

**Nazwa Wydziału**

**Nazwa Katedry**

**Katedra Inżynierii Oprogramowania**

Zarządzanie projektami (IM)

**Piotr Wrona**

Nr albumu s25773

**Tytuł pracy dyplomowej**

“Procedura bezpieczeństwa dla ciągłej integracji oprogramowania oparta na DevSecOps.”

Rodzaj pracy

magisterska

Imię i nazwisko promotora

Dr. Hab. Inż. Piotr Habela

miejsce, miesiąc, rok obrony

Spis Treści

[Wstęp 5](#__RefHeading___Toc425_451261914)

[Cel pracy 5](#__RefHeading___Toc427_451261914)

[Założenia 5](#__RefHeading___Toc427_451261914_Copy_1)

[Wprowadzenie teoretyczne do zagadnienia 7](#__RefHeading___Toc437_451261914)

[Geneza 7](#__RefHeading___Toc138_3912164325)

[Cele DevSecOps 7](#__RefHeading___Toc140_3912164325)

[Manifest DevSecOps 7](#__RefHeading___Toc142_3912164325)

[Podejścia do DevSecOps 9](#__RefHeading___Toc144_3912164325)

[Rozwiązanie projektowe 10](#__RefHeading___Toc429_451261914)

[Rozwiązanie wejściowe 10](#__RefHeading___Toc149_3912164325)

[Wybór struktury aplikacji 10](#__RefHeading___Toc733_3420670368)

[Wybór środowiska aplikacji 11](#__RefHeading___Toc735_3420670368)

[Wybór bazy danych 11](#__RefHeading___Toc737_3420670368)

[Wybór modelu i architektury dla projektu 12](#__RefHeading___Toc739_3420670368)

[Modele infrastrukturalne 12](#__RefHeading___Toc431_451261914)

[Propozycja infrastruktury 13](#__RefHeading___Toc741_3420670368)

[Wejściowy proces CI/CD dla projektu 15](#__RefHeading___Toc435_451261914)

[Cel projektowania i stosowania CI/CD 15](#__RefHeading___Toc57_1453835382)

[Propozycja procesu CI/CD dla projektu 16](#__RefHeading___Toc743_3420670368)

[Rozwiązanie wyjściowe 19](#__RefHeading___Toc151_3912164325)

[Planowanie bezpieczeństwa 19](#__RefHeading___Toc154_3912164325_Copy_1)

[Bezpieczeństwo infrastruktury i implementacja rozwiązania 19](#__RefHeading___Toc154_3912164325)

[Bezpieczeństwo infrastruktury cloudowej 20](#__RefHeading___Toc146_3912164325)

[Skanowanie infrastruktury 20](#__RefHeading___Toc156_3912164325)

[Testy penetracyjne i implementacja zabezpieczeń 20](#__RefHeading___Toc158_3912164325)

[Bezpieczeństwo aplikacji i implementacja i implementacja rozwiązania 20](#__RefHeading___Toc160_3912164325)

[Skanowanie kodu źródłowego 20](#__RefHeading___Toc162_3912164325)

[Podsumowanie 21](#__RefHeading___Toc425_451261914_Copy_1)

[Wnioski 21](#__RefHeading___Toc169_3912164325)

# Wstęp

## Cel pracy

Celem pracy jest analiza procesu ciągłego dostarczania i eksploatacji oprogramowania, a następnie wyszczególnienie miejsc podatnych z punktu widzenia jego bezpieczeństwa. Zaproponowane zostaną rozwiązania na redukcję wrażliwości i docelowo zabezpieczenie kodu źródłowego. W niektórych sytuacjach przedstawione zostanie porównanie narzędzi, z wyszczególnieniem wad i zalet.

Opracowanie ma pozwolić na jak najgłębsze przedstawienie zagadnienia oraz porządkuję zagadnienie DevSecOps. Dzięki temu, praca będzie przydatna dla osób zajmujących się procesami ciągłego dostarczania i eksploatacji oprogramowania jako uzupełnienie wiedzy, ale również będzie niezwykle przydatna dla osób poznających tematykę i chcących zaczerpnąć bazową wiedzę z zakresu bezpieczeństwa przy procesach CI/CD.

## Założenia

## Wprowadzenie teoretyczne do zagadnienia

### Geneza

### Cele DevSecOps

### Manifest DevSecOps

W celu poprawnego zrozumienia badanego zagadnienia konieczne jest przybliżenie terminu DevSecOps. Został on podobnie jak koncept podejścia Agile zebrany w zbiór reguł, które tworzą menifest[[1]](#footnote-2). W tym przypadku jego twórcami są:

* Ian Allison
* Justin Tiplitsky
* Scott Kennedy
* Nigel Kersten
* Shannon Lietz
* Fabian Lim
* Michelle Nikulshin
* Christian Price
* Ravi Dhungel
* Kyle Rose
* Brandon Sherman

Ropatrzmy poszczególne tezy stawiane przez manifest:

**„Leaning in** over Always Saying “No””

Przez to hasło należy rozumieć pewną transformację podejściową członków zespołu z częstego mówienia „nie” do świadomego wkładu w rozwijanie kwestii bezpieczeństwa. Temat ten nie może być odkładany jako nieistotny, a kwestie związane z inicjatywami w tej materii muszą być wspierane.

**„Data & Security Science** over Fear, Uncertainty and Doubt”

DevSecOps musi być oparty na profesjonalnym podejściu, wspomaganym przez rzetelne raporty wykorzystywanych narzędzi niż podejściu chaotycznym i kierowanym przez strach.

„**Open Contribution & Collaboration** over Security-Only Requirements”

Tworzenie oprogramowania opiera się na integracji wielu produkty zewnętrznych oraz wewnętrznych. Dlatego, konieczna jest świadomość wszystkich elementów oraz jak je zintegrować aby stworzyć bezpieczny system.

„**Consumable Security Services with APIs** over Mandated Security Controls & Paperwork”

Priorytem powinno być stworzenie procesu który możemy wyrażać jako pewnego rodzaju usługę zapewniania bezpieczeństwa zamiast skupiać się na narzędziach służących wykonywaniu danej funkcjonalności.

„**Business Driven Security Scores** over Rubber Stamp Security”

Konieczne jest dostosowanie naszych wymogów dotyczących bezpieczeństwa do wymagań biznesu. Każda implementacja powinna być oceniana pod względem jej użyteczności, zamiast zwykłego ohczenia działania jako zrobione.

**Red & Blue Team Exploit Testing** over Relying on Scans & Theoretical Vulnerabilities

Tworzenie zespółów czerwony/niebiescy do nauki obrony naszego systemu. W zalożeniu podejście to odwzorowuję bardziej realistyczne przebieg potencjalnego ataku przy użyciu najbardziej aktualnych narzędzi penetracyjnych.

**24x7 Proactive Security Monitoring** over Reacting after being Informed of an Incident

Ciągłe skanowanie naszego systemu pozwala nam wykrywać zdarzenia potencjalnie niebezpieczne i przeciwdziałać im przed szkodą.

**Shared Threat Intelligence** over Keeping Info to Ourselves

Ciągła wymiana informacji jest kluczem do rozwijania świadomości i poszerzania wiedzy wszystkich członków zespołu. Takie działania pozwalają na szybszą reakcję i lepszą odpowiedź w przypadku sytuacji zagrożenia.

**Compliance Operations** over Clipboards & Checklists

Ciągła kontrola ponad ufność w checklisty.

### Podejścia do DevSecOps

# Rozwiązanie projektowe

## Rozwiązanie wejściowe

### Wybór struktury aplikacji

W ostatnich latach, trend rozwijania aplikacji sukcesywnie przesuwa się z dużych monolitycznych aplikacji na aplikacje mikro serwisowe. Tabela poniżej opisuję główne różnice dla obu podejść.

Rozpatrując przedstawione porównanie, łatwo można dostrzec przeważające zalety podejścia mikro serwisowego. Z tego powodu, projekt przewiduję zastosowanie konteneryzacji zarówno dla serwera www oraz serwera baz danych.

W uproszczeniu konteneryzacja jest wydzieleniem małego środowiska uruchomieniowego, z izolacją zasobów, konfiguracji czy udostępnieniem osobnego interfejsu sieciowego. Kontener jest budowany na podstawie pliku „Dockerfile”, który zawiera predefiniowany obraz docker’a i uruchamia kod źródłowy na tak przygotowanym środowisku.

#### Wybór środowiska aplikacji

Projekt przewiduje użycie aplikacji node.js jako środowiska uruchomieniowego. Wyróżnia się ono następującymi cechami:

* wieloplatformowa
* bardzo duża szybkość obsługi zapytań
* łatwość tworzenia zaawansowanych programów
* duża liczba framework’ów przyspieszających pisanie aplikację

Node.js wykorzystuję składnie języka JavaScript, który cechuję się elastycznością (możliwość zastosowania programowania funkcyjnego czy obiektowego) oraz dużą czytelnością (jest to język wysokopoziomowy).

#### Wybór bazy danych

Aplikacja wykorzystuję bazę danych do przechowywania informacji o video, udostępnianym przez serwer www. Dla przytoczonego zastosowania, nie jest wymaga złożona relacyjność, a optymalne jest zestawienie klucz wartość. Wybór bazy padł na nierelacyjną bazę mongodb.

Baza cechuję się:

* dużą liczbą obsługiwanych typów danych
* obsługą kursorów
* zapytaniami ad-hoc
* zapytaniami do zagnieżdżonych pól dokumentów
* indeksowaniem
* wsparciem dla agregacji danych
* możliwością składowania plików w bazie
* architekturą zaprojektowana z myślą o łatwej replikacji

### Wybór modelu i architektury dla projektu

#### Modele infrastrukturalne

Główny podział infrastruktury współczesnych aplikacji opiera się na lokalizacji serwerów i dysponowaniem zakresem obowiązków dotyczących obsługi. Wyróżniamy:

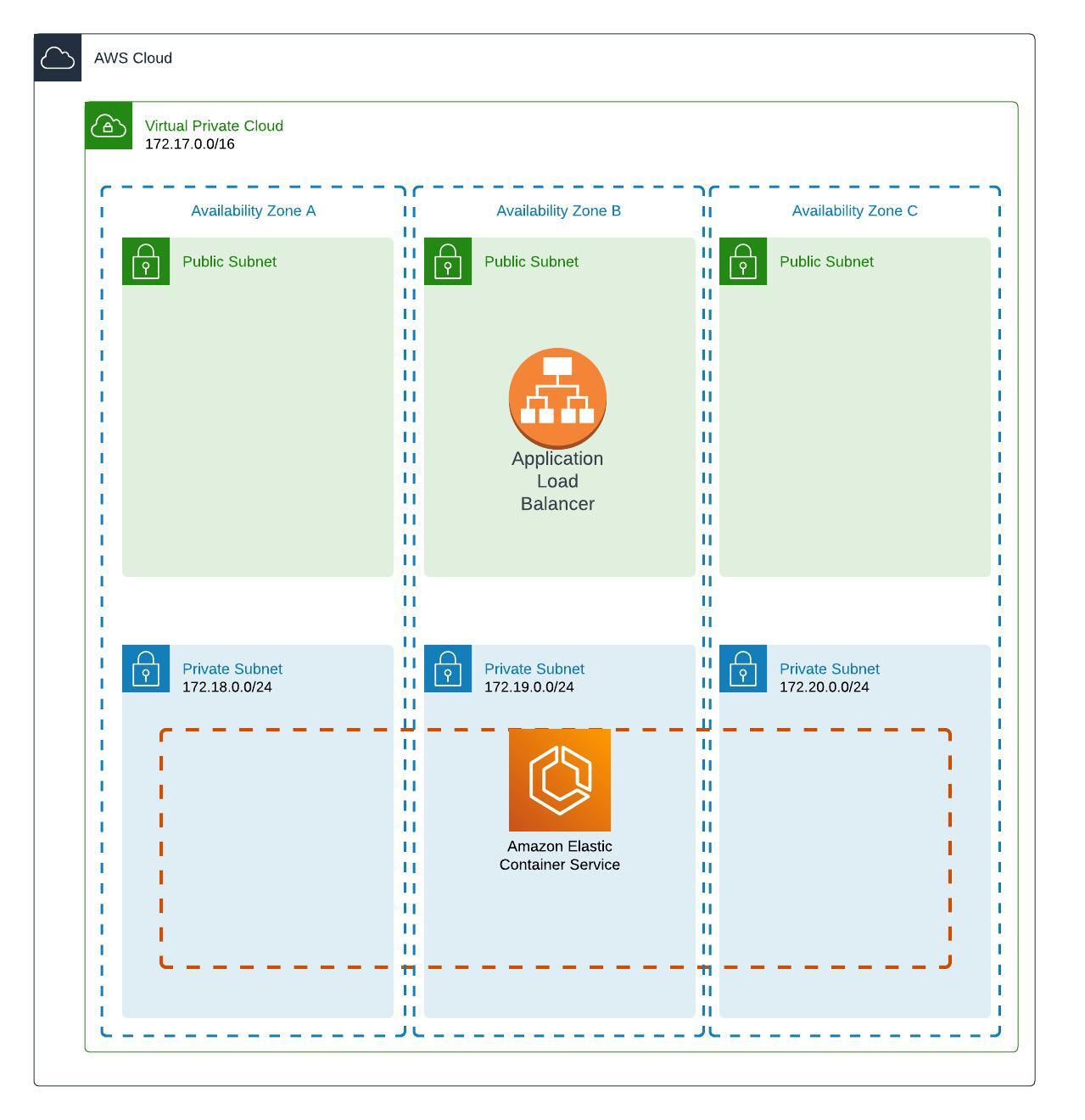
* On-premises – infrastruktura zlokalizowana w przedsiębiorstwie. Pełna obsługa hardware’u przypada na zatrudnionych administratorów sieci.
* Cloud – w tym modelu przesuwamy odpowiedzialność za obsługę sprzętu na zewnętrznych dostawców. Ze względu na zakres obowiązków, w rozwiązaniach chmurowych możemy wymienić:
  + SaaS (Software as a Service) – model zakłada kompletne przeniesienie odpowiedzialności za sprzęt, infrastrukturę oraz aplikację na rzecz dostawcy.
  + PaaS (Platform as a Service) – polega na udostępnieniu przez dostawce gotowej infrastruktury oraz sprzętu do budowy aplikacji.
  + IaaS (Infrastructure as a Service) – w tym modelu zarówno budowa infrastruktury oraz aplikacji leży po stronie użytkownika końcowego. Dostawca odpowiada tylko za sprzęt.

W praktyce zdarza się, że rolę się przeplatają a obok modelu IaaS stosujemy rozwiązania SaaS. Jednym z częstych przypadków jest stosowanie usług Elasticsearch do monitorowania naszej infrastruktury Iaas.

#### Propozycja infrastruktury

Na potrzeby rozwiązania projektowego, użyta została chmura AWS w modelu IaaS. Występują też pewne elementy zbliżone do modelu PaaS, takie jak zastowanie AWS ECS, w podejściu fargate.

Głównym logicznym podziałem naszej infrastruktury jest VPC (Virtual Private Cloud), który definiuję naszą sieć i zakres adresów. Architektura jest 2-tierowa i wyróżnia 3 podsieci publiczne oraz 3 podsieci prywatne. Cała sieć znajduję się w jednym regionie eu-central-1 (Frankfurt), co w pewien sposób powoduję podatność na awarię w obrębie regionu, natomiast pewną redundantność uzyskujemy poprzez wydzielenie 3 AZ (Avaliability zone): 1A, 1B, 1C.

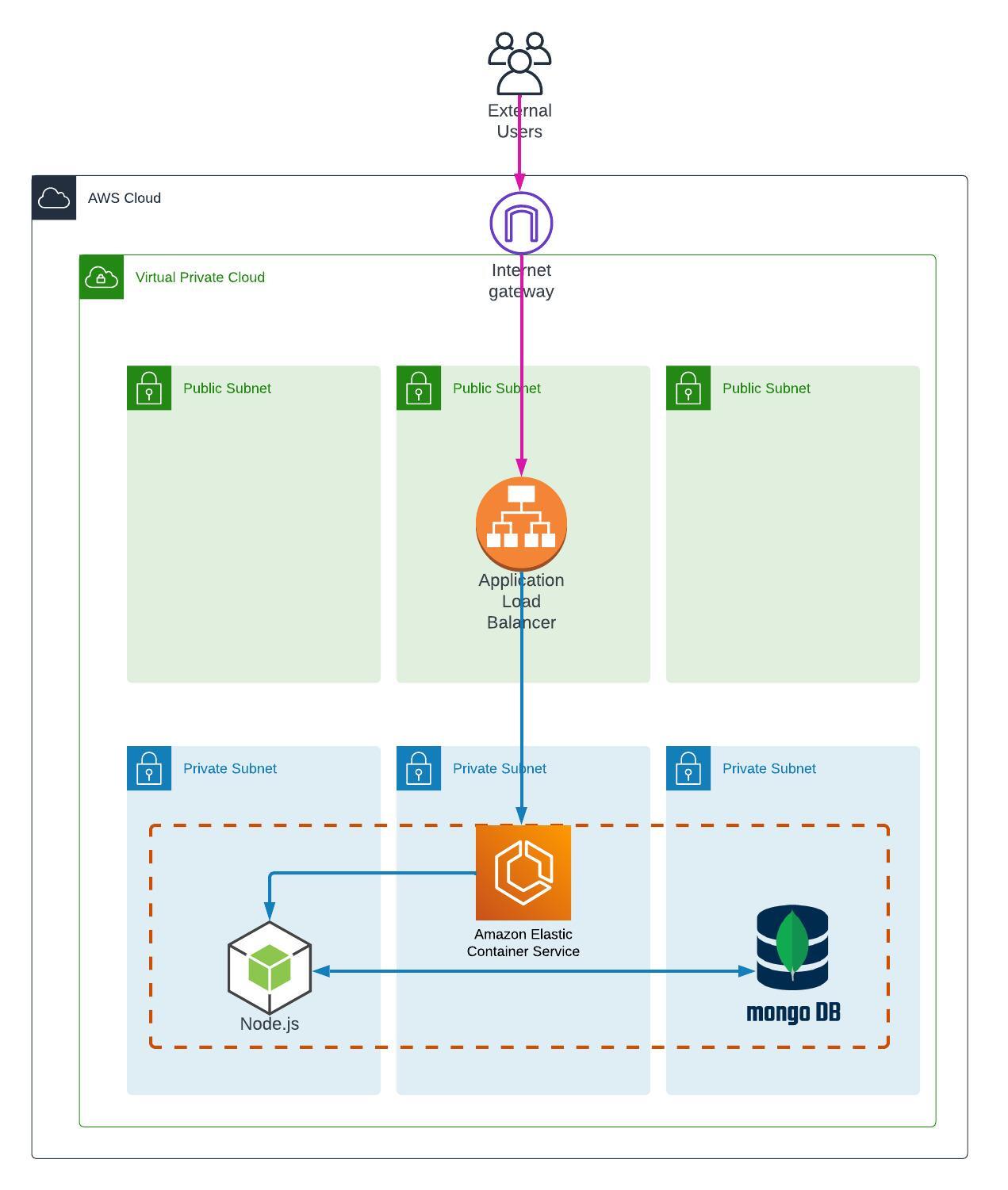


Każda podsieć publiczna ma maskę 255.255.255.0 która obejmuję 254 adresy. Dla naszego rozwiązania projektowe zakres możliwyhch destynacji jest nieproporocjonalnie duży, natomiast umożliwa to ewentualne dalsze rozwijanie infrastruktury czy mozliwości skalowania.

Wejściem do naszego systemu jest IG (Internet Gateway) który przyjmuję cały ruch publiczny i kieruje go na ALB (Application Load Balancer). Odpowiada on za dystrubucję ruchu, terminację certyfikatu SSL oraz częściową obsługę błędów (np.: poprzez odpowiedź na błąd 404).

Kontenery aplikacji znajdują się w podsieciach prywatnych i są zarządzane przez serwis ECS (Elastic Container Service) w modelu Fargate. Podejście to cechuje:

* brak konieczności administracji serwerów
* mikroserwisowość
* łatwość skalowania
* skalowalność horyzontalna
* duża dostępność poprzez dystrubucję pomiędzy 3 zony



### Wejściowy proces CI/CD dla projektu

#### Cel projektowania i stosowania CI/CD

Trend tworzenia aplikacji na przestrzeni ostatnich lat zmienił się z typowego podejścia waterfall do podejścia agile. Podejście wywodzi się od koncepcji dostarczania MVP czyli produktu o małej wartości, w krótszych odstępach. Pozwala to na szybsze zebranie opinii klienta końcowego i zabezpiecza przedsiębiorstwo przed tworzeniem funkcjonalności niedostarczających wartości dodanej.

Zmiana podejścia zrodziła również wiele wyzwań dla osób odpowiadających za cały strumień wartości i dostawę oprogramowania. Możemy wyróżnić kilka z nich:

* Integracja wielu pionów tj.: deweloperów, testerów, działu bezpieczeństwa informacji oraz biznesu
* Stworzenie oraz zarządzanie wieloma środowiskami, które muszą odpowiadać zarówno celom testerów oprogramowania czy działom operacji
* Wprowadzanie i dbanie o elementy bezpieczeństwa pomimo biznesowej presji dostarczania oprogramowania

Z tych wielu powodów konieczny jest prawidłowe zaplanowanie procesu CI/CD dla naszej aplikacji. CI czyli ciągła integracja to wszystkie czynności integrujące nasze repozytoria, aż do momentu decyzji biznesowej, która następnie inicjuję proces ciaŋłego wzdrażania, czyli CD.

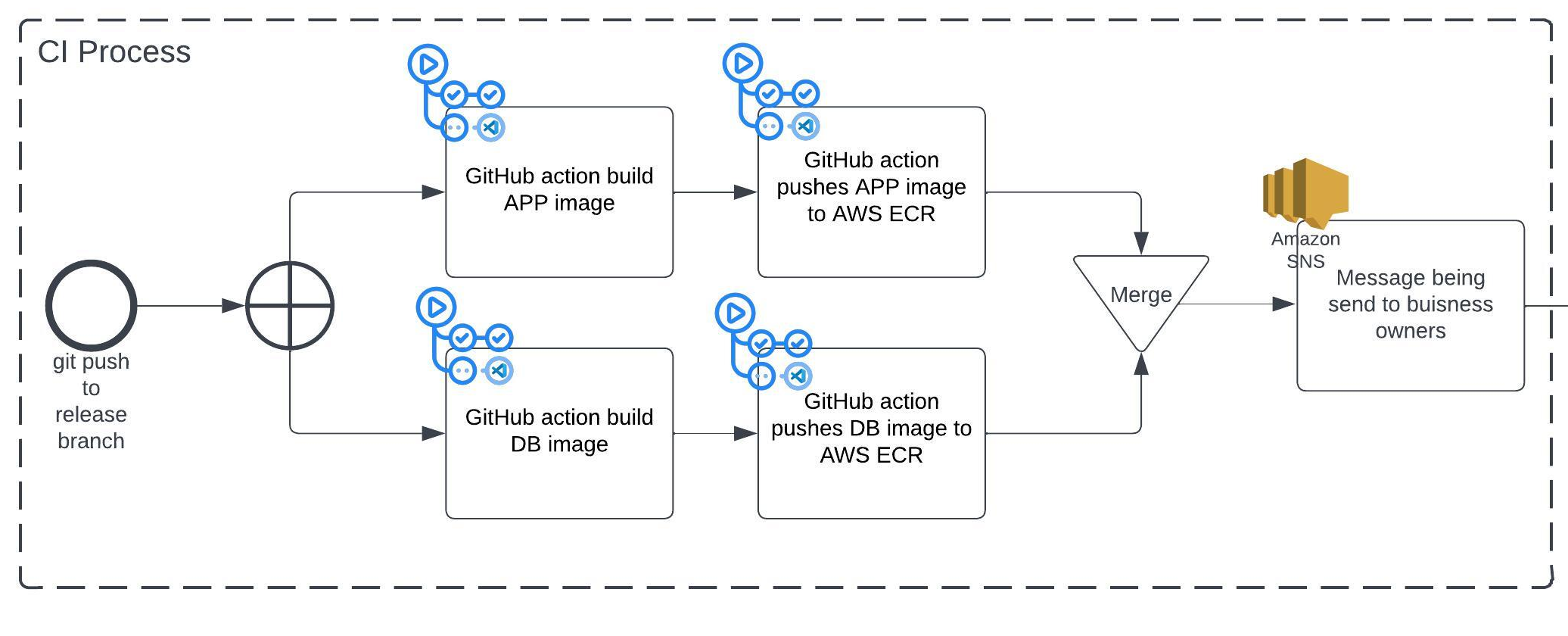
Podejście do schematu dostarczania oprogramowania jest bardzo róźne i zależy głównie od:

* Technologii wdrażanej aplikacji
* Poziomu unifikacji systemów i serwerów
* Strategii biznesowej firmy
* Ilości środowisk testowych i produkcyjnych
* Doświadczenia i testowania rozwiązań

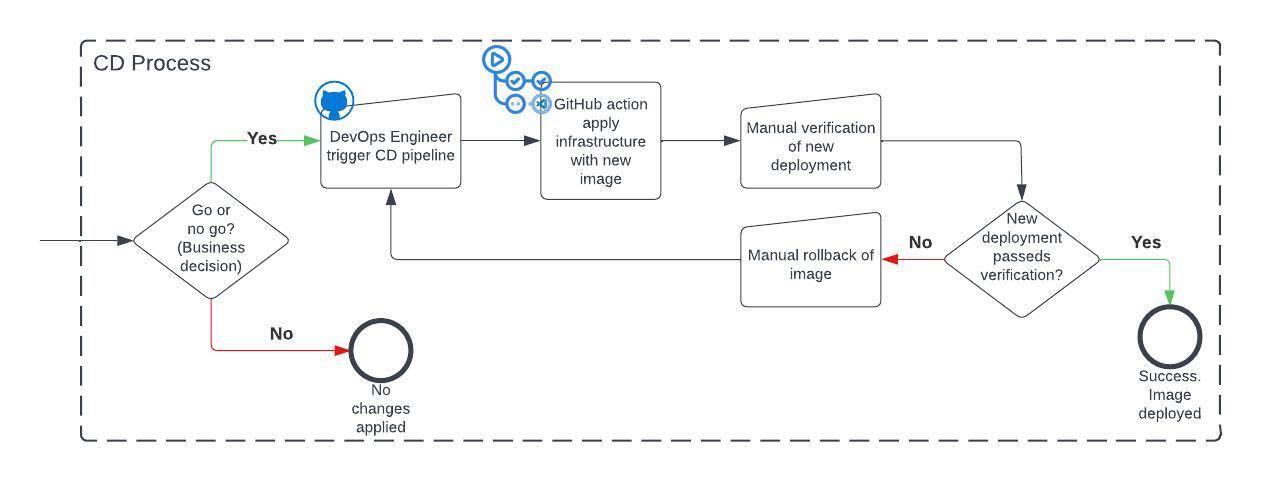
#### Propozycja procesu CI/CD dla projektu

Projekt zakłada jedno środowisko testowe (TEST ENV) oraz jedno środowisko produkcyjne (PROD ENV). Należy nadmienić, że podejście jest uproszczone w stosunku do najczęstszych wzorów komercyjnych. Dla rozwiązań rynkowych przewiduję się zazwyczaj dwa środowiska testowe (np.: QA (Quality Assurance) oraz UAT (User Acceptance Tests)), jak również dwa środowiska produkcyjne, głównie dla celów przełączania w czasie awarii.

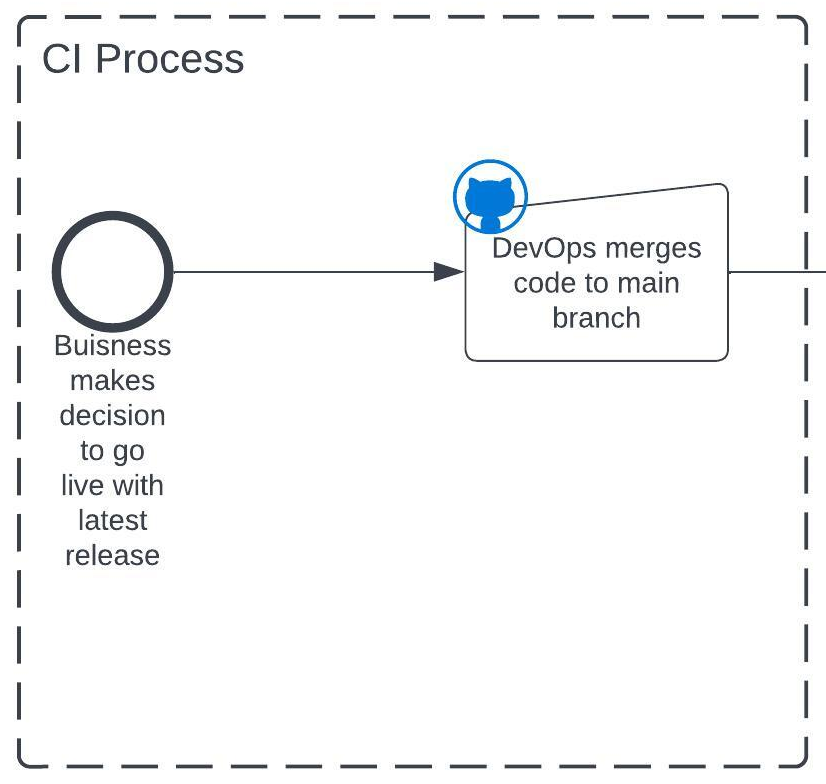
Podejście konteneryzacyjne od aplikacji, powoduję pewną centralizację procesu wokół repozytorium obrazów dockerowych. Wpływa to zarówno część integracyjną - obok udostępnia kodu, następuję udostępnienie obrazu, jak również część wdrożeniową – poprzez pobranie najnowszej wersji obrazu i zaaplikowanie na środowisku.



1. Proces rozpoczyna się od wypuszczenia kodu na tzw.: „release branch”. Zaletami tworzenia dedykowanych gałęzi git dla dannych sprintów są:
   * Lepsza kontrola nad zmianami i możliwość prowadzenia dziennika zmian
   * Możliwość łatwego rollback’u do stabilnej wersji
2. Zmiana w repozytorium inicjuje równoległe przetwarzania dwóch pipelinów – jeden do tworzenia obrazu aplikacji oraz drugi do tworzenia kontenera bazy danych.
3. Każdy z nowo utworzonych obrazów zostaję otagowany mianem „latest” i przekazany do uprzednio utworzonych repozytoriów obrazów na chmurze AWS.
4. Zakończenie procesu ładowania nowych wersji aplikacji oraz bazy jest finalizowane poprzez przekazanie do kolejki Amazon SNS, notyfikacji mailowej dla biznesu.

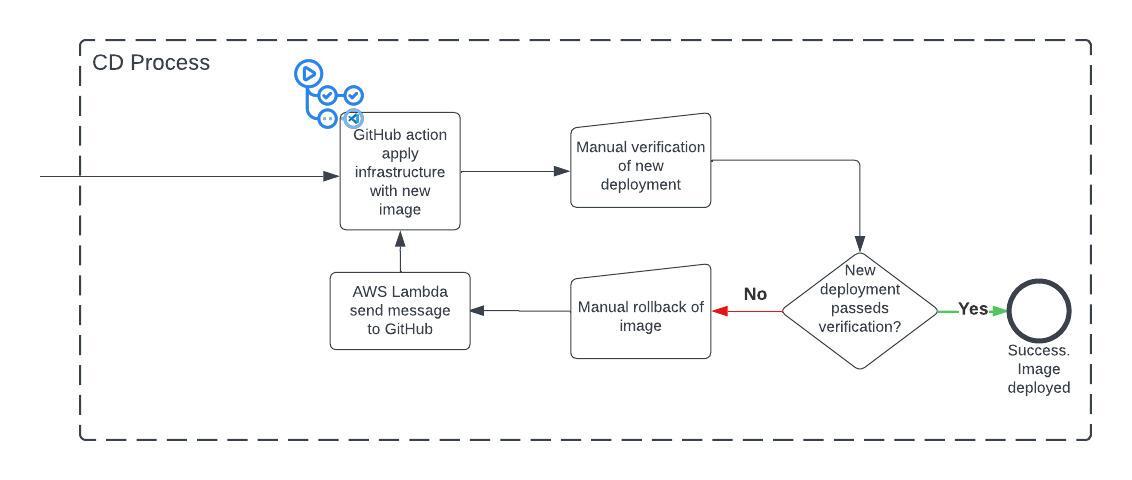


1. Proces wdrożeniowy jest inicjowany poprzez decyzję biznesową. Na tym etapie występuję możliwość rezygnacji z nowego wydania. Taka akcja nie wymaga dodatkowych kroków, gdyż rollback następuję przy kolejnej budowie obrazów.
2. Automat aplikuję nowe rozwiązanie przy użyciu narzędzia terraform.
3. Inżynier DevOps przeprowadza weryfikację środowiska pod kątem stabilności i kompletności. Rekomendowane są w tym miejscu testy automatyczne wraz z obserwacją i finalizacja przez decyzję manualną.
4. W przypadku istotnych problemów i braku akceptacji inżyniera na emisję wydania, następuję rollback obrazu w repozytorium ECR.
5. Podobnie do poprzednich kroków z sekwencji, to inżynier DevOps inicucję aplikację poprzedniej wersji obrazów oraz weryfikację po procesie. Ze względu na działanie w sytuacji DR (disaster recovery) dla środowiska testowego, zbędna jest w tej sytuacji decyzja biznesowa.
6. W przypadku akceptacji nowego wydania, zmiana jest zamykana jako sukces.



Z perspektywy środowiska produkcyjnego, cześć integracyjna procesu ma zuważalnie mniejszy stopień złożoności.

1. Inicjacja sekwencji rozpoczyna się od decyzji biznesowej, która jest podejmowana zgodnie z harmonogramem wydań oprogramowania.
2. Inżynier DevOps na podstawie zarządzenia, łączy gałąź „release” z „main”, co inicjuję część wdrożeniową.

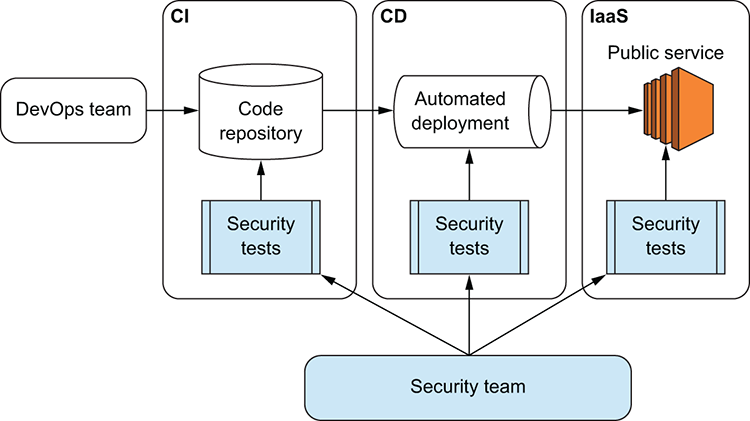


1. Część wdrożeniowa polega na automatycznej inicjacji składni yaml która podmienia obraz repozytorium produkcyjnego na wersję przetestowaną na środowisku testowym.
2. Inżynier DevOps przeprowadza weryfikację środowiska pod kątem stabilności i kompletności. Rekomendowane są w tym miejscu testy automatyczne wraz z obserwacją i finalizacja przez decyzję manualną.
3. W przypadku istotnych problemów i braku akceptacji inżyniera na emisję wydania, następuję rollback obrazu w repozytorium ECR.
4. AWS Lambda wykrywa zmianę w repozytorium kodu i wysyła sygnał do GitHub w celu rozpoczęcia emisji cofniętego obrazu dockerowego.
5. Inżynier DevOps przeprowadza weryfikację środowiska pod kątem stabilności i kompletności. Rekomendowane są w tym miejscu testy automatyczne wraz z obserwacją i finalizacja przez decyzję manualną.
6. W przypadku akceptacji nowego wydania, zmiana jest zamykana jako sukces.

## Rozwiązanie wyjściowe

### Planowanie bezpieczeństwa

#### Zarządzanie sekretami



### Bezpieczeństwo infrastruktury i implementacja rozwiązania

#### Bezpieczeństwo infrastruktury cloudowej

1. Podział infastruktury na prywatne i publiczne podsieci pozwala na lepszą kontrolę dostępu do backendu. Wszelki ruch wejściowy powinien być w naszej infrastrukturze poprzez 443 na ALB.
2. Dopiero w tym miejscu wykonujemy tzw. SSL termination co oznacza po prostu dekrypcję transferu i następnie jest przekazywany do prywatnych podsieci
3. .

#### Skanowanie infrastruktury

#### Testy penetracyjne i implementacja zabezpieczeń

### Bezpieczeństwo aplikacji i implementacja i implementacja rozwiązania

#### Skanowanie kodu źródłowego

# Podsumowanie

## Wnioski

Table of Figures

1. https://www.devsecops.org/ [↑](#footnote-ref-2)