**Wstęp**

Replikacja danych polega na przechowywaniu kopii danych na wielu komputerach. Istnieje wiele korzyści z zastosowania tej techniki:

**zwiększona wydajność** - większe obciążenia (na przykład żądania klientów) mogą być tolerowane, ponieważ obciążenie jest dzielone między kilka procesów. Ponadto, zmniejszone może być opóźnienie replikacji danych bliżej użytkownika. Niestety, korzyści są mniejsze, jeżeli replikowane dane są do odczytu / zapisu.

**zwiększona dostępność** - replikacja pomaga w tolerowaniu poszczególnych awarii serwera. Jeśli połączenie z serwerem nie powiedzie się z prawdopodobieństwem p, liczba serwerów potrzebnych do zapewnienia określonego poziomu usług to - Dostępność = 1-PN.

**odporność na awarie** - podobnie jak dostępność, ale gwarantuje poprawność oprócz dostępności. Na przykład, jeśli serwer grupy serwerów n posiada złą informację inne mogą przegłosować nieprawidłowy serwer w celu zapewnienia poprawnych danych dla klienta.

Z punktu widzenia klienta widzenia, jest tylko jedna logiczna kopia danych. Jeśli klient dokonuje aktualizacji danych to zmiany powinny znaleźć odzwierciedlenie we wszystkich istniejących replikach.

**Podstawowy model** zarządzania replikowanymi danymi:

- Klienci wysyła żądanie do systemu(front-end).  
- System(front-end) zapewnia przejrzystość, ukrywając fakt, że dane są replikowane.  
- System(front-end) komunikuje się z jednym lub więcej menedżerów repliki do pobierania / zapisywania danych.  
- Zarządcy replik współdziałają w celu zapewnienia, że dane są spójne.

Istnieje kilka modeli zapewniających replikację: pasywne, aktywne i leniwe.