview – <https://docs.confluent.io/current/kafka-rest/api.html#content-types>

Avro Messages - <https://docs.confluent.io/current/kafka-rest/quickstart.html#produce-and-consume-avro-messages>

Binary Messages - <https://docs.confluent.io/current/kafka-rest/quickstart.html#produce-and-consume-binary-messages>

## Kafka

### Producing Messages

To connect from a client pod:

1. Deploy a kafka client pod with configuration:

kubectl apply -f clients/kafka-client.yaml -n kafka

2. Log into the Pod

kubectl exec -it kafka-client -n kafka -- /bin/bash

### Creating a Topic

kafka-topics --zookeeper confluent-cp-zookeeper-headless:2181 --topic confluent-topic --create --partitions 10 --replication-factor 2 --if-not-exists

### Producing Messages

MESSAGE="`date -u`"

echo "$MESSAGE" | kafka-console-producer --broker-list confluent-cp-kafka-headless:9092 --topic confluent-topic

### Consuming Messages

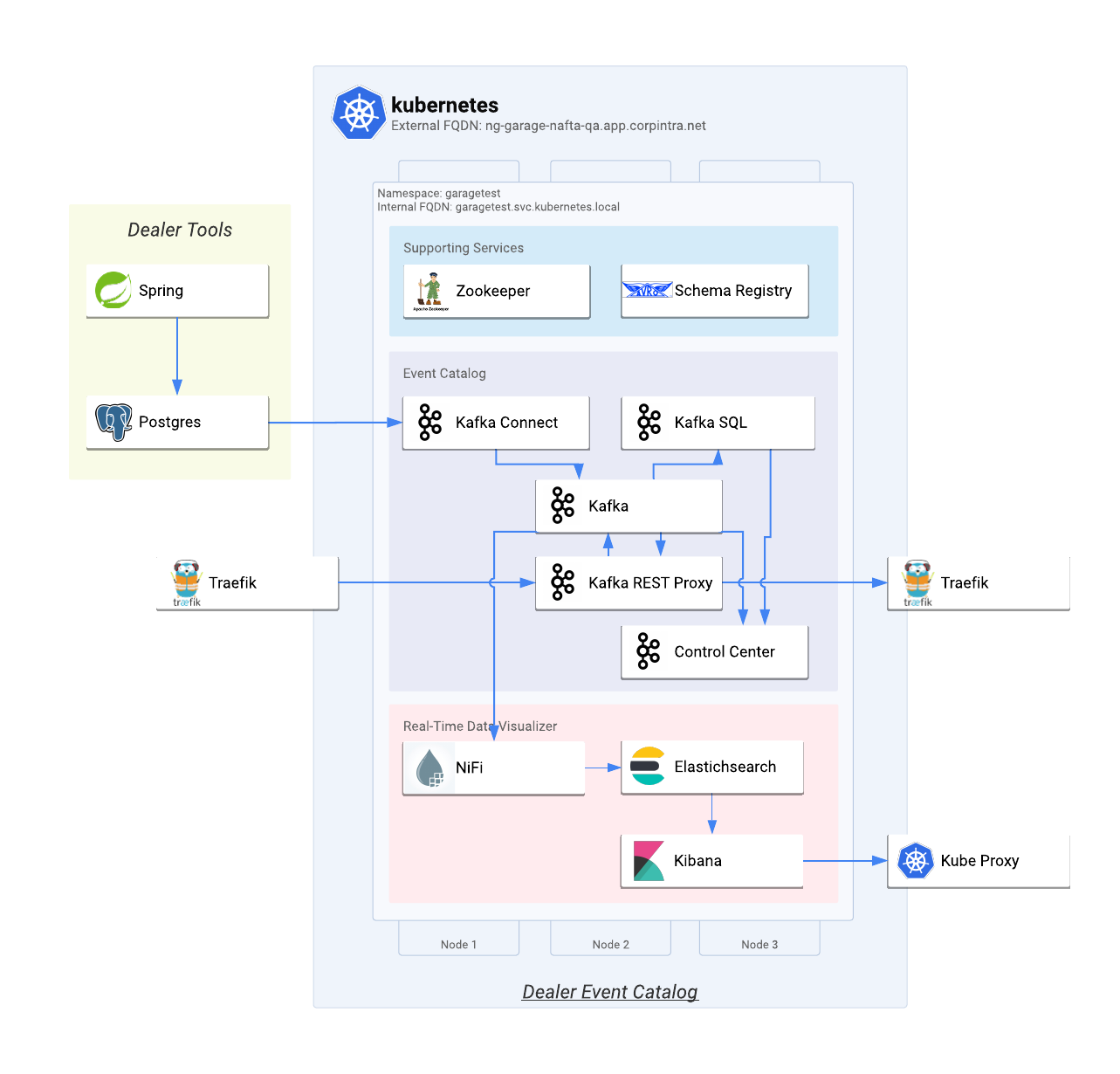
kafka-console-consumer --bootstrap-server confluent-cp-kafka-headless:9092 --topic confluent-topic --from-beginning --timeout-ms 2000 --max-messages 1 | grep "$MESSAGE"

# Dealer Event Catalog

## Overview

Dealer Event Catalog is a central event storage system designed to capture updates to dealer information originating from various existing systems. Raw event data is then converted into a standard format, enhanced and augmented to produce a canonical model representation of the event. Both raw and canonical model events are then made available to consumers interested in ongoing dealer updates in real time.

## Sample Flow



|  |  |
| --- | --- |
| **Dealer App**  Dealer management Spring web application |  |
| **Dealer Database**  Postgres database with Logical Decoding enabled |  |
| **Database Change Stream**  Postgres logical replication protocol used to stream changes | . |
| **Database Change Sink**  Kafka Connect with Postgres listening and relaying Postgres events to Kafka. |  |
| **Stream Processor** | Kafka SQL engine used to combine incoming dealer change events with current dealer records. |
| **Dealer Event Consumer** | Kafka REST Consumer receiving dealer changes coming from Dealer Tools. |

## Deployment

### Elasticsearch

Create a mapping for the timestamp field, letting Elasticsearch know it is of type date.

kubectl exec elasticsearch-0 -n kafka -- curl -XPUT "localhost:9200/dealer\_events?pretty" -H 'Content-Type: application/json' -d '{"mappings": {"message": {"properties": {"timestamp": { "type": "date" }}}}}'

### Kibana

Forward Kibana to local port 5601.

kubectl port-forward -n kafka kibana-0 5601:5601

### NiFI

Forward NiFi to local port 8080 and create a flow for storing Kafka messages in Elasticsearch

kubectl port-forward -n kafka nifi-0 8080:8080

response=$(curl -F template=@dealer-catalog/flow-dealer-updates.xml http://localhost:8080/nifi-api/process-groups/root/templates/upload -L -k)

template=$(echo ${response}|sed -n 's:.\*<id>\(.\*\)</id>.\*:\1:p')

data="{ \"originX\": 0.0, \"originY\": 0.0, \"templateId\": \"${template}\" }"

curl http://localhost:8080/nifi-api/process-groups/root/template-instance \

-H "Content-Type: application/json" \

-d "${data}"

### Kafka Connect

kubectl port-forward svc/confluent-cp-kafka-connect 8083:8083 -n kafka

curl http://localhost:8083/connectors \

-H "Content-Type: application/json" \

-d@dealer-catalog/connector-dealer-db.json

# Environment Specific Changes

This section outlines modifications made to official Docker images and Helm charts in order to leverage them in the ng-garage-nafta-qa environment. Modified source code for images and charts is included in this project.

## Image Repos

All images had to be pulled from official repos, re-tagged and pushed to reg-dhc-americas.app.corpintra.net prior to being usable in Kubernetes deployments. This example demonstrates the process for confluentinc/cp-zookeeper:5.2.1 image.

# pull the image

docker pull confluentinc/cp-zookeeper:5.2.1

# tag image for internal repo

docker tag confluentinc/cp-zookeeper:5.2.1 reg-dhc-americas.app.corpintra.net/myanama/cp-zookeeper:5.2.1

# push image to internal repo

docker push reg-dhc-americas.app.corpintra.net/myanama/cp-zookeeper:5.2.1

## Security Context

Majority of charts were modified to include a security context section in order to work around ng-garage’s Kubernetes cluster policy of not using root users. For example:

...

spec:

securityContext:

fsGroup: 1000

runAsUser: 1000

containers:

- name: prometheus-jmx-exporter

...

## Chart Image Names

All charts values.yaml files were updated to use images from reg-dhc-americas.app.corpintra.net repo and with original image tags. For example:

...

cp-zookeeper:

enabled: true

servers: 3

image: reg-dhc-americas.app.corpintra.net/myanama/cp-zookeeper

imageTag: 5.2.1

...

## Network Policy

By default, components are not able to communicate with other components due to ng-garage-nafta-qa‘s default network policy. With the exception of Postgres, NetworkPolicy‘s had to be created for each component, here’s a sample one for Kafka service.

apiVersion: networking.k8s.io/v1

kind: NetworkPolicy

metadata:

name: cp-kafka-network-policy

spec:

podSelector:

matchLabels:

app: cp-kafka

ingress:

- {}

## Images

### Control Center

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-control-center |
| Original image | confluentinc/cp-enterprise-control-center:5.2.0 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-enterprise-control-center:5.2.0 |

### Kafka

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-kafka |
| Original image | confluentinc/cp-enterprise-kafka:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-enterprise-kafka:5.2.1 |

### Kafka Connect

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-kafka-connect |
| Original image | confluentinc/cp-enterprise-kafka-connect:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-enterprise-kafka-connect:5.2.1 |
| Image changes | * Installed Debezium Postgres connector |

### Kafka REST

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-kafka- rest |
| Original image | confluentinc/cp-enterprise-kafka- rest:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-enterprise-kafka- rest:5.2.1 |

### KSQL Server

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-ksql-server |
| Original image | confluentinc/cp-ksql-server:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-ksql-server:5.2.1 |
| Image changes | * Created kafka group * Created kafka user * Granted ownership of KSQL Server directory to kafka user |

### Schema Registry

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-schema-registry |
| Original image | confluentinc/cp-schema-registry:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-schema-registry:5.2.1 |

### Zookeeper

|  |  |
| --- | --- |
| Chart | helm-charts/confluent/charts/cp-zookeeper |
| Original image | confluentinc/cp-zookeeper:5.2.1 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/cp-zookeeper:5.2.1 |

### Postgres

|  |  |
| --- | --- |
| Chart | helm-charts/postgres |
| Original image | debezium/postgres:10 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/postgres:10 |

### NiFi

|  |  |
| --- | --- |
| Chart | helm-charts/nifi |
| Original image | apache/nifi:1.8.0 |
| New image | reg-dhc-americas.app.corpintra.net/myanama/nifi:1.8.0 |

### Elasticsearch

|  |  |
| --- | --- |
| Chart | helm-charts/elasticsearch |
| Original image | docker.elastic.co/elasticsearch/elasticsearch-oss:6.6.1 |
| New image | myanama/elasticsearch-oss:6.6.1 |

### Kibana

|  |  |
| --- | --- |
| Chart | helm-charts/kibana |
| Original image | docker.elastic.co/kibana/kibana-oss:6.6.1 |
| New image | myanama/kibana-oss:6.6.1 |

# Future tasks

* Operations
  + Logging
  + Monitoring
  + Alerting
* Clustered/HA
  + NiFi
  + Elasticsearch
  + Kibana
* Persistence
  + All components
* Encryption
  + In-flight
  + At-rest
* Role based access control
  + Authentication
  + Authorization
* Ingress
  + TCP
  + Subdomain-based FQDNs
* Testing
  + Message size
  + Message volume
  + Producer count
  + Consumer count
  + Endurance
* Licenses
  + Confluent
  + Elastic