WYŻSZA SZKOŁA BANKOWA
Warszawa

Programowanie aplikacji w Java

Maciej Gowin

Zjazd 10 - dzień 1

Zawartość

- Wdrażanie i hosting aplikacji
- CI/CD
- Strategie integracji kodu
- Strategie wdrażania aplikacji
- Zadanie z CI/CD
- Bazy NoSQL
- Zadanie z NoSQL

Wdrażanie aplikacji

Do tej pory zajmowaliśmy się implementacją kodu naszej aplikacji w Springu oraz uruchamialiśmy ją lokalnie, na naszych komputerach.

Załóżmy, że skończyliśmy właśnie pracę nad jej pierwszą wersją.

Naszym kolejnym zadaniem jest dostarczenie jej do użytkowników.

Rozważymy teraz istniejące opcje dostarczenia naszej aplikacji do użytkowników oraz przejdziemy przez cały proces wdrażania aplikacji.

Hosting

Udostępnienie aplikacji z naszego własnego komputera wymaga dużo wysiłku związanego z administracją serwera. O ile nie
jesteśmy firmą, która chce poświęcić wiele zasobów na zbudowanie i utrzymywanie własnej serwerowni, najlepszą opcją
będzie skorzystanie z usług firm zewnętrznych.

Na rynku jest wiele firm, które posiadają serwerownie złożone z wielu komputerów, które udostępniają swoich klientom moc
obliczeniową potrzebną do funkcjonowania aplikacji. Udostępnianie zasobów serwera klientowi nazywamy **hostingiem**.

Istnieje wiele rodzajów hostingu, które różnią się sposobem rozliczania płatności oraz podziałem odpowiedzialności za
poszczególne czynności związane z funkcjonowaniem naszej aplikacji.

Rodzaje hostingu

- Serwer dedykowany
- Hosting współdzielony
- Serwer VPS (Virtual Private Server)
- Hosting zarządzany
- Kolokacja
- Hosting w chmurze

Serwer dedykowany

Serwer dedykowany jest dostępny tylko dla nas i nasza aplikacja jest jedyną uruchomioną na nim.

W przypadku hostingu na serwerze dedykowanym mamy największą kontrolę nad serwerem, na którym uruchomiona jest nasza
aplikacja. Mamy pełen dostęp administratora do wszystkich zasobów serwera.

Serwer dedykowany wymaga najwięcej wysiłku oraz wiedzy z zakresu instalacji i zarządzania nim.

Hosting współdzielony

W przypadku hostingu współdzielonego mamy bardzo ograniczone możliwości administracji serwerem, ponieważ na jednym
serwerze uruchomione może być wiele aplikacji różnych klientów.

Serwery współdzielone pozwalają na lepsze wykorzystanie zasobów poprzez uruchamianie wielu aplikacji na jednej maszynie.
Dzięki temu są znacznie tańsze od serwerów dedykowanych.

Oprócz ograniczonej kontroli nad serwerem, wadą tego typu hostingu jest możliwy wzajemny wpływ aplikacji uruchomionych
na tym samym serwerze.

Serwer VPS

Serwer VPS jest rozwiązaniem pomiędzy serwerem dedykowanym i hostingiem współdzielonym.

Na jednej maszynie uruchomione jest wiele aplikacji, ale każda z nich jest uruchomiona w odizolowanej przestrzeni (
maszynie wirtualnej).

Dzięki temu możliwa jest kontrola nad serwerem zbliżona do serwera dedykowanego, przy jednoczesnym wykorzystaniu zasobów
podobnym do hostingu współdzielonego.

Maszyna wirtualna pozwala na uruchomienie więcej niż jednego systemu operacyjnego na tym samym serwerze, jednak oznacza
to również, konieczność alokacji oddzielnych zasobów dla każdego z nich.

Hosting zarządzany

W przypadku hostingu zarządzanego oprócz serwera otrzymujemy również wsparcie techniczne związane z konfiguracją sprzętu
oraz oprogramowania.

Wsparcie techniczne może obejmować monitorowanie czy instalacje aktualizacji poprawiających bezpieczeństwo serwera.

W zależności od dostawcy, pakiety usług wchodzących w skład hostingu zarządzanego mogą się różnić.

Kolokacja

Kolokacja polega na wynajęciu fizycznej przestrzeni na serwery.

Dzięki temu możemy sami zdecydować, z jakiego rodzaju sprzętu chcemy korzystać.

Rola firmy zewnętrznej ogranicza się do utrzymania dostępu do energii, internetu oraz chłodzenia naszego serwera.

Hosting w chmurze

Hosting w chmurze charakteryzuje się możliwością uruchomienia aplikacji z użyciem połączonych zasobów wielu komputerów.

Dostawcy chmury dostarczają narzędzia developerskie oraz usługi takie jak np. bazy danych, które mogą zostać wdrożone
bez konieczności ręcznej instalacji oraz pozwalają ograniczyć czynności związane z ich utrzymaniem.

Dzięki temu ułatwione jest budowanie i zarządzanie infrastrukturą naszej aplikacji. Pozwala to poprawić niezawodność
naszej aplikacji.

Rodzaje hostingu w chmurze - 1

- **Chmura publiczna** - zasoby chmury są współdzielone przez użytkowników
  - **IaaS** (Infrastructure as a Service) - otrzymujemy serwer wirtualny, ale jesteśmy odpowiedzialni za jego
    konfigurację oraz zarządzanie środowiskiem dla naszej aplikacji
  - **PaaS** (Platform as a Service) - oprócz serwera wirtualnego otrzymujemy gotowe środowisko do uruchomienia naszej
    aplikacji
  - **FaaS** (Function as a Service) - pozwala uruchomić nasz kod na żądanie, bez rezerwowania żadnego serwera (
    serverless). Płacimy tylko za zużyte przez naszą aplikację zasoby
  - **SaaS** (Software as a Service) - otrzymujemy gotowe do użycia oprogramowanie

Rodzaje hostingu w chmurze - 2

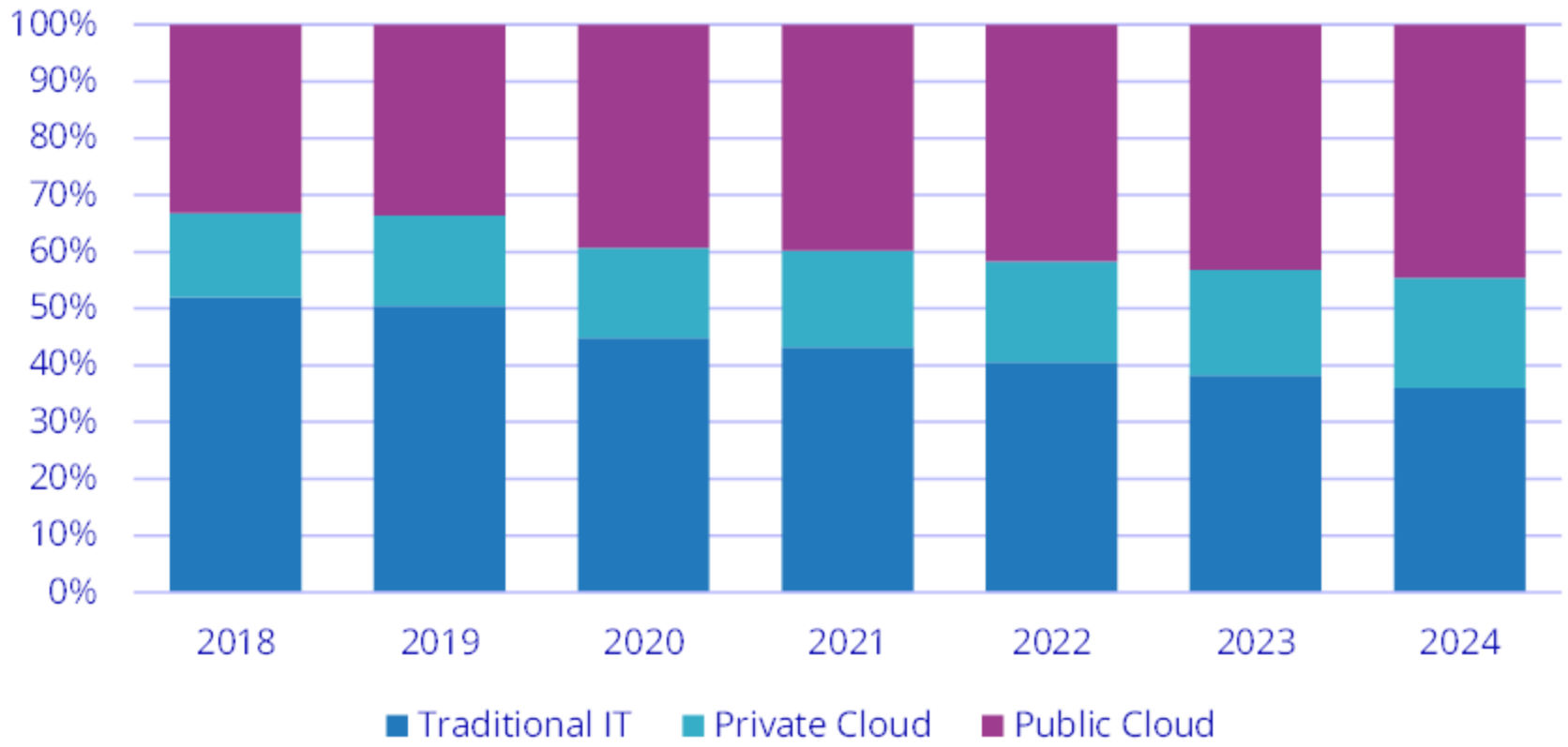
- **Chmura prywatna** - pozwala na korzystanie z usług chmurowych na sprzęcie, który posiadamy na własność
- **Chmura hybrydowa** - nasze zasoby składają się zarówno z chmury publicznej jak i prywatnej
- **Multicloud** - nasze zasoby składają się z usług więcej niż jednego dostawcy chmury

Podejścia do chmury

- **Cloud native** - używamy usług specyficznych dla danej chmury
  - Trudniejsza zmiana dostawcy chmury (vendor locking)
  - Lepsze wykorzystanie zasobów chmury poprzez wykorzystanie zoptymalizowanych dla niej serwisów
- **Cloud agnostic** - używamy usług, które mogą zostać wdrożone na różnych chmurach
  - Łatwiejsza zmiana dostawcy chmury
  - Niekorzystanie z zoptymalizowanych przez dostawcę usług, powoduje zazwyczaj gorszą wydajność oraz wyższy koszt
    rozwiązania

IDC
ANALYZE THE FUTURE

Worldwide Cloud IT Infrastructure Market Forecast by Deployment Type, 2018- 2024 (shares based on Value)



Source: IDC 2021

CI/CD - 1

Do tej pory zaimplementowaliśmy naszą aplikację, wybraliśmy hosting i dostarczyliśmy jej pierwszą wersję do naszych
użytkowników. Jako zespół pracujący nad tą aplikacją planujemy teraz dodanie nowych funkcjonalności do jej kolejnej
wersji.

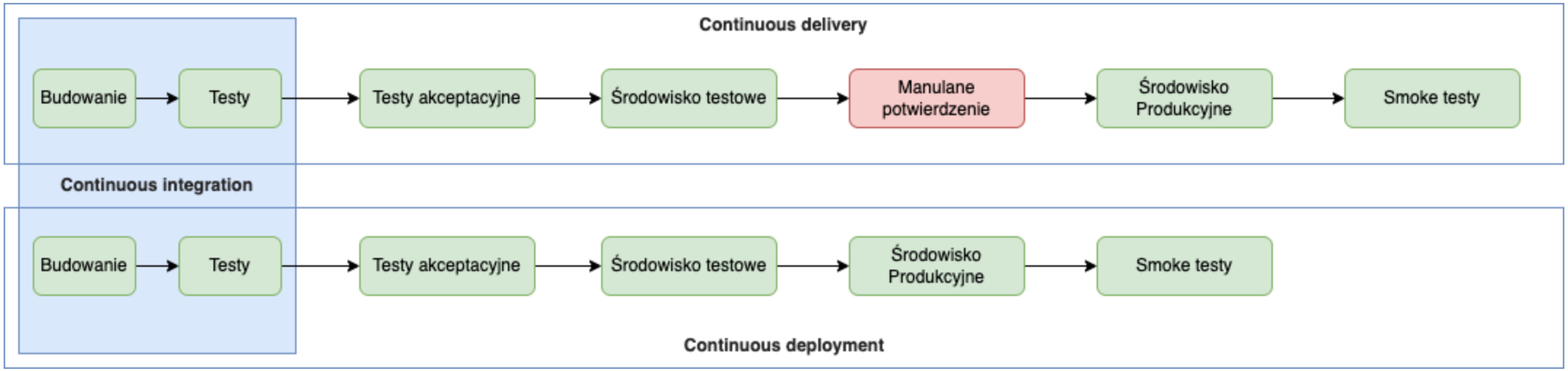
Aby zaoszczędzić nasz czas, chcemy zautomatyzować process wdrażania kolejnych wersji aplikacji. W tym celu zamierzamy
stworzyć process CI/CD.

CI/CD -2

CI/CD oznacza:

- **CI** (Continuous Integration) - regularne budowanie i testowanie zmian w kodzie różnych członków zespołu pracujących
  nad aplikacją.
- **CD** (Continuous Delivery) - zmiany są automatycznie wdrażane i testowane na środowisku testowym. Przed
  dostarczeniem nowej wersji dla użytkownika konieczne jest manualne potwierdzenie.
- **CD** (Continuous Deployment) - w porównaniu do Continuous Delivery, po przejściu przez środowisko testowe zmiany są
  automatycznie dostarczane do użytkowników aplikacji.

CI/CD - Continuous Integration vs Continuous Delivery vs Continuous Deployment

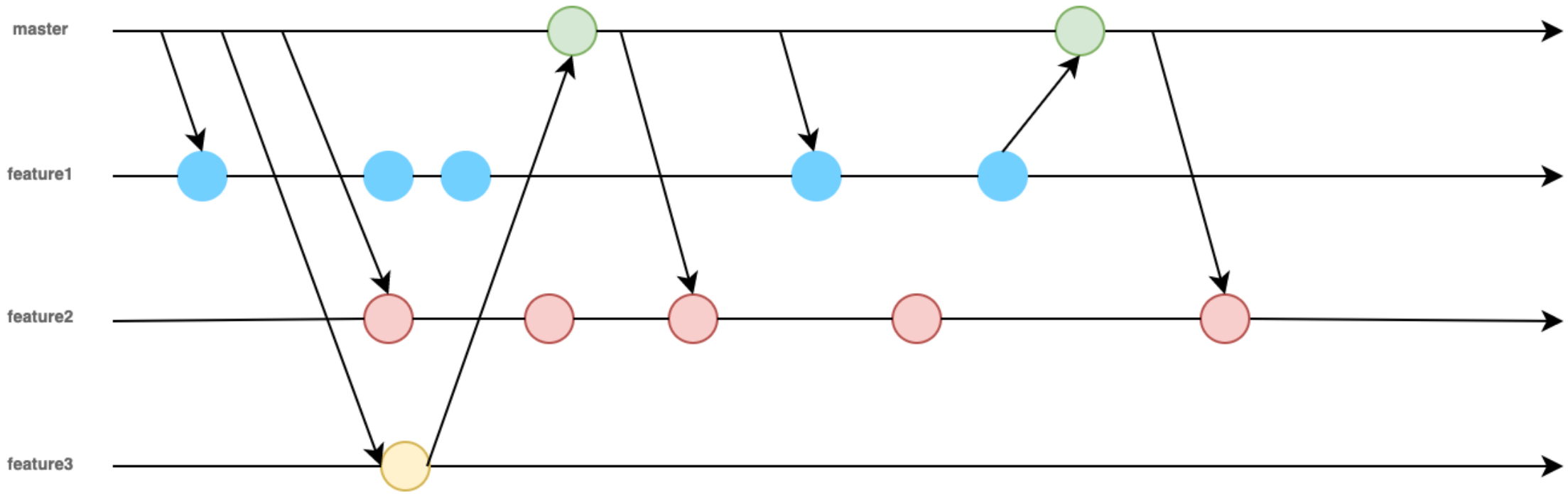


Strategie integracji kodu

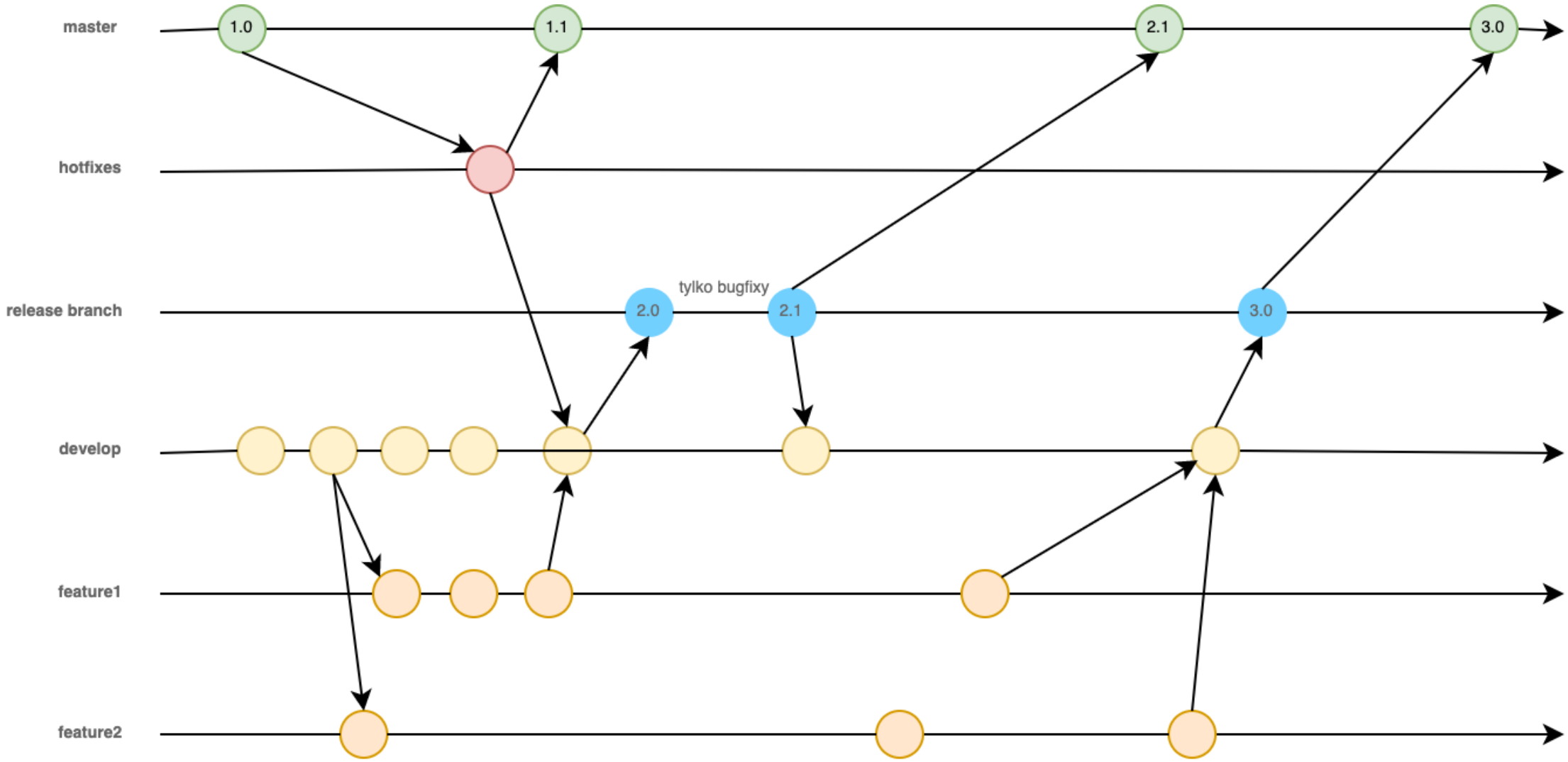
Najpopularniejsze strategie integracji kodu to:

- Trunk-based development
- Git-flow
- Github-flow
- Gitlab-flow

Trunk-based development



Git-flow

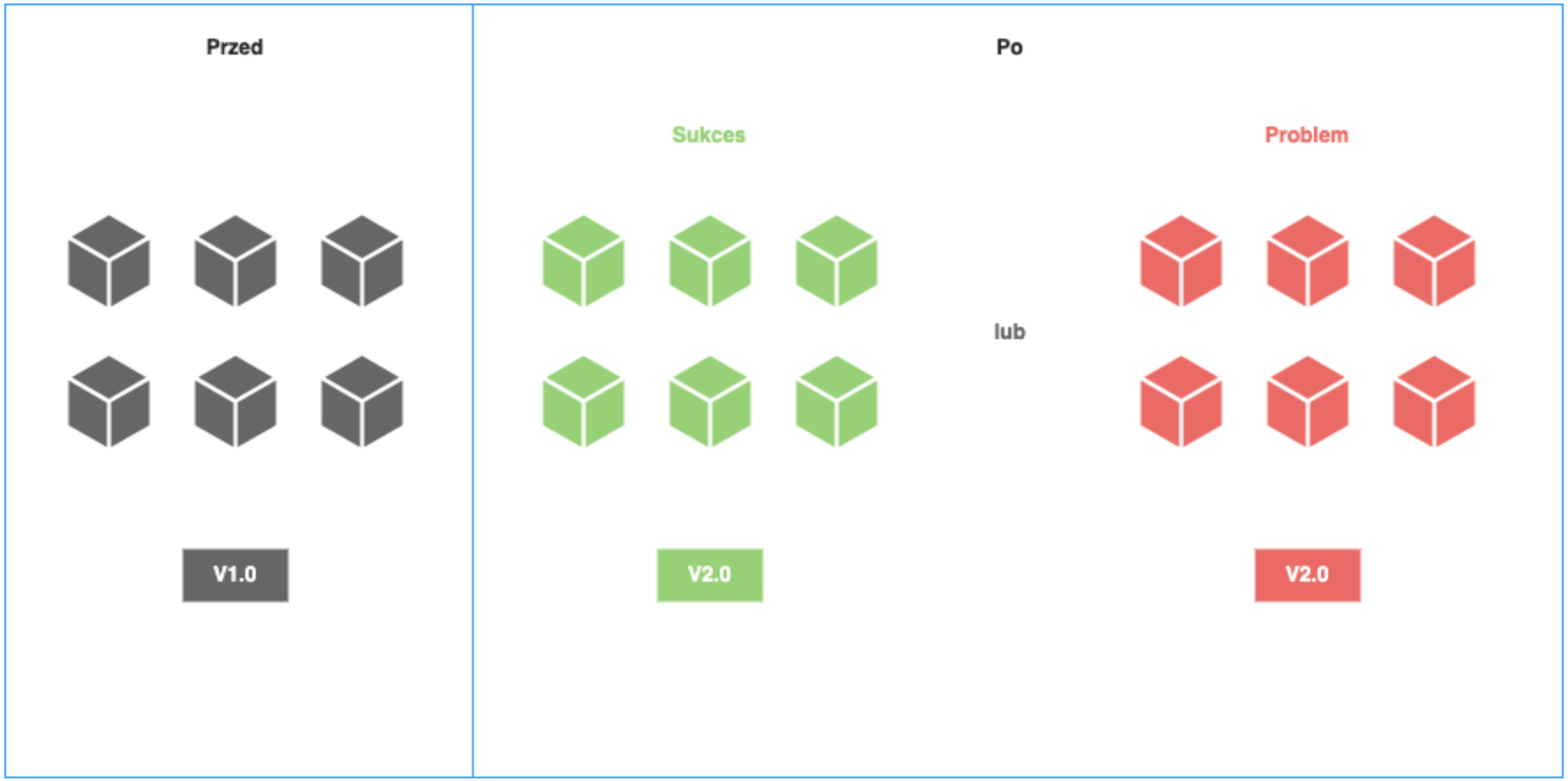

master
1.0
1.1
2.1
3.0
hotfixes
release branch
tylko bugfixy
2.0
2.1
3.0
develop
feature1
feature2

Strategie dostarczania kolejnych wersji aplikacji

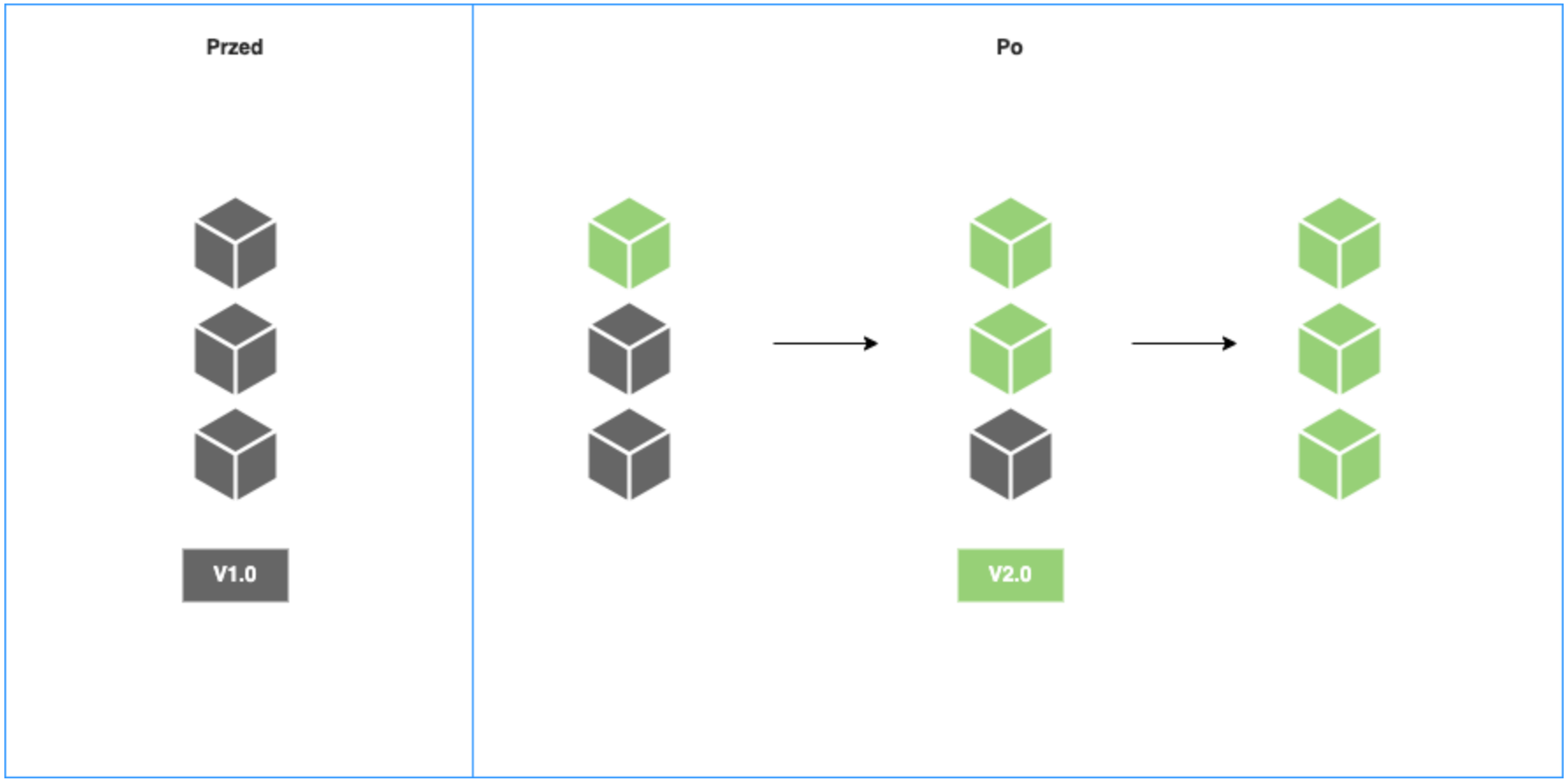
Najczęściej używanymi strategiami dostarczania nowych wersji aplikacji do użytkownika są:

- Basic Deployment
- Rolling update
- Blue-Green deployment
- Canary deployment
- A/B Testing

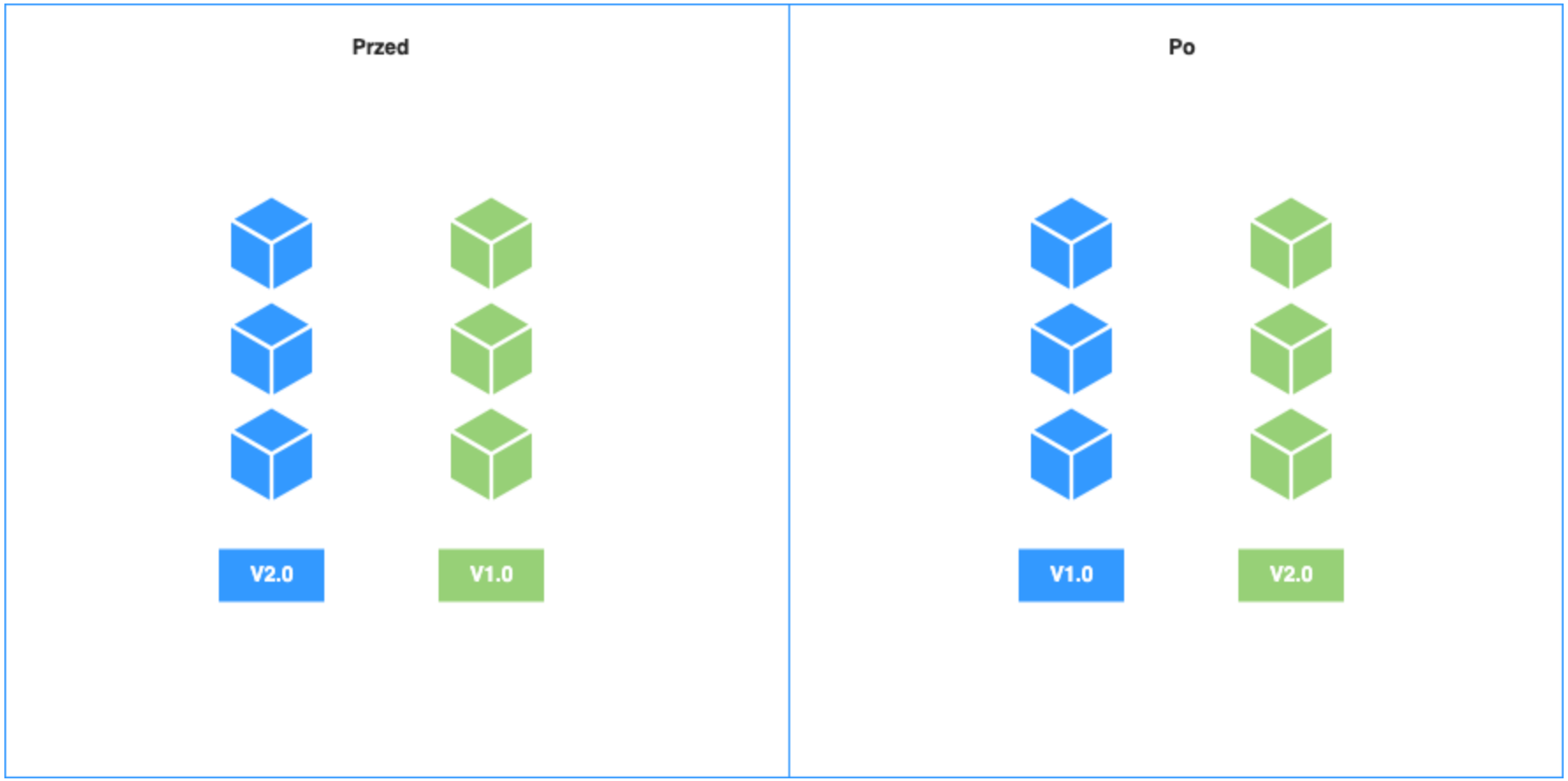
Basic deployment



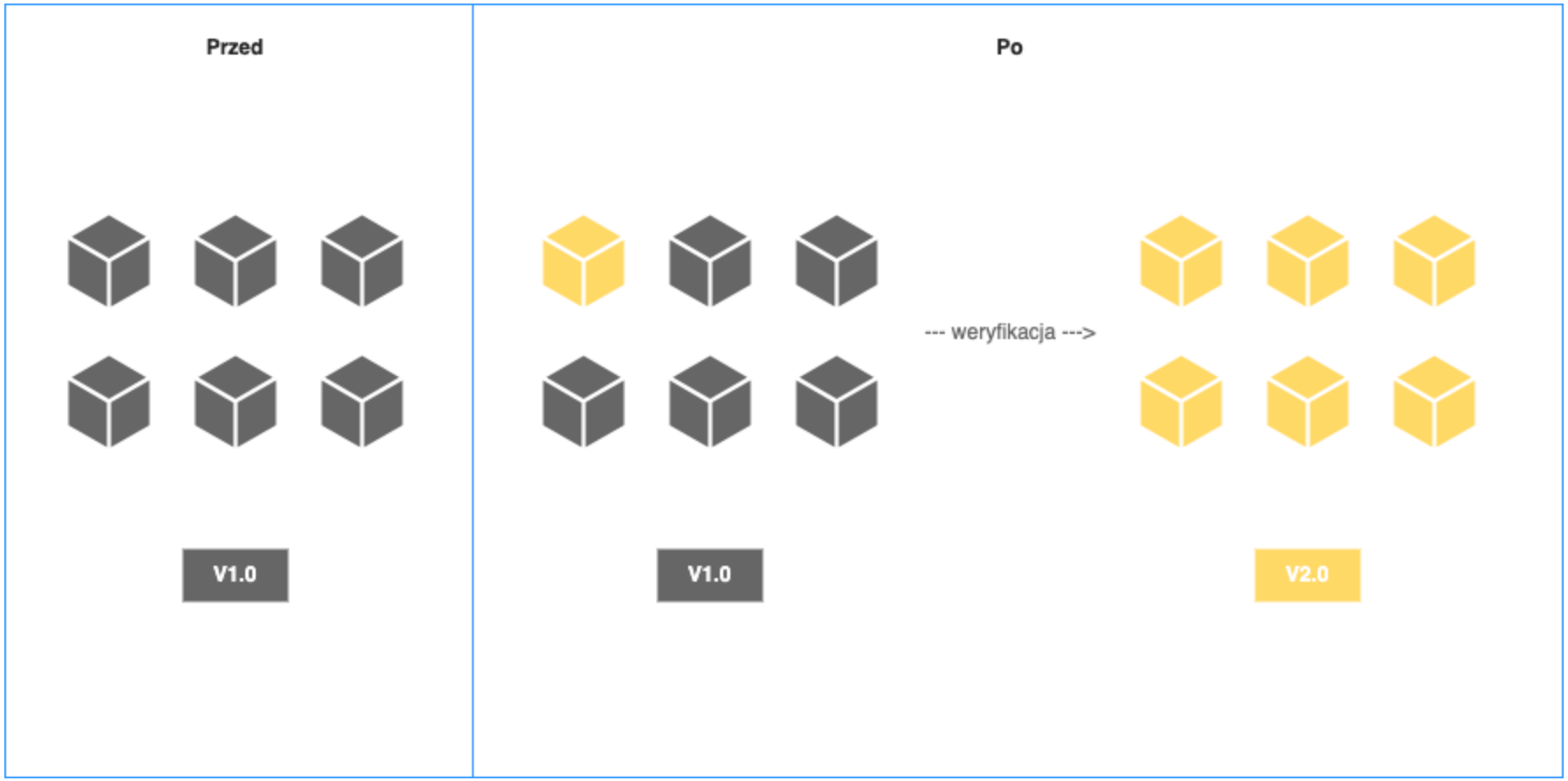
Rolling update



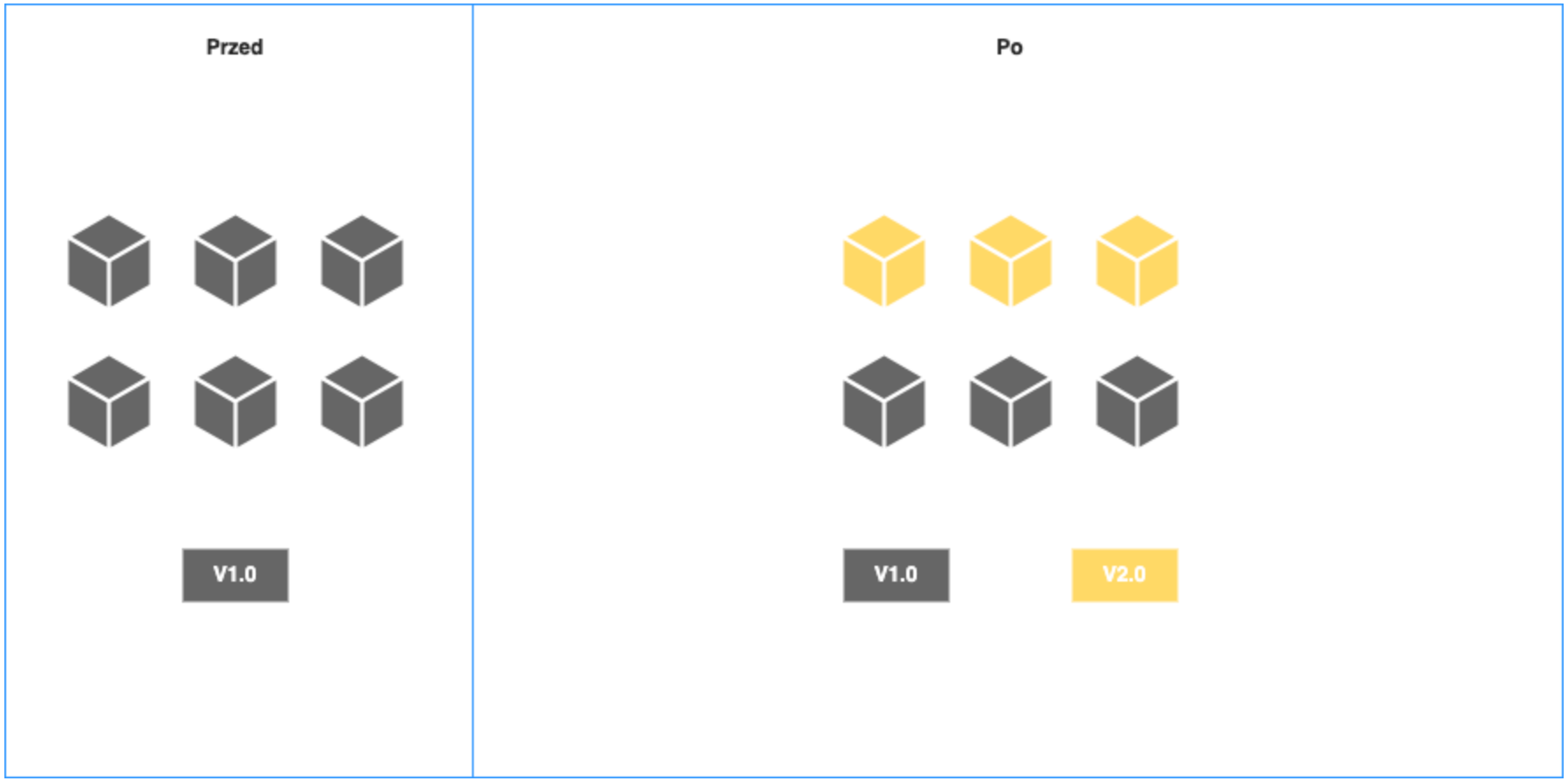
Blue-Green deployment



Canary deployment



A/B Testing



Narzędzia

Do implementacji naszego procesu CI/CD możemy użyć jednego z popularnych narzędzi, np.

- Gitlab CI/CD
- Jenkins
- Atlassian Bamboo
- Circle CI
- Buddy
- AWS CodePipeline

Zadanie

Zaimplementuj uproszczoną wersję procesu CI/CD z użyciem Gitlab CI/CD. Proces ten powinien pobrać najnowszą wersję kodu
aplikacji napisanej w Javie, zbudować ją oraz uruchomić testy.

Bazy NoSQL

Nadszedł czas rozpoczęcia kolejnego projektu. Z wymagań wynika, że musimy być przygotowani na obsługę dużego ruchu do
naszej aplikacji oraz przechowywanie bardzo dużej ilości danych.

Bazy danych SQL, które poznaliśmy do tej pory, oferują możliwości skalowania w celu obsługi większego ruchu i ilości
danych, jednak są w tym aspekcie dosyć ograniczone i wymagają dodatkowych czynności w celu jej zapewnienia.

Dla takich zastosowań powstały właśnie bazy NoSQL. Dostarczają one znacznie lepsze możliwości skalowania bez
podejmowania dodatkowych działań.

Ponadto w wielu przypadkach znacznie ułatwiają modelowanie danych przechowywanych w bazie danych.

Typy baz NoSQL

Najpopularniejsze typy baz NoSQL to:

- **Klucz-wartość** - klucz jest znany, ale wartości nie. Sprawdzają się w przechowywaniu nieustrukturyzowanych danych.
- **Dokumentowe** - rozszerzenie baz klucz-wartość, oferują możliwość zagnieżdżania par klucz-wartość i używania ich w zapytaniach.
- **Rodziny kolumn** - dane są przechowywane w postaci rodziny kolumn. Pozwala to na szybkie przeszukiwanie dużej ilości danych, ale wymaga to dobrze przemyślanego schematu bazy.
- **Grafowe** - bazy reprezentowane w postaci grafu. Węzły grafu reprezentują dane, a krawędzie relacje między nimi.

Bazy klucz wartość i dokumentowe

```
{
  "id": "1",
  "nazwa": "Hotel WSB",
  "adres": {
    "ulica": "Sportowa 12",
    "miasto": "Warszawa",
    "kraj": "Polska"
  },
  "pokoje": [
    {
      "numer": 1,
      "liczbaPokoi": 2
    },
    {
      "numer": 2,
      "liczbaPokoi": 3
    }
  ]
}
```

Rodziny kolumn - 1

Klucz składa się z wielu posortowanych kolumn

Przykład:

Klucz = Firma/Linia Autobusowa/Czas/Numer rejestracyjny

Klucz	Lokalizacja
MTA/M86-SBS/2020-01-01T13:01:00/NYCT_5824	(40.781212,-73.961942)
MTA/M86-SBS/2020-01-01T13:02:00/NYCT_5840	(40.780664,-73.958357)
MTA/M86-SBS/2020-01-01T13:03:00/NYCT_5867	(40.780281,-73.946890)

Rodziny kolumn - 2

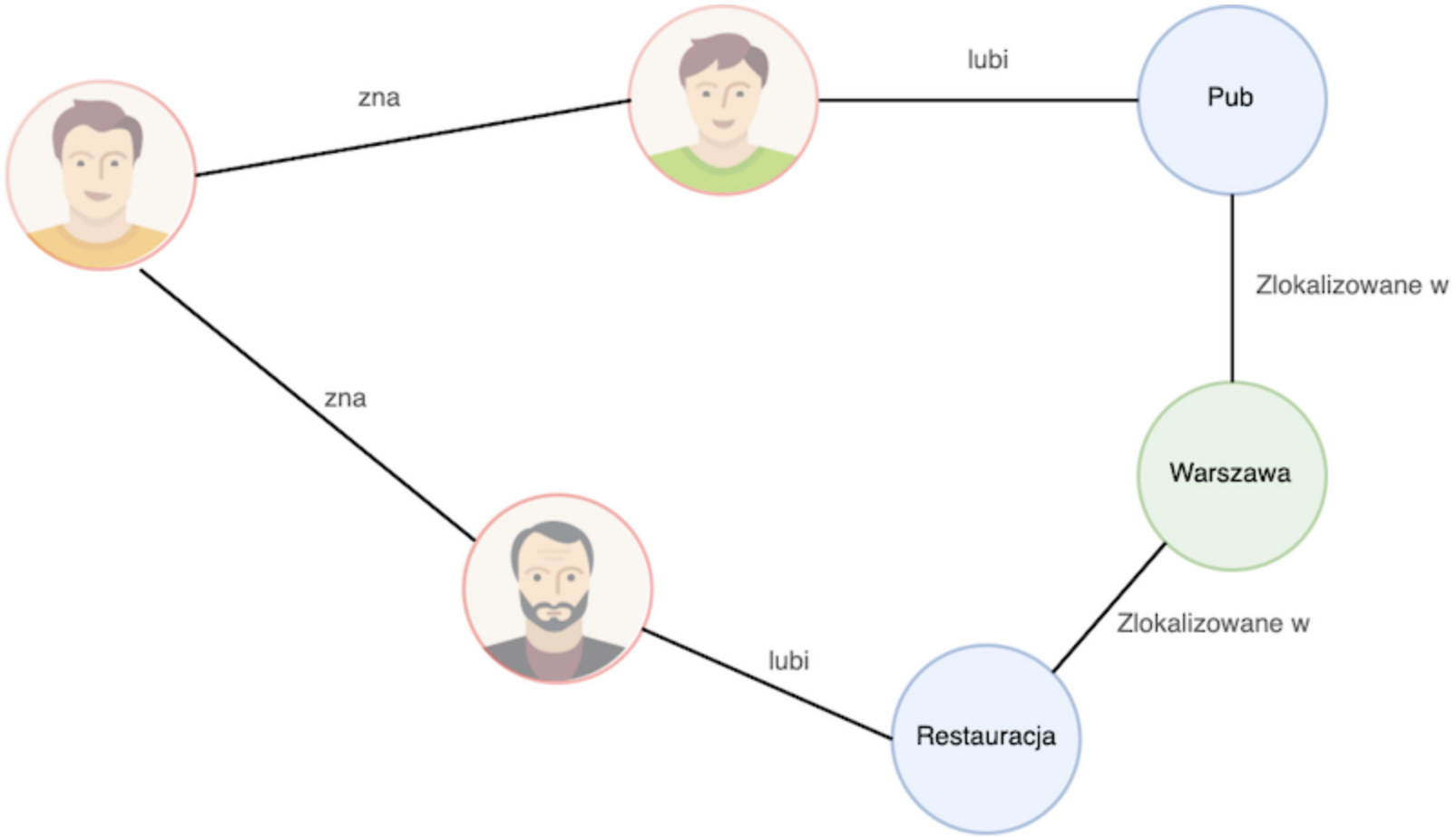
Przykład wydajnego zapytania:

- Lokalizacje konkretnego busa w danym zakresie czasu

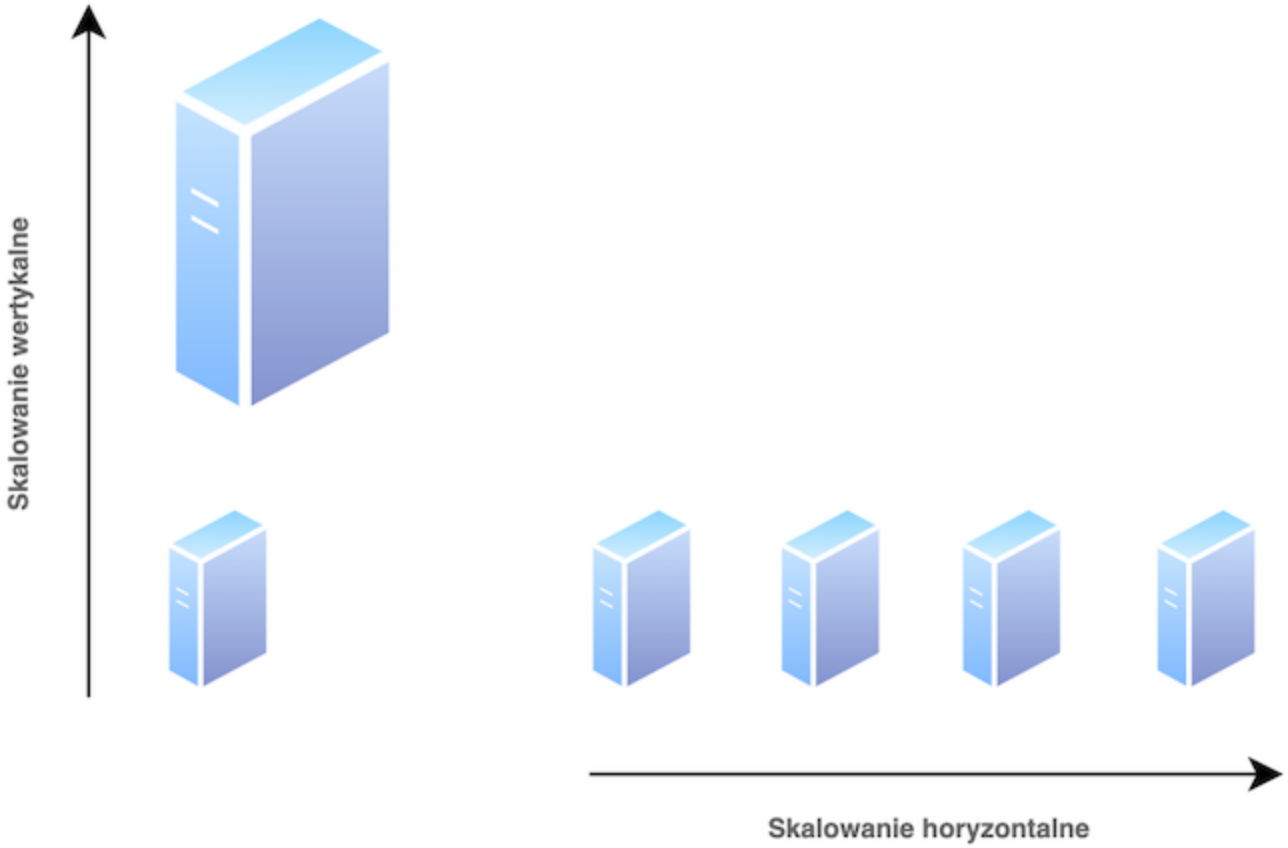
Przykład niewydajnego zapytania:

- Nazwy busów znajdujących się w prostokącie pomiędzy P1(40, -73), P2(41, -74)

Bazy grafowe



Skalowanie wertykalne (w górę), a skalowanie horyzontalne (wszerz)



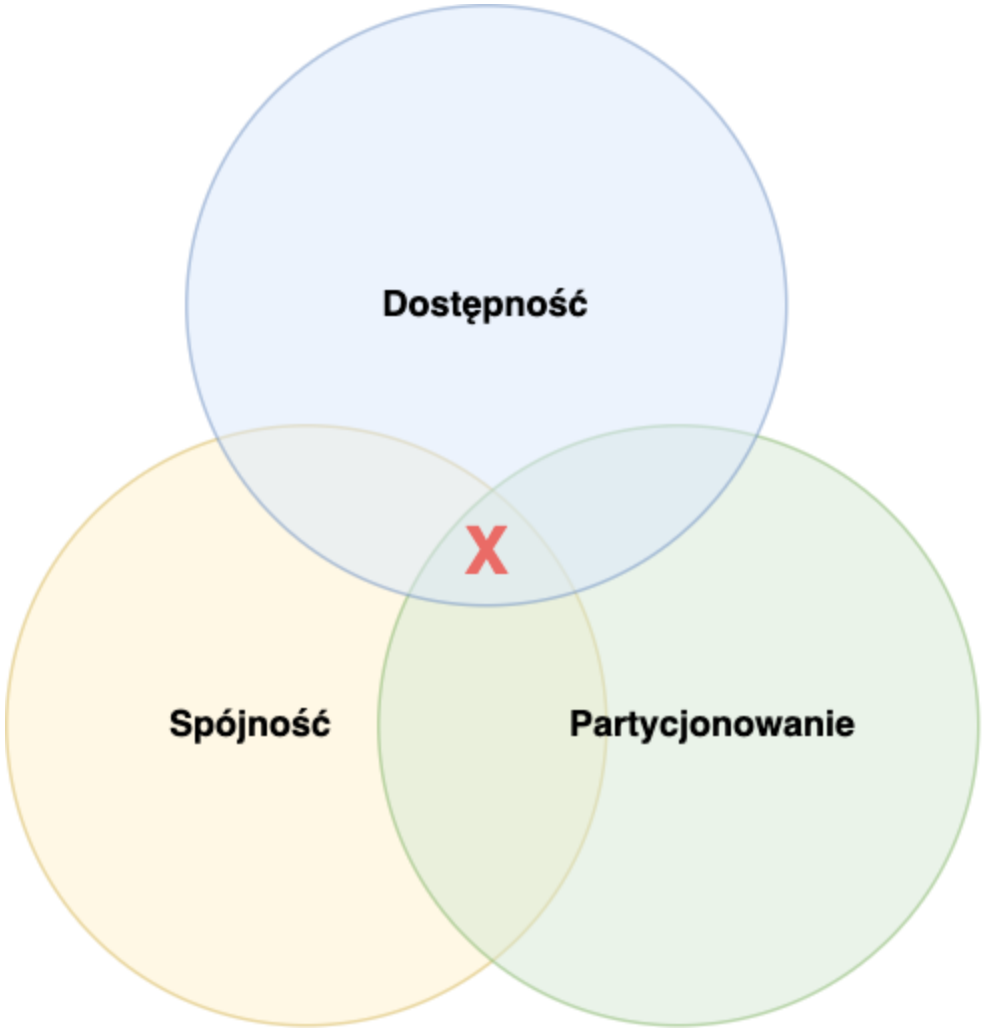
ACID - przypomnienie

- (A)tomicity - Niepodzielność
- (C)onsistency - Spójność
- (I)solation - Izolacja
- (D)urability - Trwałość

Twierdzenie CAP - 1

- (C)onsistency - każda część rozproszonego systemu zwraca dane w najnowszej wersji.
- (A)vailability - system rozproszony zwróci dane nawet pomimo problemu z jego częścią, jednak nie ma gwarancji, że będą
  one w najnowszej wersji.
- (P)artition-tolerance - system rozproszony działa poprawnie pomimo problemów z jego częścią.

Twierdzenie CAP - 2



Zadanie

Naszym zadaniem jest utworzenie bazy danych NoSQL MongoDB na platformie https://mlab.com załadowanie do niej danych oraz napisanie kilku prostych zapytań.