tcVISION for Apache Kafka

A great part of the added value of modern IT systems is the latency-free data and process integration of transactional and analytical areas. The cross-system integration platform tcVISION is unique, efficient and reliable. With tcVISION mainframe data can be fast and easily integrated into BigData based operative applications or Business Intelligence and Analytics in near real-time.

The tcVISION solution is practice approved and is constantly developed further to meet the requirements of the new technologies. The result is the support of BigData in tcVISION version 6.

In the current version of tcVISION V6 BigData is now a fully integrated output platform and supports the integration with Apache Kafka. Consequently tcVISION supports direct streaming of changed data into a BigData Apache Kafka environment. Apache Kafka is an open source data streaming platform developed by the Apache Software Foundation.

The software stands out, because it is a distributed system and real-time scalable. Thus it is best suited to meet the challenges of BigData requirements.

As with all output platforms provided by tcVISION, the data streaming via Apache Kafka is based on official standard interfaces. The implementation of the Apache Kafka interface is fast and easy.

Data streaming to Apache Kafka as a transport layer enhances the BigData connectivity of tcVISION.

In addition to Apache Kafka, transport layers to BigData include the creation of files, direct output to a hadoop file system (HDFS) as well as the output to MongoDB.

The currently used protocols for data transfer to BigData are JSON, Avro and CSV. The main focus of the tcVISION integration platform is the supply of a real-time synchronization to integrate mainframe data into BigData based solutions.

Features

The tcVISION integration platform consists of a variety of state-of-the-art technology components which cover far more than an ETL process.

- With tcVISION data exchange in the sense of real-time synchronization and replication turns into a single-step operation.
- No additional middleware is required.
- Diverse Change Data Capture technologies allow an efficient selection of the required data from the source system with focus on the changed data. The data exchange process is reduced to the necessary minimum which results in lower costs for the data exchange.
- tcVISION can also use backup and recovery files (e.g. image copies, log files, UNLOADs, etc.) as a source for replication. Production data does not need to be touched.
- tcVISION enables the fast and efficient load of large volumes of mainframe data into BigData (streaming). The processor costs of the mainframe are low and negligible.
- An integrated data repository guarantees transparent data management across platforms.
- Mainframe knowledge is not necessarily required for the replication.
- tcVISION includes a rule engine to transform data into a target compliant format or allows user specific processing via supplied APIs.

- The integrated staging concept supports the offload of changed data in raw format to less expensive processor systems. This reduces costs and mainframe processor resources to a minimum.
- The preparation of the data for the target system can be performed on a less expensive platform (Linux, UNIX or MS-Windows).
- The transfer to and streaming of data into BigData is part of the tcVISION data exchange process. No intermediate files are required.
- The exchange of large volumes of data between a production mainframe environment and BigData can run in parallel processes to reduce latency to a minimum.
- The tcVISION integration platform contains comprehensive control mechanisms and monitoring functions for an automated data exchange.

- tcVISION has been designed in a way that BigData based projects can be deployed with complete project autonomy and maximum reduction of mainframe resources.

The Benefits

- Near real-time replication of mainframe data to BigData allows actual real-time analytics.
- The relocation of mainframe applications (e.g. internet applications like online banking, e-Government, etc.) to BigData with synchronous data on both platforms is also possible.
- Because of the concentration on changed data, the costs of the data exchange are reduced to a minimum.
- The utilization of mainframe resources is reduced as far as possible to avoid costs for mainframe knowledge and mainframe MIPS.
- Data exchange processes can be deployed and maintained with tcVISION without mainframe knowledge, thus costs can be saved and BigData projects can be developed and put to production faster.
- The near real-time replication of tcVISION from mainframe to BigData allows the relocation of BI, reporting and analytic applications to the more cost-effective and – for these applications – more powerful BigData platform.
- Compensation of decreasing mainframe knowledge