

## **Rozproszone i obiektowe systemy baz danych**

### **Opis zadania projektowego**

Dany jest zbiór baz danych { B1, B2, ..., BN } rozproszonych w węzłach sieci komputerowej. Spełnione są następujące założenia:

- dostęp do każdej bazy danych jest możliwy za pośrednictwem serwera zarządzającego bazą danych (np. MySQL, Sybase, Oracle, PostgreSQL, MS SQL Server, Db4o, ObjectStore, Objectivity/DB) i/lub z wykorzystaniem serwera aplikacji (np. JBoss, Glassfish), oraz innych technologii umożliwiających projektowanie wielowarstwowych aplikacji rozproszonych (np. ODBC, JDBC, CORBA, DCOM, MS .NET, ADO, PHP, ASP, JSP, Windows Forms, Windows Presentation Foundation (WPF), Windows Communication Foundation (WCF), Windows Identity Foundation (WIF));
- rozproszona i/lub obiektowa baza danych jest kolekcją baz danych a nie kolekcją plików;
- bazy danych integruje odpowiedni system zarządzania, mechanizmy umożliwiające przetwarzanie oraz dostęp do zdalnych danych (np. łączniki, migawki, perspektywy), mechanizmy sterowania przepływem danych (np. replikacje, transakcje), a także specjalizowane oprogramowanie sieciowe.

Opracować aplikację umożliwiającą dostęp do rozproszonych baz danych, a także odczyt oraz modyfikację odpowiednich informacji w bazach w zależności od zastosowania (np. wyprowadzanie informacji o stanie bankomatów, magazynów, zasobów książkowych, wypożyczalni płyt, połączeniach komunikacyjnych, ofertach biur podróży, dostępności pokoi hotelowych, biletów, itp.). W szczególności dane mogą być publikowane na stronach internetowych, udostępniane przez aplikacje lub usługi sieciowe, aplikacje mobilne.

Do realizacji systemu wykorzystać metody rozpraszania danych, sterowania przepływem oraz przetwarzania danych (np. fragmentacja, alokacja, replikacja, transakcje rozproszone, strumienie, procedury, migawki, perspektywy, łączniki, synonimy) dostępne w wybranym systemie bazodanowym.

**Ustalić skład grup projektowych (grupy 1-2 osobowe), szkolenie BHP.**

**Przedstawić podczas zajęć i wysłać na maila osoby prowadzącej zajęcia opis wstępnych założeń projektowych (wzór na stronie WWW; terminy: 6 zajęcia – grupy parzyste; 7 zajęcia – grupy nieparzyste).**

**Realizować oraz prezentować podczas zajęć, w podanych terminach, kolejne fazy rozwoju systemu rozproszonej bazy danych i aplikacji klienckiej (np. demonstracja działania, prezentacje slajdów z przykładami).**

### **Harmonogram**

- **Zajęcia (1):** sprawy organizacyjne; bhp; ustalanie składu i numerów grup projektowych; w kolejnych tygodniach grupy o numerach nieparzystych przychodzą na zajęcia o numerach nieparzystych, tj. 3, 5, 7, 9, 11, 13, a także na zajęcia 14 i 15, a grupy o numerach parzystych przychodzą na zajęciach o numerach parzystych, tj. 2, 4, 6, 8, 10, 12, a także na zajęcia 14 i 15 (uwaga: jeśli grupa nie może przyjść na swój termin lub termin wypadnie z powodu, np. godzin rektorskich lub innych, to można przyjść w kolejnym tygodniu).
- **Zajęcia (2, 3, 4, 5):** analiza literaturowa, definiowanie tematów i wstępnych założeń projektowych; grupy nieparzyste – zajęcia 3 i 5; grupy parzyste – zajęcia 2 i 4.
- **Zajęcia (6):** przesłanie sprawdzonej, finalnej wersji opisu założeń do projektu (według podanego wzoru) na maila prowadzącego; prezentacja wymagań funkcjonalnych i

niefunkcjonalnych rozproszonej aplikacji bazodanowej – diagram przypadków użycia, wybrane scenariusze użycia; grupy parzyste.

- Zajęcia (7): przesłanie sprawdzonej, finalnej wersji opisu założeń do projektu (według podanego wzoru) na maila prowadzącego; prezentacja wymagań funkcjonalnych i niefunkcjonalnych rozproszonej aplikacji bazodanowej – diagram przypadków użycia, wybrane scenariusze użycia; grupy nieparzyste.
- Zajęcia (8): testowanie mechanizmów replikacji i/lub rozpraszania danych w wybranych środowiskach bazodanowych (np. PostgreSQL) oraz środowiskach wirtualizacji (np. Docker, Kubernetes); grupy parzyste.
- Zajęcia (9): testowanie mechanizmów replikacji i/lub rozpraszania danych w wybranych środowiskach bazodanowych (np. PostgreSQL) oraz środowiskach wirtualizacji (np. Docker, Kubernetes); grupy nieparzyste.
- Zajęcia (10): projekt i implementacja rozproszonej bazy danych; grupy parzyste.
- Zajęcia (11): projekt i implementacja rozproszonej bazy danych; grupy nieparzyste.
- Zajęcia (12): projekt i implementacja prostej aplikacji wykorzystującej rozproszoną bazę danych; projekt oraz implementacja mechanizmów loadbalancingu (np. HAProxy, Nginx), zapewniających wysoką dostępność systemu oraz równoważenie obciążenia serwerów bazodanowych i/lub serwerów aplikacji; grupy parzyste.
- Zajęcia (13): projekt i implementacja prostej aplikacji wykorzystującej rozproszoną bazę danych; projekt oraz implementacja mechanizmów loadbalancingu (np. HAProxy, Nginx), zapewniających wysoką dostępność systemu oraz równoważenie obciążenia serwerów bazodanowych i/lub serwerów aplikacji; grupy nieparzyste.
- Zajęcia (14\_obie): testy i ocena działania systemu rozproszonego; prezentacja działania aplikacji, spisu treści i/lub wstępnej wersji sprawozdania z projektu; grupy parzyste i nieparzyste.
- Ostatnie zajęcia (15\_obie): testy i ocena działania systemu rozproszonego; prezentacja działania aplikacji, spisu treści i/lub wstępnej wersji sprawozdania z projektu; grupy nieparzyste i parzyste.

**Przedstawić podczas zajęć stronę tytułową oraz spis treści sprawozdania, a następnie wysłać na adres e-mail osoby prowadzącej końcowe sprawozdanie z projektu, które powinno zawierać następujące elementy (przykładowy spis treści znajduje się na stronie WWW).**

1. Stronę tytułową.
2. Spis treści, spis rysunków, spis tabel, spis listingów (zobacz przykład na stronie WWW).
3. Cel i zakres projektu.
4. Opis działania systemu. Wymagania funkcjonalne (funkcje systemu) i niefunkcjonalne (narzędzia i technologie, parametry wydajnościowe systemu, i inne).
5. Założenia przyjęte podczas realizacji, np. rodzaj klienta, architektura wykorzystywanych baz danych (jednorodne - homogeniczne, niejednorodne – heterogeniczne systemy baz danych; obiektowe, obiektowo-relacyjne, relacyjne bazy danych), sposób rozproszenia danych (fragmentacja, alokacja, replikacja), mechanizmy sterowania przepływem i przetwarzania danych (transakcje, strumienie, wyzwalacze, procedury, migawki, perspektywy, łączniki, synonimy).
6. Wykorzystywane serwery baz danych, serwery aplikacji, metody, narzędzia, technologie projektowania oraz implementacji systemu (np. język specyfikacji UML, wzorce projektowe, stosowane narzędzia wspomagające projektowanie).
7. Projekt bazy danych. Modele conceptualne, logiczne i fizyczne rozproszonych baz danych. Schemat globalny bazy danych, schematy lokacji lokalnych. Normalizację schematów relacji.
8. Architekturę systemu i schemat komunikacji w rozproszonym systemie baz danych. Mechanizmy rozpraszania oraz integracji danych, sterowania przepływem i przetwarzania danych, mechanizmy loadbalancingu (zapytania, perspektywy, migawki, wyzwalacze, transakcje, indeksowanie, pliki parametrów, procesy bazodanowe), a także wykorzystywane protokoły komunikacyjne.
9. Projekt aplikacji (interfejsu) użytkownika. Sposoby dostępu do baz danych (przetwarzanie transakcyjne, perspektywy, migawki, zapytania SQL, funkcje API, mapowanie relacyjno-obiektowe, warstwa pośrednicząca, serwery aplikacji, serwery WWW) oraz wykorzystane mechanizmy bezpieczeństwa (np. sposób uwierzytelniania, uprawnienia grup użytkowników - role, szyfrowanie danych w bazie danych, szyfrowanie połączeń sieciowych).

10. Przykłady implementacyjne: prosta, rozproszona baza danych, wybrane elementy aplikacji użytkownika (interfejsu).
11. Sposób instalowania, uruchamiania i testowania aplikacji. Konfigurowanie środowiska testowego (np. maszyny wirtualne, sieć komputerowa, loadbalancing) i/lub produkcyjnego.
12. Ocenę wydajności i optymalizację działania systemu (czasy przetwarzania zapytań, stopień obciążenia systemu, mechanizmy zwiększające wydajność i dostępność systemu, itp.)
13. Podsumowanie – wnioski końcowe, właściwości wykorzystywanych technologii i mechanizmów przetwarzania rozproszonego.
14. Wykorzystane pozycje literaturowe.

## Przykładowe tematy

- Aplikacja bazodanowa z mechanizmem replikacji i loadbalancingu (wybrane środowiska, np. MySQL, PostgreSQL, Oracle, MS SQL Server, Db4o, bazy typu No-SQL).
- Rozproszona (obiektowa, relacyjna) baza danych połączeń kolejowych, autobusowych, lotniczych.
- System obsługi bankomatów, kas biletowych, itp.
- System monitorowania ofert hoteli, schronisk, biur podróży zintegrowany z rozproszoną bazą danych.
- System monitorowania stanu magazynów, hurtowni, urzędzeń.
- Rozproszona baza danych wypożyczalni samochodów, książek, kaset DVD.
- System zdalnego rozliczania stanu liczników wody, energii cieplnej, elektrycznej, czasu pracy stacji roboczych.
- System zakupów internetowych zintegrowany z rozproszoną bazą danych.
- Projektowanie mechanizmów przetwarzania transakcyjnego, mechanizmu strumieni, zapytań rozproszonych w rozproszonych systemach baz danych.
- Projektowanie „przeźroczystego” dostępu do rozproszonych (obiektowych, relacyjnych) baz danych.
- Projektowanie bezpiecznych systemów rozproszonego przetwarzania danych wykorzystujących mechanizmy kryptograficzne, technologię blockchain.
- Rozproszone bazy danych w zastosowaniach Internetu Rzeczy (IoT).

## Literatura

- [1] Fowler M., Rice D., Foemmel M., Heatt E., Mee R., Stafford R., Architektura systemów zarządzania przedsiębiorstwem. Wzorce projektowe, Helion, Gliwice, 2005.
- [2] Wrembel R., Bębel B., Oracle. Projektowanie rozproszonych baz danych, Helion, Gliwice, 2003.
- [3] Date C. J., Wprowadzenie do systemów baz danych, WNT, Warszawa, 2000.
- [4] Beynon-Davies P., Systemy baz danych, WNT, Warszawa, 2000.
- [5] Coulouris G., Dollimore J., Kindberg T., Systemy rozproszone - podstawy i projektowanie, WNT, Warszawa, 1998.
- [6] Banachowski L., Bazy danych - tworzenie aplikacji, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1998.
- [7] Henderson K., Bazy danych w architekturze klient/serwer, Wydawnictwo Robomatic, 1997.
- [8] Silberschatz A., Peterson J. L., Galvin P. B., Podstawy systemów operacyjnych, WNT, Warszawa, 1993.
- [9] Bell D., Grimson J., Distributed Database Systems, Addison Wesley, 1992.
- [10] Ozsu T. M., Valduriez P., Principles of Distributed Database Systems, Prentice Hall, 1999.
- [11] Kim W., Wprowadzenie do obiektowych baz danych, WNT, Warszawa, 1996.
- [12] Harrington J.L., Obiektowe bazy danych dla każdego, MIKOM, Warszawa, 2001.
- [13] Strona internetowa: <http://www.db4o.com>
- [14] Strona internetowa: <http://wazniak.mimuw.edu.pl/index.php> - systemy rozproszone, zaawansowane systemy baz danych.