**-** **USE CASE -****REAL TIME DATA PROCESSING WITH KAFKA, NIFI, CLICKHOUSE AND SUPERSET**

Thực hiện : Nguyễn Văn Tuấn - TuanNV90

Quản lý : Ứng Việt Anh - AnhUV

Chủ đề : Use Cases for Apache Nifi

**THUẬT NGỮ VÀ TỪ VIẾT TẮT**

none

**USE CASE: REAL TIME DATA PROCESSING WITH KAFKA, NIFI, CLICKHOUSE AND SUPERSET**

1. **Mô tả usecase**

chúng ta sẽ xử lý dữ liệu xe ra/vào theo thời gian thực từ 5 bãi đỗ xe khác nhau và tạo một dòng dữ liệu và báo cáo đầu cuối

Thu thập dữ liệu từ api bằng nifi, đưa dữ liệu vào kafka topics, phân luồng dữ liệu để xử lí , lưu dữ liệu sau xử lí vào Clickhouse và trực quan hoá bằng Superset

1. **Quy trình**

* Using Nifi to retrieve messages from API and push the data to the Kafka Producer.
* Consuming Kafka messages in NiFi
* Storing data in ClickhouseDB
* Build a dashboard using Superset ( chưa thực hiện)

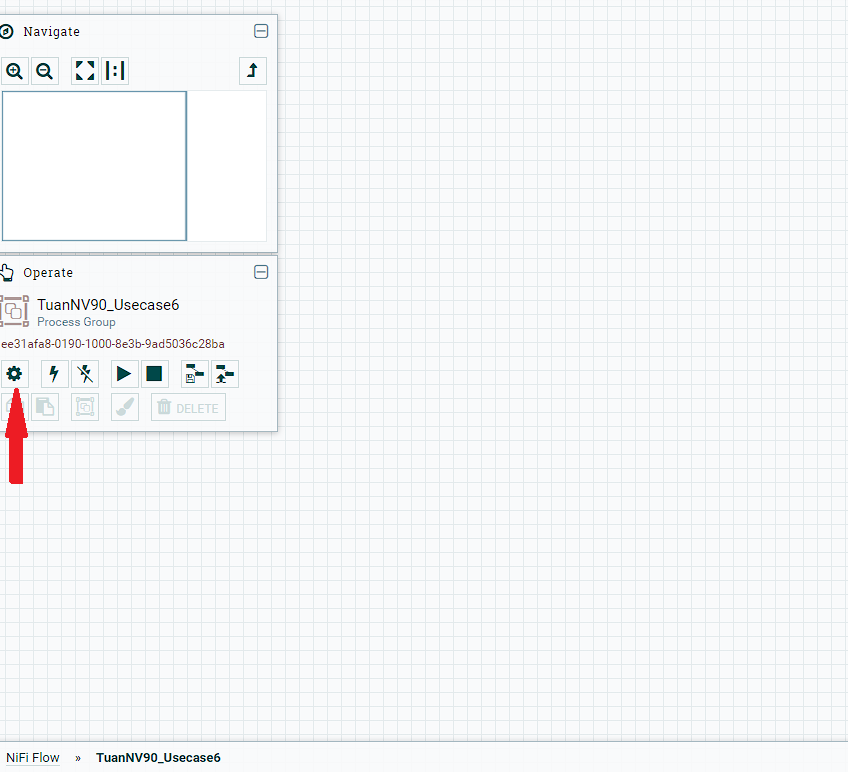
1. **Chuẩn bị**

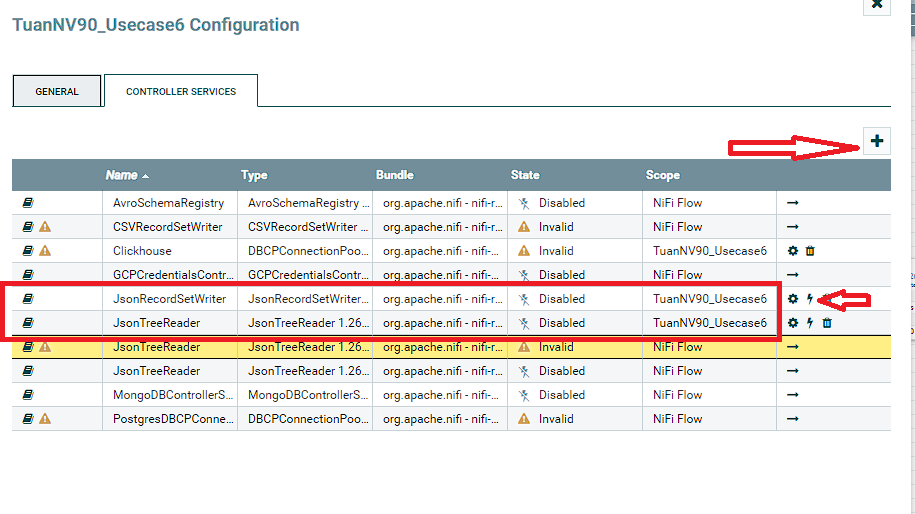
* Api: mỗi bãi đỗ xe sẽ có 1 api (giả lập API thay đổi theo thời gian -random)
* <https://tuannv90.pythonanywhere.com/park-1>
* <https://tuannv90.pythonanywhere.com/park-2>
* <https://tuannv90.pythonanywhere.com/park-3>
* <https://tuannv90.pythonanywhere.com/park-4>
* <https://tuannv90.pythonanywhere.com/park-5>
* Apache Nifi đã dựng sẵn
* Clickhouse
* Apache Superset đã dựng sẵn
* Kafka đã dựng sẵn

1. **Thiết lập và cài đặt:**

**Bước 1: Thiết lập, cấu hình các processor và triển khai luồng trên Nifi**

Bước 1.1:





Bước 1.2: Đầu tiên tạo một Process Group bằng việc kéo và thả vào giao diện, sau đó hãy đặt tên cho nó. Process Group này sẽ chứa các Processor và các kết nối giữa Processor của chúng ta, việc tạo Process Group giúp chúng ta quản lí dễ dàng hơn

A screenshot of a computer

Description automatically generated

Bước 1.3: Double-click vào Process Group vừa tạo, chúng ta sẽ kéo mục Processor để thêm các Processor và cấu hình nó như sau:

Tạo 5 processor “InvokeHTTP” và thiết lập “HTTP URL” bằng 5 api tương ứng

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Tiếp theo cấu hình các process như dưới:

ConvertRecord:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

PublishKafkaRecord:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

ConsumeKafka:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

jA screenshot of a computer

Description automatically generated

EvaluateJsonPath:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

JoltTransformJSON:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Mã Jolt Specifition:  
[{

"operation": "default",

"spec": {

"Category": "Category",

"City": "City",

"ID": "ID",

"Make": "Make",

"Model": "Model",

"Name": "Name",

"Plate": "Plate",

"TS": "TS",

"Year": "Year",

"Exit\_TS": "${now():format('yyyy-MM-dd HH:mm:ss')}"

}

}]

A screenshot of a computer

Description automatically generated

RouteOnAttribute:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

JotlTransformJSON:



A screenshot of a computer

Description automatically generated

[{

"operation": "default",

"spec": {

"Category": "Category",

"City": "City",

"ID": "ID",

"Make": "Make",

"Model": "Model",

"Name": "Name",

"Plate": "Plate",

"TS": "TS",

"Year": "Year",

"Exit\_TS": "${now():format('yyyy-MM-dd HH:mm:ss')}",

"Price":150

}

}]

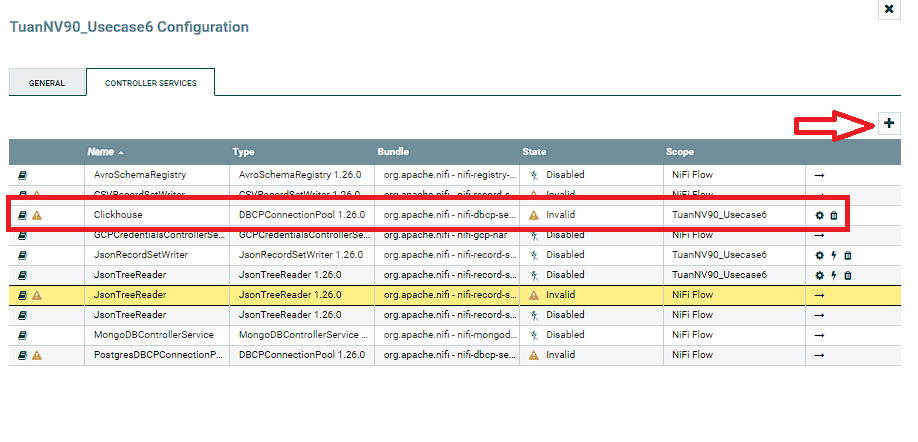
A screenshot of a computer

Description automatically generated

PutDatabaseRecord

A screenshot of a computer

Description automatically generated



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Luồng chúng ta sẽ triển khai:  
A screenshot of a computer

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated