

Installation einer Testumgebung für Apache - Kafka und Apache NiFi

Motivation:

- Bau einer „Spielwiese“ für Kafka – Cluster und NiFi Testumgebung auf Windows PCs
- Verwendung der Open Source Distributionen der benötigten Produkte
- PoC für Windows Entwicklungsumgebung
 - Kafka – Cluster / 3 Broker & Apache NiFi Singlenode

Verfasser: Martin.Mayer-Brooks@telekom.de

Stand: 2022-01-12

zu installierende Komponenten:

- ➔ Apache Zookeeper
- ➔ Apache Kafka
- ➔ Apache NIFI

Hinweise:

Alle drei Komponenten sind als JAVA_Programme implementiert und somit plattformunabhängig

Und können für Tests relativ einfach installiert werden, wenn eine kompatible Java Laufzeitumgebung installiert ist.

- a) Z.B. Unter Windows
- b) Unter verschiedenen Linux Distributionen oder
- c) Z.B. als Docker Container z.B. in K8S – Cluster z.B. zu finden unter:
 - a. https://hub.docker.com/_/zookeeper
 - b. <https://hub.docker.com/r/wurstmeister/kafka>
 - c. <https://hub.docker.com/r/apache/nifi>

Aber ganz ohne Anpassungen geht es u.U. auf einfachen Windows-PCs nicht.

Der Mitschrieb zeigt solchen Themen und Lösungsansätze auf.

Rahmenbedingungen / Stolperfallen

Installierte JRE macht Probleme (Vermutung c:\program files(86)... -> Sonderzeichen in Skripten ?)

Abhilfe: eigene Java-Laufzeitumgebung installieren 😊

Generelles Installationsprinzip

Unter „C:\Users\<myUSER>\myApp“ wird eine „Heimat“ für eigene versch. portable Programme eingeführt und darunter wird eine freie Java – Distribution installiert.

⇒ Hier verwendet: „openjdk-x64-13.0.1“ (<https://openjdk.java.net/>)

Benutzer - Umgebungsvariable JAVA_HOME = C:\Users\<myUSER>\myApps\openjdk-x64-13.0.1 setzen

- ➔ Damit wird aber die Systemumgebung von der IT-Abteilung überschrieben! 😊
- ➔ Ggf. besser: JAVA_HOME nur in neu bereit zu stellender Umgebung konfigurieren und Standardeinstellungen des PC belassen.
 - Aber u.U. aufwendig 😞
 - Anleitung geht von gesetzter Umgebungsvariable aus, wenn nichts anderes explizit gesagt wird.


URLs : Software Quellen und Tutorial

- Apache Zookeeper
 - Quelle: <https://zookeeper.apache.org/releases.html>
 - Tutorial: https://www.tutorialspoint.com/apache_kafka/index.htm
- Apache Kafka
 - Quelle: <http://kafka.apache.org/downloads>
 - Tutorial: https://www.tutorialspoint.com/apache_kafka/index.html
 - <https://kafka.apache.org/quickstart>
 - Beispiele mit Linux Shell -> aber unter .\bin\windows\sind die korrespondierenden Batchfiles
- Apache NIFI
 - Quelle: <https://nifi.apache.org/download.html>
 - Tutorial: https://www.tutorialspoint.com/apache_nifi/index.htm
 - Online Doku = <https://nifi.apache.org/docs/nifi-docs/html/user-guide.html>

PoC-Mitschriebe – benötigte Java - Programme unter Windows installiert.

Wurzelverzeichnis der Kafka – Cluster – Software = T:\myApps

Hintergrund: Dateinamen / Pfade unter C:\Users\<myUSER>\myApps/* werden u.U. zu lang werden

1. Installation - Apache Zookeeper (1 Zookeeper; keine Cluster in Testumgebung)
 - a. Installationsordner = T:\myApps\zkeeper\zookeeper-current0
 - b. CONFIG = .\conf\zoo.cfg → change dataDir=./data (relative zu Installationsordner)
 - i. 1 Zookeeper auf default Port = 2181; Admin-Port 8080
 - ii. Log4j-properties
 - iii. **zkEnv.cmd -> log4j-Properties und JAVA ENV hardcoted anpassen**
 1. siehe set JAVA_HOME=C:\Users\<myUSER>\myApps\openjdk-x64-13.0.1
- c.  zkEnv.cmd
- d. Start mit: (Verknüpfung)
 - i. T:\myApps\zkeeper\zookeeper-current0\bin\zkServer.cmd
 - ii. Ausführen in : T:\myApps\zkeeper\zookeeper-current\bin

iii. Beispielskript mit Echo-Ausgaben



zkEnv.cmd

2. Installation Apache Kafka

- a. Installationsordner = T:\myApps\kafka\kafka-current
- b. CONFIG -> server0.properties , server1.properties, server1.properties
 - i. Unterscheid in Broker-ID , und Portkonfiguration => 9090 – 9092



server0.properties

ii. Muster siehe

c. Start der Kafkabroker0..2 mit „Startverknüpfungen“

- i. kafka-server0-start
 1. Ziel: (in einer Zeile: Skript und Konfigfile des Brokers)
 - a. T:\myApps\kafka\kafka-current\bin\windows\kafka-server-start.bat
T:\myApps\kafka\kafka-current\config\server0.properties
 2. Ausführen in: T:\myApps\kafka\kafka-current

d. Hinweis:

- i. Bei Installation unter C:\Users\<myUSER>\myApp\ kafka\kafka-current
Brachte Fehlermeldung in Startskripte „Paramterliste zu lang... ?

Abhilfe: Installation unter T\myApps -> keine Probleme 😊 Ursache ?

Test vgl. <https://kafka.apache.org/quickstart>

3. Apache Nifi (7 GB !)

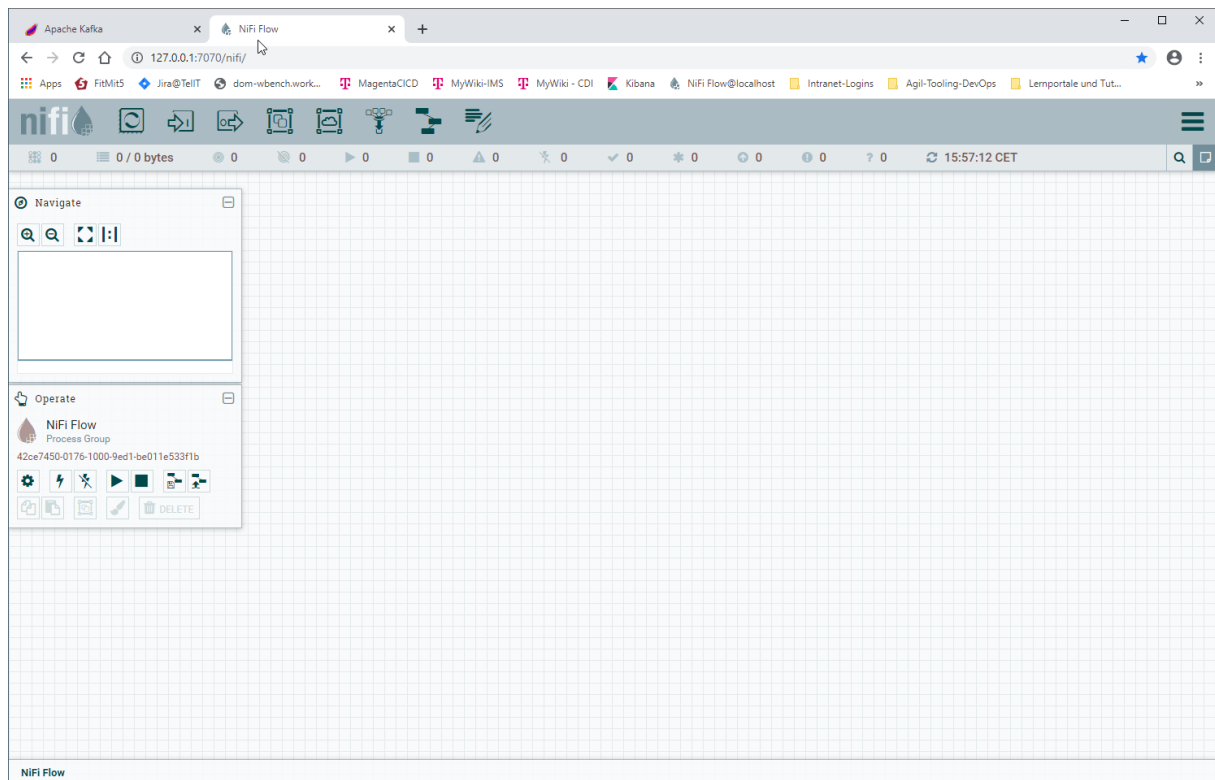
- a. Installationsort: C:\Users\<myUSER>\myApps\nifi\nifi-current
- b. Configuration .\conf\ nifi.properties
 - i. nifi.web.http.port = 7070 zu ändern → Port 8080 ist auch Zookeeper-Adminport

Hinweis:

Beim Starten von NiFi sieht man den vollständigen Start des Applikationsservers (Jetty) nicht in der STDOUT, sondern im Applikationslogfile

(hier: C:\Users\<myUSER>\myApps\nifi\nifi-current\logs\nifi-app.log)

Die NiFi Webkonsole sollte dann so aussehen:



NiFi stellt sog. Prozessoren für unterschiedlichste Anwendungsfälle zur Verfügung.

Für Tests mit unserem installierten Kafka - Cluster sind die Typen

„PublishKafka_<version>“ (Producer) und „ConsumeKafka__<version>“ (Consumer) interessant

Einstiegstip: <https://nifi.apache.org/docs/nifi-docs/html/getting-started.html>

Viel Spaß und Erfolg beim Bau und Verwendung der Umgebung

VG Martin