Replikacja bazy na SQL Server

Replikacja bazy danych polega na kopiowaniu i przesyłaniu danych lub obiektów bazodanowych pomiędzy wieloma serwerami oraz synchronizowaniu danych znajdujących się na biorących udział w replikacji serwerach bazodanowych. Dane kopiowane nazywamy danymi źródłowymi, a dane docelowe – repliką

# Gdzie wykorzystuje się replikację?

Replikację danych najczęściej wykorzystuje się w systemach rozproszonych baz danych, gdzie z jednego zdalnego więzła kopiuje się dane do innych zdalnych węzłów.

Replikując dane:

* Umożliwiamy lokalny dostęp do danych użytkownikom z oddalonych oddziałów firmy
* Pozwalamy na częściową niezależność serwerów bazodanowych, które mogą cały dzień przechowywać kopie danych, a jedynie nocą, podczas niskiego obciążenia systemu, zsynchronizować informacje
* Możemy podzielić dane w sposób odpowiadający strukturze firmy
* Możemy fizycznie rozdzielić serwery bazodanowe realizujące różne zadania w oparciu o te same dane

# Model replikacji

Na Microsoft SQL Server stosowany jest model wydawca – dystrybutor – subskrybent. Opisuje to trzy role które mogą zostać przypisane serwerom bazodanowym

## Wydawca

Wydawca (zwany także publikatorem) to serwer, na którym znajduje się źródłowa baza danych, udostępniona do publikacji. Zadaniami serwera pełniącego rolę wydawcy są:

* Utrzymywanie wzorcowej wersji danych
* Udostępnianie danych źródłowych innym serwerom
* Monitorowanie zmian w replikowanych danych i informowanie o nich serwer pełniący role dystrybutora

## Dystrybutor

Dane udostępnione prze wydawcę zostają przesłane na zarządzający replikacją serwer pełniący rolę dystrybutora. Dystrybutor zarządza przepływem danych między wydawcą a subskrybentami. Zadaniami tego serwera są:

* Utrzymywanie systemowej bazy danych dystrybucyjnej (distribution), która jest tworzona automatycznie podczas przypisywania serwerowi roli dystrybutora
* Przechowywanie i zarządzanie metadanymi związanymi z replikacją danych. Takimi jak historia zmian danych, transakcje przeprowadzone na poszczególnych serwerach, konfiguracja serwerów biorących udział w replikacji itd.

## Subskrybent

Subskrybent to serwer, który pobiera dane od publikatora poprzez dystrybutor i przechowuje je w swojej bazie danych, udostępniając je użytkownikom. Podstawowe zadania subskrybenta to:

* Odbieranie zmian replikowanych danych przeprowadzonych na innych serwerach
* Przechowywanie lokalnej kopii tych danych. Dane mogą zostać udostępnione użytkownikom tylko do odczytu lub do odczytu i modyfikacji

# Publikacje i artykuły

Dane replikowane pomiędzy serwerami bazodanowym składają się z tzw. Publikacji. Publikacją jest dowolny zbiór artykułów pobranych z bazy danych. Artykuł to pojedyncza tabela, widok lub procedura składowana przeznaczona do replikacji.

## Rodzaje replikacji

Na SQL Server mamy trzy rodzaje replikacji:

* Replikację migawkową (snapshot replication) – okresowo jest tworzony obraz (migawka) zawierajązy dane wszystkich artykułów w publikacji, i wysyłany jest w całości subskrybentów, nadpisując ich dane. Wadą tego rozwiązania jest wymóg dużej przepustowości, jak i tego, że migawka jest tworzona w całości od początku
* Replikację transakcyjną (transactional replication) – polega na bieżącym śledzeniu i przechwytywaniu przez wydawcę wszelkich zmian typu INSERT, UPDATE, DELETE i zachowywaniu ich w dystrybucyjnej bazie danych.
* Replikację scalającą (merge replication) – dane zmodyfikowane na jednych serwerze bazodanowym po określonym czasie są przesyłane do dystrybutora,który rozwiązuje konflikty i wysyła spójny obraz danych do wszystkich serwerów bazodanowych