

Rozproszone bazy danych

Wprowadzenie do systemów baz
danych

Fragmentacja i replikacja danych

- **Fragment** danych stanowi pewien podzbiór wszystkich danych całej bazy danych
- **Replika** danych stanowi kopię całości lub części danych przechowywanych w innej części całej bazy danych

Cechy rozproszonych baz danych

- Przezroczystość lokalizacji – użytkownicy nie muszą wiedzieć, w którym miejscu są przechowywane dane
- Przezroczystość fragmentacji – użytkownicy nie muszą wiedzieć, w jaki sposób dane są podzielone
- Przezroczystość replikacji – użytkownicy nie muszą wiedzieć, w jaki sposób dane są powtarzane (replikowane)
- Autonomia lokalna
- Nie opieranie się na lokalizacji centralnej

Typy rozproszonych baz danych

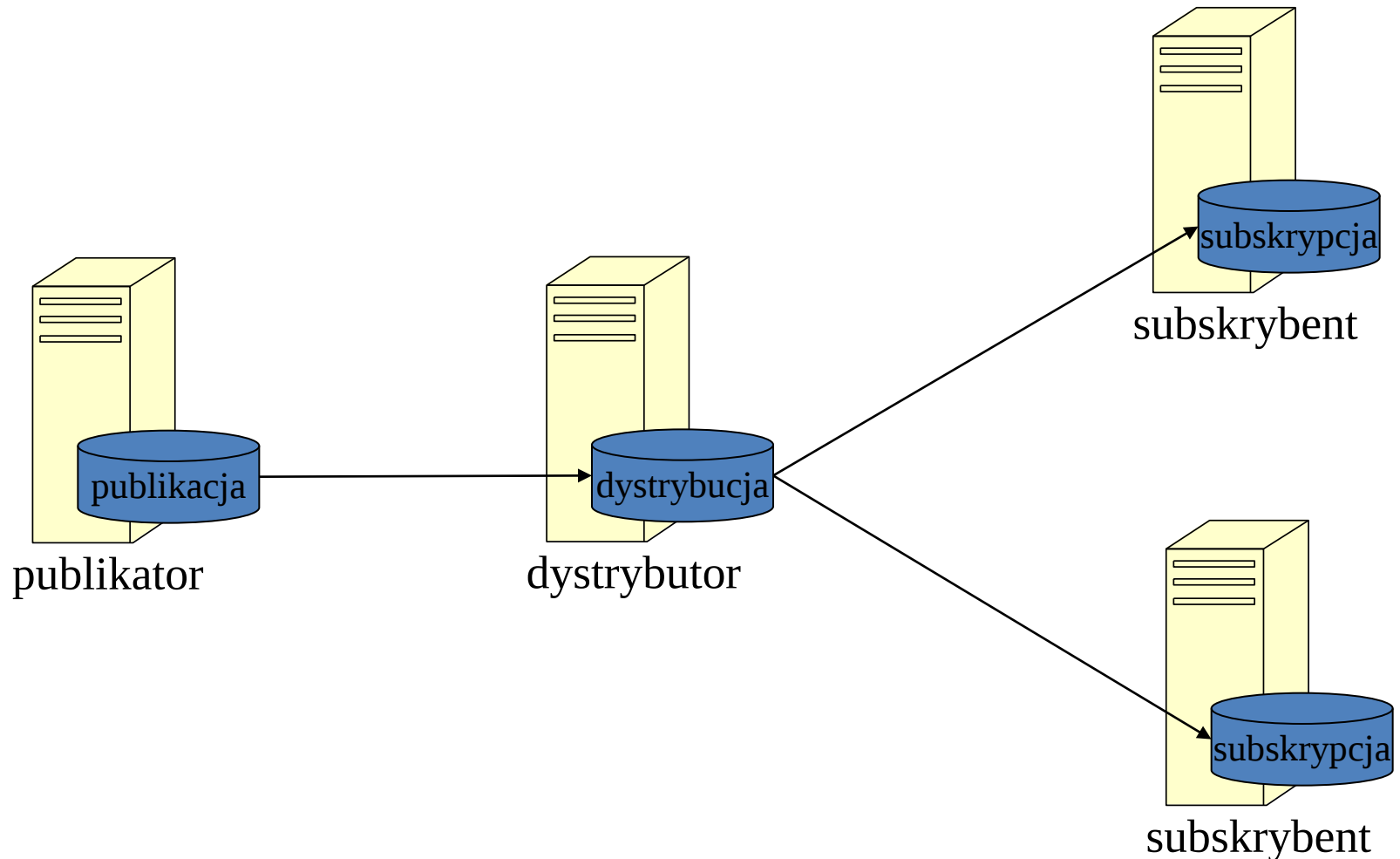
- Jednorodna rozproszona baza danych
 - Oparta na tym samym, dla wszystkich serwerów, systemie zarządzania bazą danych, np. MS SQL Server albo ORACLE
- Niejednorodna rozproszona baza danych
 - Oparta na różnych systemach jednocześnie, np. MS SQL Server i ORACLE
- Federacyjny (wielobazowy) system bazy danych
 - Połączenie względnie niezależnych autonomicznych baz danych

Role serwerów w środowisku replikacji danych

(MS SQL Server)

- **Publikator** (*serwer publikacji*) – serwer, na którym znajduje się źródłowa baza danych, która jest udostępniona do publikacji.
- **Dystrybutor** (*serwer dystrybucyjny*) – serwer, na którym znajduje się dystrybucyjna baza danych. Dystrybucyjna baza danych jest stosowana do składowania replikowanych danych. Zadaniem dystrybutora jest pobieranie, utrzymywanie, a następnie dystrybucja danych.
- **Subskrybent** (*serwer subskrypcyjny*) – serwer odbierający replikowane dane i zachowujący je w replikowanej bazie danych.

Publikator, dystrybutor i subskrybent

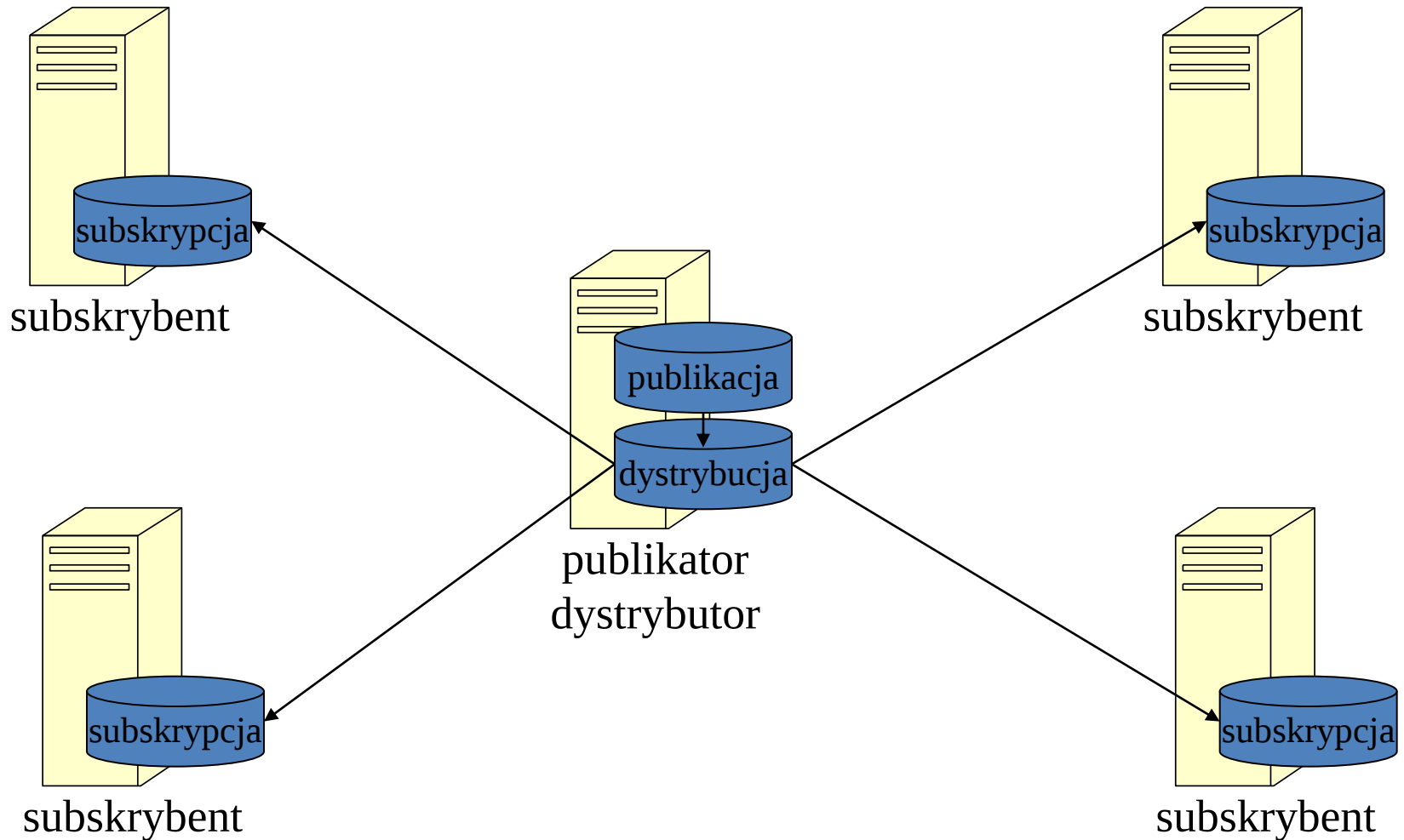


(Dystrybucja = dystrybucyjna baza danych)

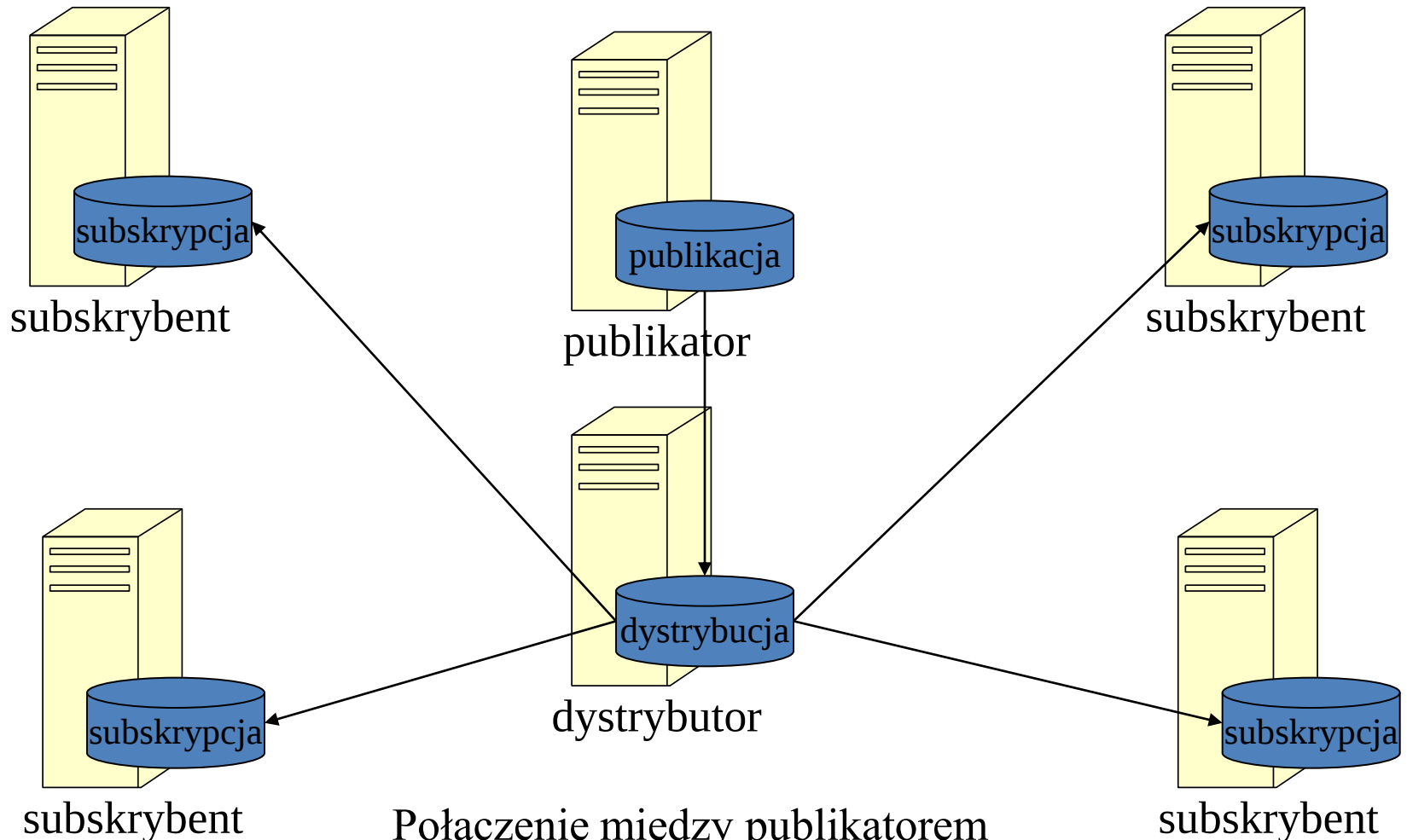
Scenariusze replikacji

- Centralny publikator
- Centralny publikator ze zdalnym dystrybutorem
- Publikujący subskrybent
- Centralny subskrybent
- Wiele publikatorów lub subskrybentów
- Subskrybent wprowadzający zmiany

Centralny publikator

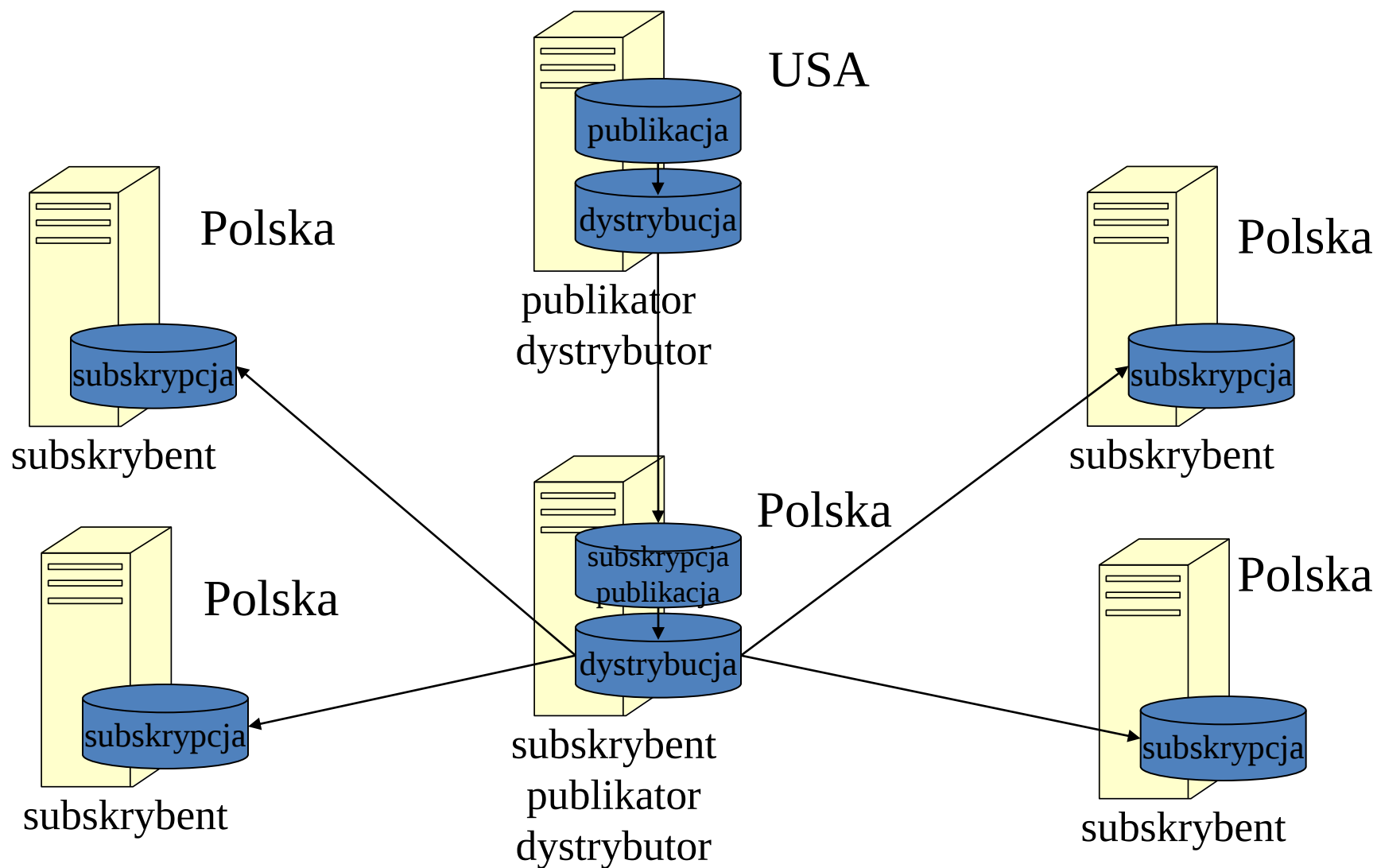


Centralny publikator ze zdalnym dystrybutorem

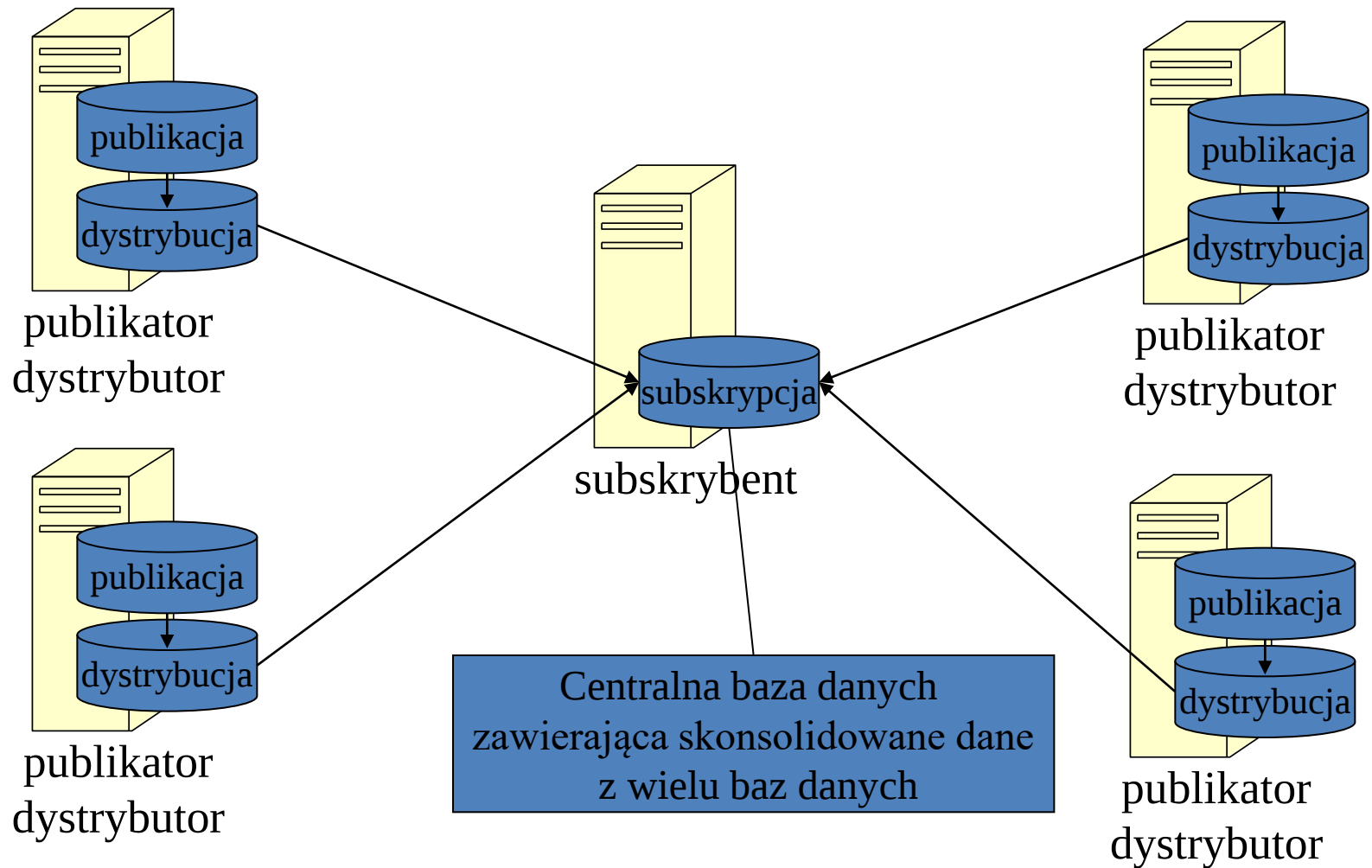


Połączenie między publikatorem
i dystrybutorem musi być szybkie i niezawodne

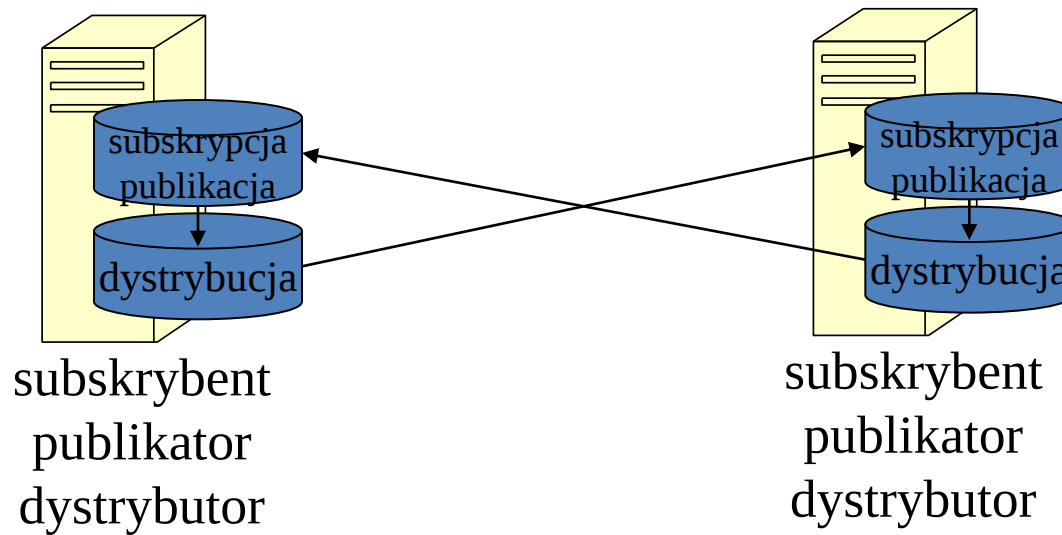
Publikujący subskrybent



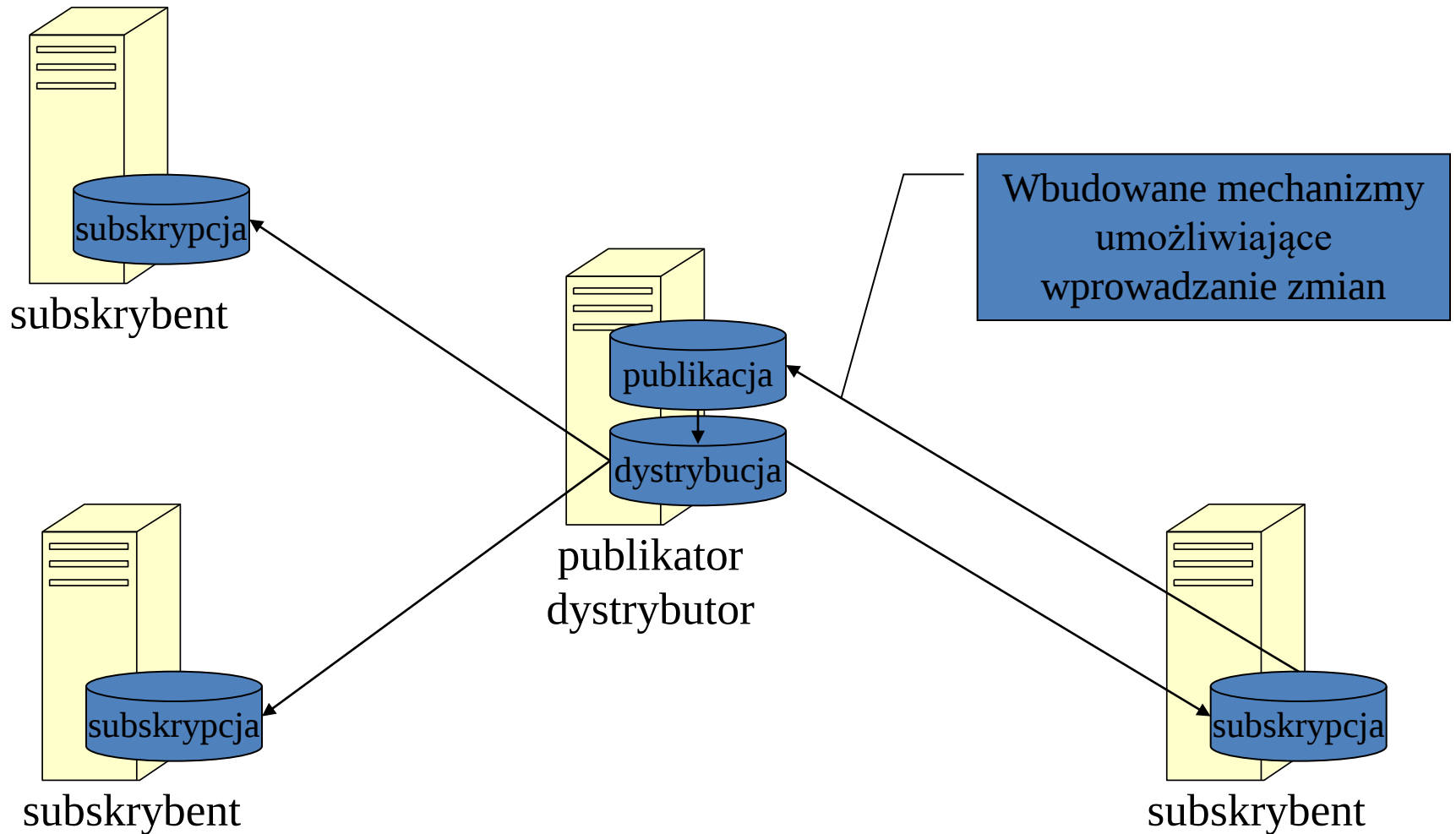
Centralny subskrybent



Wiele publikatorów wiele subskrybentów



Modyfikujący subskrybent



Typy replikacji

(MS SQL Server)

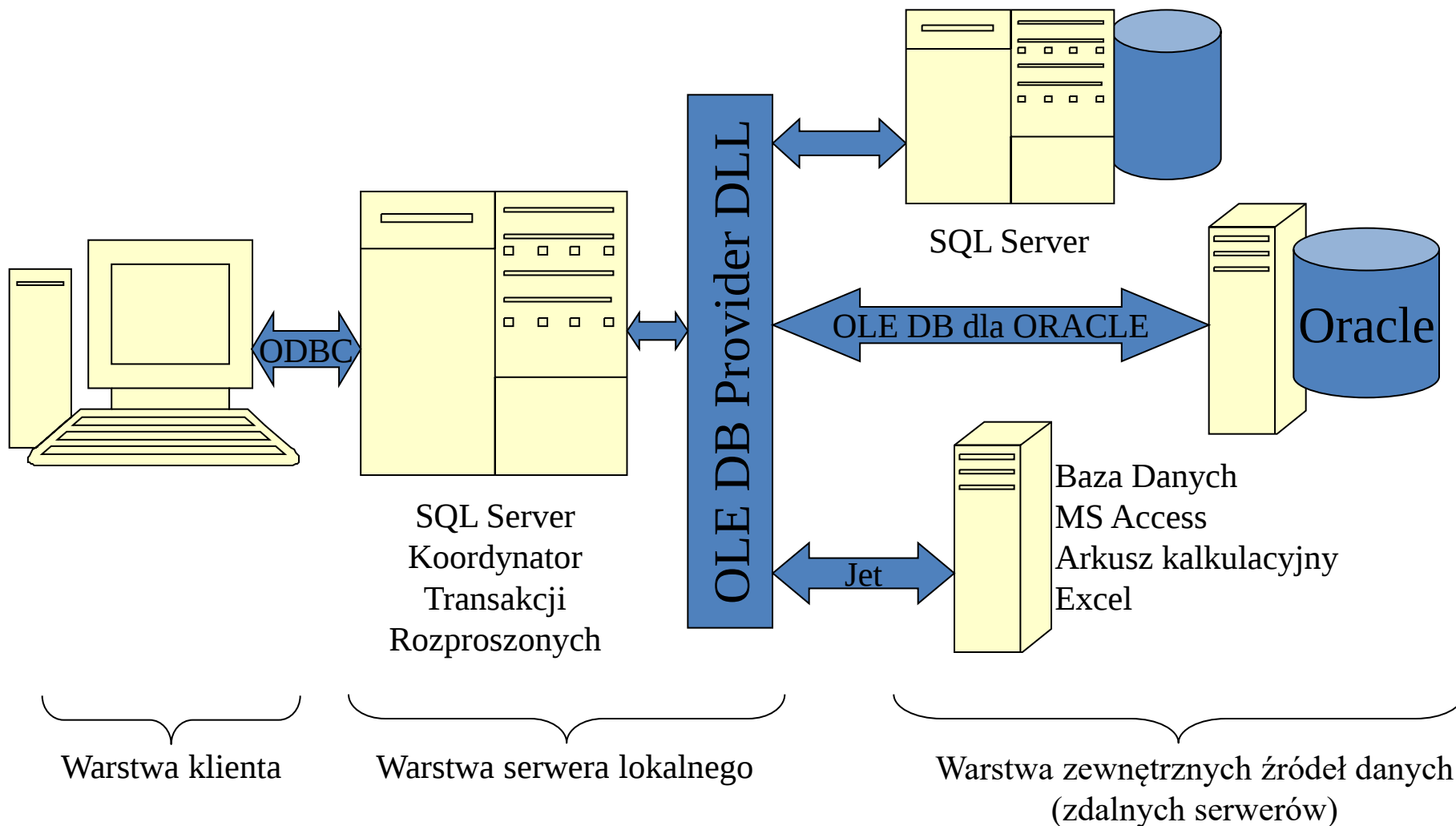
- **Replikacja migawkowa** (*snapshot replication*) – polega na pobraniu obrazu (sporządzeniu migawki) bazy danych i rozprowadzenia jej wśród subskrybentów.
- **Replikacja transakcyjna** (*transactional replication*) – śledzi i rozprowadza transakcje wśród subskrybentów; utrzymuje publikatora i subskrybenta w prawie takim samym stanie.
- **Replikacja scalająca** (*merge replication*) – okresowo przesyła zmiany. Modyfikacji bazy danych może dokonywać publikator i subskrybent – możliwa jest synchronizacja wielu systemów.

Fragmentacja danych

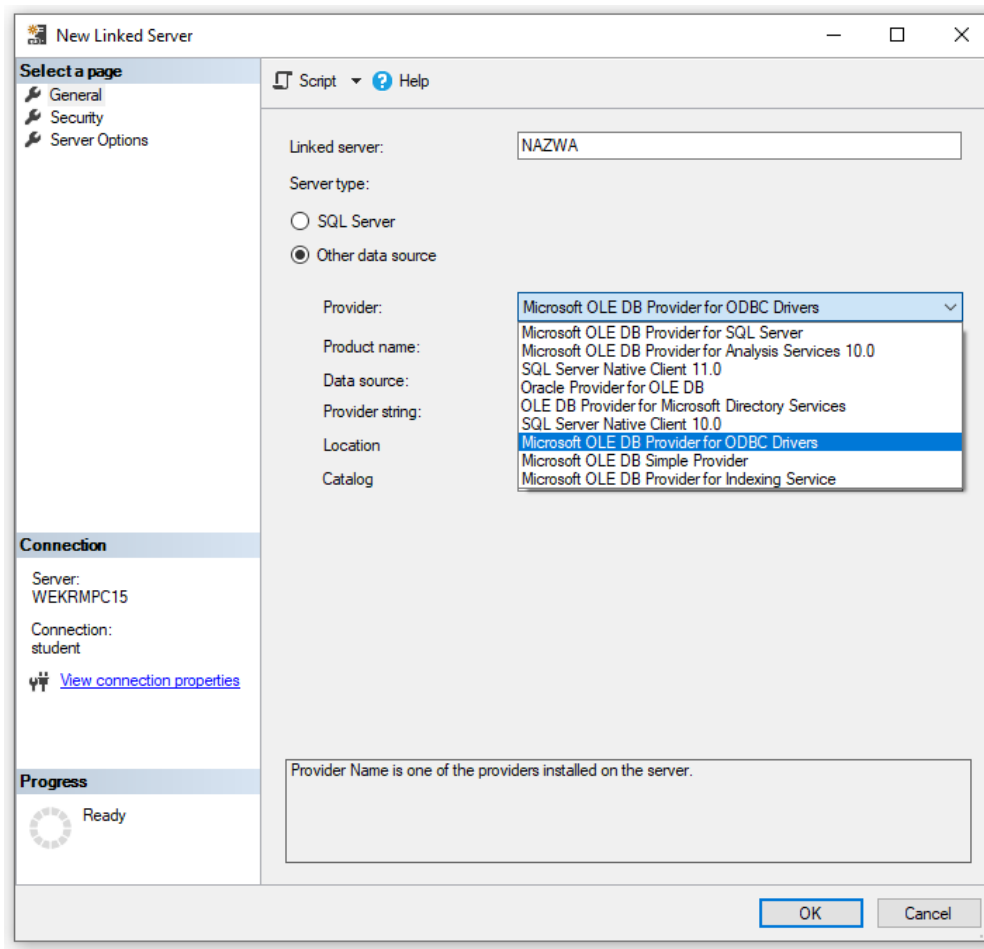
- Fragmentacja jest podziałem danych pomiędzy różne serwery
- Klient łączy się z jednym wybranym serwerem, a ten ściąga potrzebne dane z właściwych serwerów zewnętrznych i przesyła do klienta
- Informacja o lokalizacji danych przechowywana jest na serwerze
- Klient nie musi wiedzieć w jaki sposób dane zostały podzielone – wszystkie dane otrzymuje za pośrednictwem jednego serwera (przezroczystość fragmentacji)
- MS SQL Server wykorzystuje dwa mechanizmy łączenia z innymi serwerami
 - Serwery przyłączone (sprzężone) – nowy sposób umożliwiający wykonywanie zdalnych procedur i bezpośredni dostęp do zdalnych tabel, wymaga konfiguracji tylko serwera lokalnego

Architektura serwerów przyłączonych

<http://lukan.sytes.net/www/WebApplicationRembud/Strony/WebFormSQL>



Konfiguracja serwera przyłączonego na MS SQL Server



Transakcje rozproszone

- Rozproszone bazy danych umożliwiają wykonywanie zapytań rozproszonych (SELECT, INSERT, UPDATE i DELETE) tak, jakby dane zgromadzone były na jednym serwerze – jedyną różnicę stanowić może konwencja nazywania tabel
- Komercyjne serwery baz danych obsługują transakcje rozproszone, czyli transakcje obejmujące modyfikacje danych ulokowanych na różnych serwerach
- MS SQL Server wyposażony jest w usługę MS DTS (*Distributed Transaction Coordinator*) do kontroli transakcji rozproszonych i gwarantuje spójność wszystkich transakcji na SQL Serverze i serwerach z nim sprzężonych
- Warunkiem poprawnego wykonania transakcji rozproszonej jest jej zatwierdzenie na wszystkich serwerach