

Wrocławska Wyższa Szkoła Informatyki Stosowanej

Bazy danych

Dr hab. inż. Krzysztof Pieczarka

Email: kpieczarka@horyzont.eu
krzysztof.pieczarka@gmail.com

Literatura:

- Connolly T., Begg C., Systemy baz danych Praktyczne metody projektowania, implementacji i zarządzania, Tom 1, RM i Instytut Informatyki Uniwersytetu Śląskiego, Warszawa, 2004.
- Roszkowski J., Analiza i projektowanie strukturalne, Helion, Gliwice, 2004.
- Któraś z książek do MySQL-a najlepiej zawierająca omówienie podstaw języka SQL.
- P. Beynon-Davies: Systemy baz danych, WNT 2003
- J. Ullman: Systemy baz danych, WNT, 1989
- H. Garcia-Molina, J. D. Ullman, J. Widom: Systemy baz danych. Kompletny podręcznik. Wydanie II, Helion
- R. Elmasri, S. B. Navathe: Wprowadzenie do systemów baz danych. Wydanie VII, Helion

Zaliczenie:

Zadania kontrolne po poszczególnych zajęciach.

Projekt.

Zagadnienia realizowane w ramach przedmiotu:

- Przystwojenie wiedzy z zakresu podstaw modelowania danych i relacyjnego.
- Modelu danych, oraz struktury i implementacji bazy danych.
- Zapoznanie się z problematyka podstaw projektowania baz danych.
- Wykształcenie umiejętności definiowania schematu bazy i manipulacji danymi w języku SQL. Opanowanie języka zapytań SQL.
- Przystwojenie podstawowej wiedzy o projektowaniu systemów bazodanowych.

Definicje bazy danych

Bazą danych (Data Base, Database - DB) nazywamy zbiór danych o określonej strukturze, zapisany na wewnętrznym nośniku pamięciowym komputera, mogący zaspokoić potrzeby wielu użytkowników korzystających z niego w sposób selektywny w dogodnym dla siebie czasie.

Baza danych to zbiór informacji zapisanych w ściśle określony sposób w strukturach odpowiadających założonemu modelowi danych. W potocznym ujęciu obejmuje dane oraz program komputerowy wyspecjalizowany do gromadzenia i przetwarzania tych danych.

Pojęcie bazy danych

- Baza danych: strukturalizacja zarządzania informacją.
- Trwałość danych: zbiór danych istniejący przez długi czas.
- Zgodność z rzeczywistością: dotyczy pewnego fragmentu rzeczywistości związanego z firmą, organizacją, itp. Jest z reguły częścią systemu informacyjnego obsługującego zapotrzebowania informacyjne związane z tym fragmentem rzeczywistości.

System bazodanowy (System bazy danych - SBD, Data Base System - DBS) to:

- baza danych,
- środki programowe umożliwiające współbieżne operowanie na niej.

Podstawowe rodzaje funkcjonalności SBD to:

- współbieżne gromadzenie,
- wyszukiwanie,
- aktualizowanie informacji zawartych w bazie.

Systemem zarządzania bazą danych (ang. DataBase Management System, DBMS) nazywany jest program komputerowy wyspecjalizowany do gromadzenia i przetwarzania tych danych.

W ścisłej nomenklaturze baza danych oznacza zbiór danych, który zarządzany jest przez system DBMS.

Inne określenie: **System bazodanowy** (System bazy danych - SBD, Data Base System - DBS) to: baza danych oraz środki programowe umożliwiające współbieżne operowanie na niej.

Zastosowanie systemów bazodanowych

Bardzo różnorodne!!! Niemal wszędzie, gdzie mamy do czynienia z dużą ilością informacji zastosowanie znajdują bazy danych.

Przykładowe zastosowania bazy danych:

- system zarządzania kontaktami z klientami
- rejestracja różnego rodzaju usług
- kalendarz realizacji i rozliczanie usług (np. przypominanie o cyklicznych przeglądach)
- najróżniejsze kartoteki (klientów, pacjentów, studentów)
- sklepy internetowe (e-commerce)
- różnego rodzaju archiwa

Właściwości bazy danych

Współdzielenie danych

- współbieżny dostęp do danych przez wielu użytkowników

Kontrolowanie replikacji danych

- baza jest zbiorem danych bez niepotrzebnie powtarzających się lub zbędnych informacji

Spójność reprezentacji danych

- reprezentacja w oparciu o jeden model danych

Abstrakcja danych

- baza nie jest dokładnym modelem rzeczywistości, a jedynie jej wycinkiem

Bezpieczeństwo danych

Niezależność danych

- celem jest sytuacja, w której organizacja danych jest niewidoczna dla użytkowników i programów korzystających z danych

Model danych

Model danych jako architektura - zbiór zasad posługiwania się danymi:

- Definicja danych. Zbiór reguł określających, jaka jest struktura danych.
- Operowanie danymi. Zbiór reguł określających, jak operuje się danymi.
- Integralność danych. Zbiór reguł określających, które stany bazy danych są poprawne.

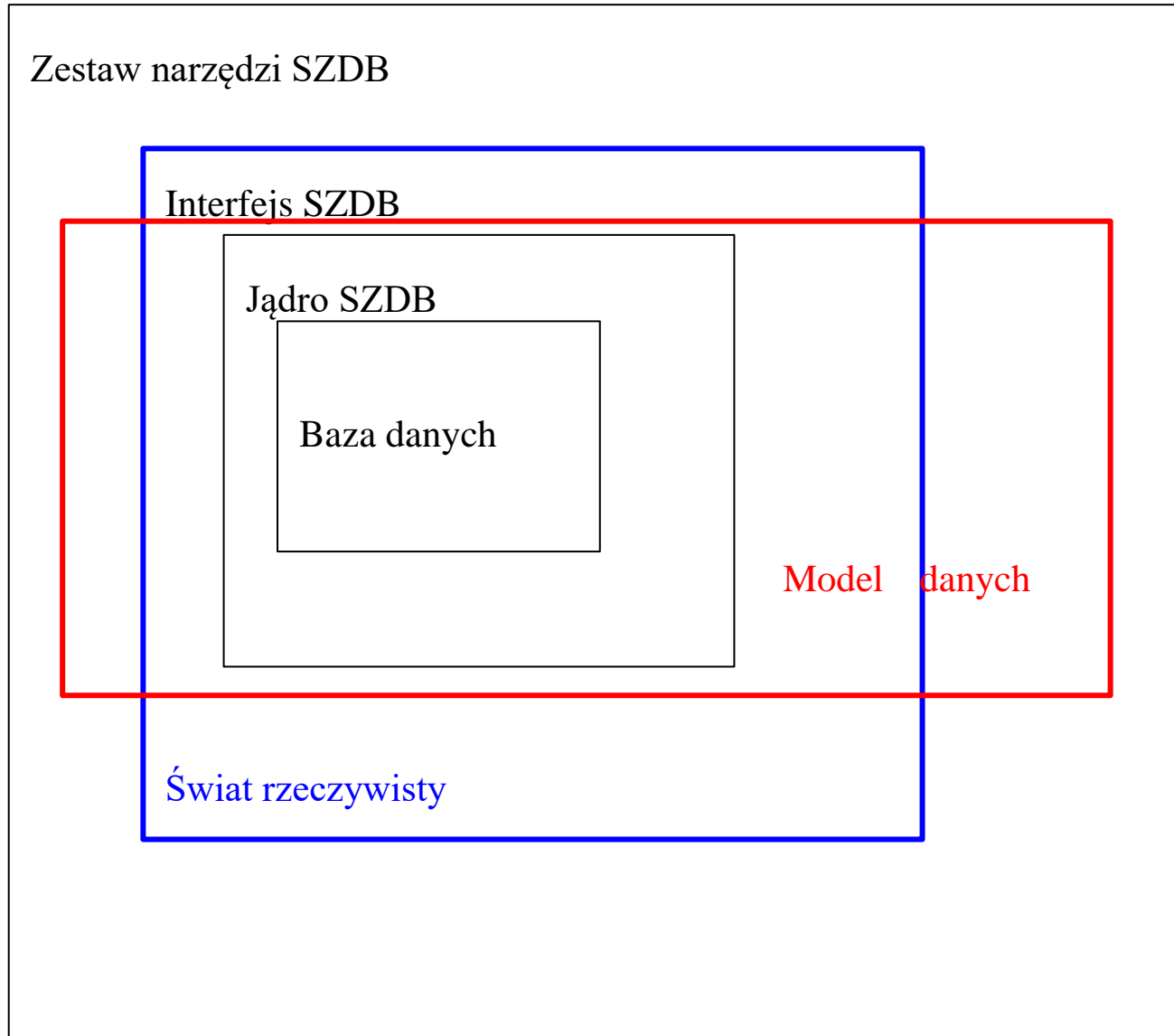
System zarządzania baza danych (SZBD) Database management system (DBMS)

Zorganizowany zbiór narzędzi umożliwiający dostęp i zarządzanie jedną lub wieloma bazami danych. SZBD jest powłoka, która otacza bazę danych i za pomocą której dokonują się wszystkie operacje na bazie danych.

Funkcje SZBD:

- Umożliwienie utworzenia nowej bazy i określenie jej schematu
- Przechowywanie danych
- Obsługa zapytań (ang. query)
- Zapewnienie wielodostępności
- Ochrona i zapewnienie integralności danych

Konstruowanie bazy danych



Moduł zarządzania transakcjami

Transakcja - nieformalna grupa operacji jednostkowych przeznaczonych do wykonania razem w jednym ciągu, jako duża operacja.

Właściwości (ACID):

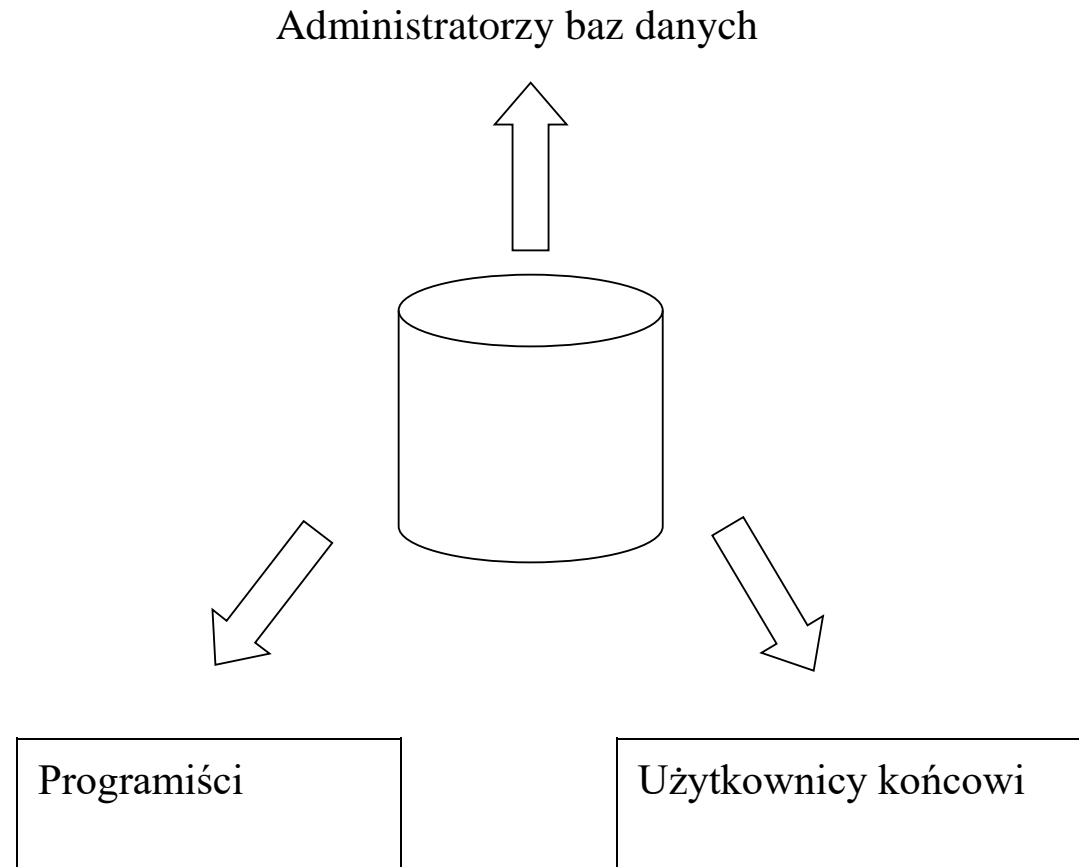
- niepodzielność (Atomicity) - cała transakcja powinna zostać przeprowadzona, albo żaden z jej elementów nie zostanie wykonany.
- spójność (Consistency) - np. miejsce w danym rejsie lotniczym nie może być przydzielone dwóm różnym pasażerom.
- izolacja (Isolation) - brak wpływu transakcji na siebie przy jednoczesnym ich przetwarzaniu.
- trwałość (Durability) - po zakończeniu transakcji jej wynik nie może zostać utracony.

Transakcje cd.

Podstawowe rozwiązania techniczne zapewniające właściwości ACID:

- Blokady - blokowanie elementu, którego dotyczy wykonywana właśnie transakcja. Po założeniu blokady dane są niedostępne dla innych transakcji.
- Logi - dokumentowanie operacji tzn. rozpoczęcie każdej transakcji, zmiany dokonywane w bazie danych przez transakcje oraz zakończenie transakcji Log jest przechowywany w pamięci stałej.
- Zatwierdzanie transakcji - gdy transakcja kończy działanie, jest gotowa do zatwierdzenia zmiany kopiowane są do logu, dopiero potem następuje aktualizacja danych.

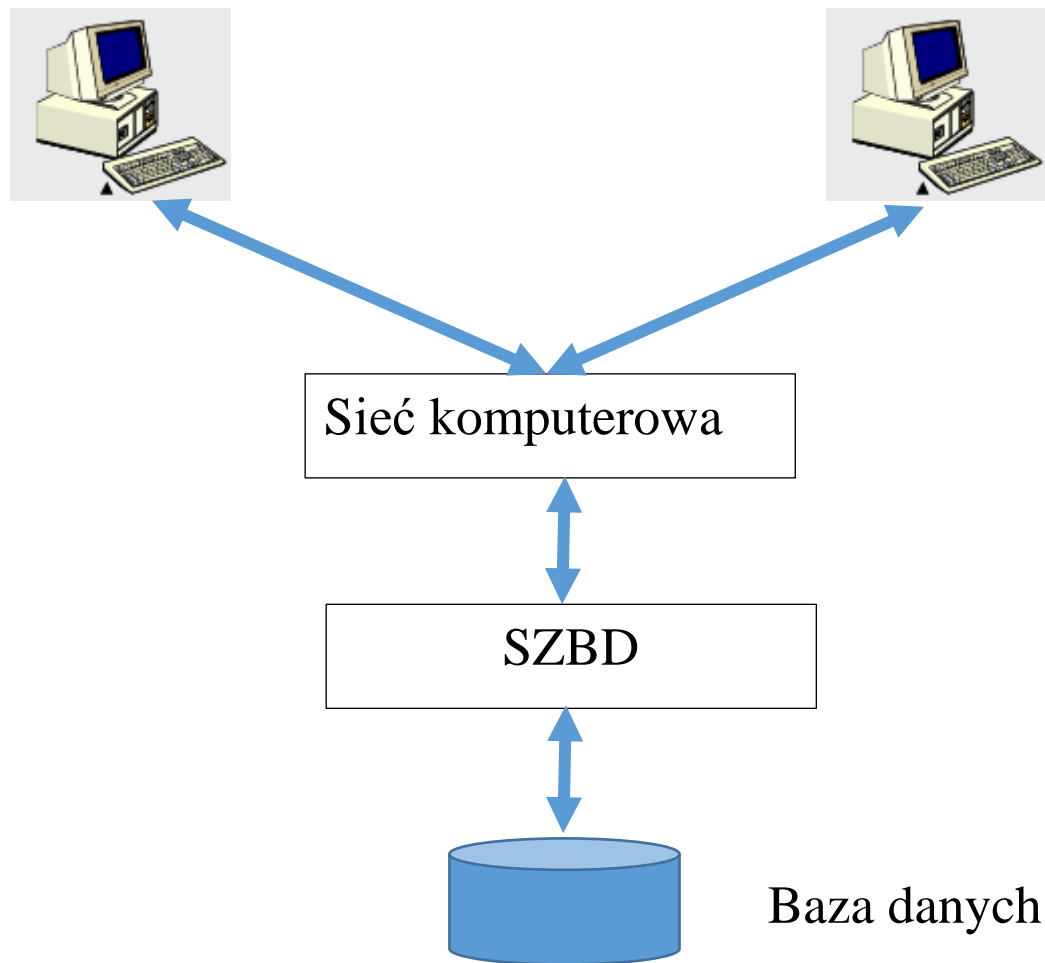
Użytkownicy baz danych



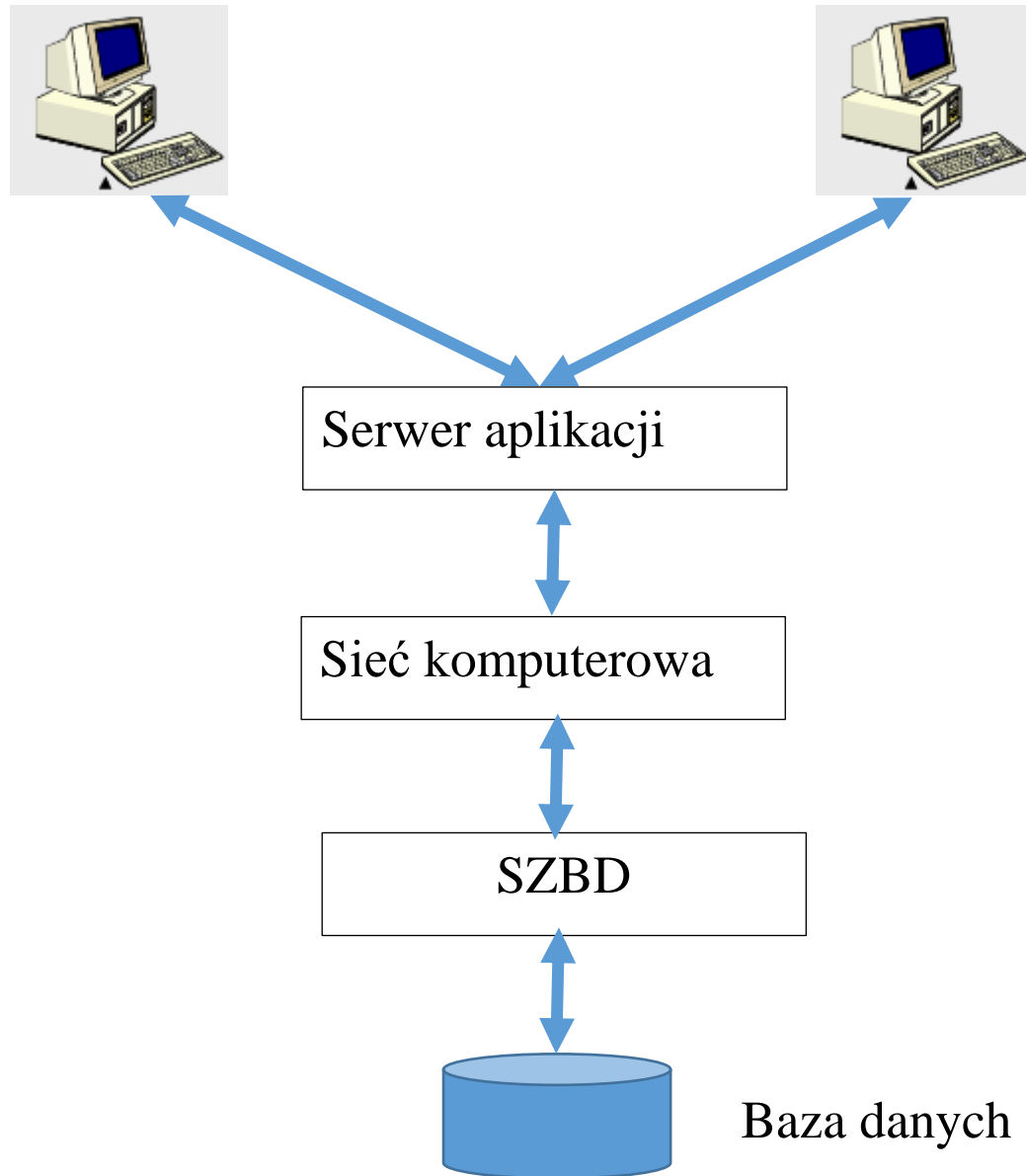
Architektura komunikacyjna

- Architektura 2 – warstwowa.
Klient – DBMS.
- Architektura 3 – warstwowa.
Klient – serwer aplikacji – DBMS.

Architektura systemu klient-serwer



Klient – serwer aplikacji – DBMS

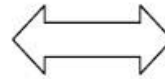


Funkcjonalności Klient - Serwer

Serwer

Typowe funkcje strony serwera aplikacji:

- przechowywanie i organizacja dostępu do danych
- wykonywanie instrukcji języka SQL
- sprawowanie kontroli nad spójnością danych
- zarządzanie zasobami bazy danych, w tym kontami użytkowników



Klienci

Typowe funkcje aplikacji po stronie klienta:

kontakt z użytkownikiem:
przyjmowanie od niego zleceń na operacje,
wykonywanie tych zleceń lub przesyłanie ich w postaci instrukcji języka SQL do serwera bazy danych

Podział baz danych – rodzaj wykonywanych operacji

- Bazy analityczne.
- Bazy operacyjne.

Podział baz danych – liczba baz danych

- Bazy scentralizowane.
- Bazy rozproszone.

Podział baz danych – model

Bazy proste:

- bazy kartotekowe
- sieciowe bazy danych
- hierarchiczne bazy danych

Bazy złożone:

- bazy relacyjne
- bazy obiektowe
- bazy relacyjno-obiektowe
- strumieniowe bazy danych
- temporalne bazy danych
- bazy NoSQL

Bazy kartotekowe (proste) - każda tablica danych jest samodzielnym dokumentem i nie może współpracować z innymi tablicami. Do baz tego typu należą liczne programy typu - książka telefoniczna, książka kucharska, spis książek, kaset lub płyt. Wspólną cechą tych baz jest ich zastosowanie w jednym wybranym celu.

Sieciowe bazy danych

Model historyczny, pozwalał tylko na związki binarne; wiele do jeden.

Hierarchiczne bazy danych

Przykładem hierarchicznej bazy danych jest opracowana przez [IBM](#) baza [IMS](#) (ang. *Information Management System*).

Bazy obiektowe

W bazach obiektowych dane przechowywane są w strukturach obiektowych (zdefiniowanych jako klasy). Koncepcje akademickie dotyczące baz obiektowych były popularne w latach 90., obecnie prace nad nimi są w zaniku.

Bazy relacyjno-obiektowe

Bazy relacyjno-obiektowe pozwalają na manipulowanie danymi jako zestawem obiektów, posiadają jednak bazę relacyjną jako wewnętrzny mechanizm przechowywania danych.

Temporalna baza danych jest odmianą bazy relacyjnej, w której każdy rekord posiada stempel czasowy, określający czas w jakim wartość jest prawdziwa. Posiada także operatory algebry relacyjnej, które pozwalają operować na danych temporalnych (wyciągać historię).

Bazy NoSQL

Skrót NoSQL (Not Only SQL), brak powiązania z relacyjnym modelem. Nazwa często błędnie kojarzona z brakiem języka zapytań SQL. Technologia ta w żaden sposób nie wyklucza elementów języka SQL lub podobnych języków zapytań.

Rozwój baz NoSQL został napędzony obecną potrzebą gromadzenia bardzo dużej ilości danych, które często nie posiadają jednolitej struktury.

W modelu tym największą uwagę zwraca się na dane, a dopiero później na schemacie.

Najpopularniejsze bazy NoSQL

Klucz-wartość - mają bardzo ograniczone możliwości. Przykładem przechowywanych danych może być usługa DNS (Nazwa - Adres IP). Są bardzo wydajne w zapisie i odczycie danych.

Kolumnowe - dane przechowywane są w kolumnach, a nie w wierszach. Korzyścią takiego podejścia jest lepsza wydajność przy operacjach agregujących oraz kompresja. Bazy te jednak są słabym rozwiązaniem przy częstym wykonywaniu operacji DML.

Dokumentowe - dane przechowywane są w dokumentach. Najczęściej ich struktura jest oparta na formatach XML lub JSON. Bazy tego typu nie zapewniają dobrej obsługi transakcji, ale pomimo tego mogą implementować transakcje ACID.

Grafowe - świetnie nadają się w sytuacjach, gdy relacje pomiędzy encjami są ważniejsze niż one same. Wykorzystywane są w sieciach społecznościowych. Podstawą tych baz danych jest teoria grafów.

Bazy relacyjne (złożone) - wiele tablic danych może współpracować ze sobą.

Bazy relacyjne posiadają wewnętrzne języki programowania, wykorzystujące zwykle SQL do operowania na danych, za pomocą których możemy tworzyć zaawansowane funkcje obsługi danych.

Relacyjne bazy danych (jak również przeznaczony dla nich standard SQL) oparte są na kilku prostych zasadach:

1. Wszystkie wartości danych oparte są na prostych typach danych.
2. Wszystkie dane w bazie relacyjnej przedstawiane są w formie dwuwymiarowych tabel (w matematycznym żargonie noszących nazwę "relacji"). Każda tabela zawiera zero lub więcej wierszy (w tymże żargonie - "krotki") i jedną lub więcej kolumn ("atrybuty"). Na każdy wiersz składają się jednakowo ułożone kolumny wypełnione wartościami, które z kolei w każdym wierszu mogą być inne.

3. Po wprowadzeniu danych do bazy możliwe jest porównywanie wartości z różnych kolumn, zazwyczaj również z różnych tabel, i scalanie wierszy, gdy pochodzące z nich wartości są zgodne. Umożliwia to wiązanie danych i wykonywanie stosunkowo złożonych operacji w granicach całej bazy danych.
4. Wszystkie operacje wykonywane są w oparciu o logikę, bez względu na położenie wiersza tabeli. Wiersze w relacyjnej bazie danych przechowywane są w porządku zupełnie dowolnym - nie musi on odzwierciedlać ani kolejności ich wprowadzania, ani kolejności ich przechowywania.
5. Z braku możliwości identyfikacji wiersza przez jego pozycję pojawia się potrzeba obecności jednej lub więcej kolumn niepowtarzalnych w granicach całej tabeli, pozwalających odnaleźć konkretny wiersz. Kolumny te określa się jako "klucz podstawowy" (primary key) tabeli.