

Dokumentacja projektu RSO

Grupa C (godzina 16.00)

Domagała Bartosz, Kornata Jarosław, Modzelewski Jędrzej, Marcin Stepnowski

Usługa bezpiecznej niezawodnej dystrybucji przetworzonej chronionej informacji

Prowadzący projekt:
dr inż. Tomasz Jordan Kruk

Spis treści

1. Opis projektu	1
1.1. Treść zadania	1
1.2. Wstępny opis rozwiązania	2
1.3. Wstępny projekt architektury	2
Warstwa wewnętrzna serwera	2
Warstwa zewnętrzna serwera (warstwa przetwarzająca)	2
Aplikacja kliencka	3
1.4. Opis szczegółowy	4
1.4.1. Wymagania funkcjonalne	4
1.4.2. Wymagania нефункционалне	7
1.4.3. Dodatkowe założenia	8
1.5. Potencjalne problematyczne scenariusze	10
1.6. Organizacja środowiska programistycznego	11
1.6.1. Język programowania Java	11
1.6.2. IntelliJ IDE	11
2. Organizacja projektu	12
2.1. Podział zadań	12
2.2. Planowany harmonogram pracy	12
2.3. Sprawozdania ze spotkań	14
2.3.1. Spotkanie I - 31 marca 2015	15
2.3.2. Spotkanie II - 8 kwietnia 2015	15
2.3.3. Spotkanie X1 - 14 kwietnia 2015	16
2.3.4. Spotkanie III - 14 kwietnia 2015	17
2.3.5. Spotkanie X2 - 21 kwietnia 2015	17
2.3.6. Spotkanie IV (TeamSpeak) - 22 kwietnia 2015	18
2.3.7. Spotkanie V - 27 kwietnia 2015	18
2.3.8. Spotkanie X3 - 28 kwietnia 2015	19
2.3.9. Spotkanie V (TeamSpeak) - 28 kwietnia 2015	19
3. Narzędzia i zewnętrzne biblioteki	20
3.1. Docker	20
3.1.1. Czym jest docker?	20
3.1.2. Spójność	20
3.1.3. Cykl Dockera	20
3.1.4. Centralne repozytorium	21
3.1.5. Przeprowadzone testy	21
3.2. Google Protocol Buffers	22
3.3. MySQL	23

1. Opis projektu

1.1. Treść zadania

Zaprojektować i zaimplementować usługę udostępniania danych statystycznych na temat studentów uczelni wyższej, przy założeniu bezpiecznej i niezawodnej dystrybucji danych wrażliwych - poufnych informacji o studentach.

Usługa powinna:

- mieć dokładnie jeden ten sam tekstowy plik konfiguracyjny dla części serwerowej i klientów usługi,
- składać się po stronie serwerowej z dwóch warstw: (1) wewnętrznej wytwarzającej, przechowującej i przesyłającej przetworzoną wrażliwą informację, (2) warstwy zewnętrznej do udostępniania przetworzonej informacji oprogramowaniu klienckiemu usługi,
- oprogramowanie realizujące część serwerową powinno być współbieżne,
- składać się po stronie klienckiej z prostych narzędzi opatulających zaproponowane API protokołu komunikacyjnego,
- zapewniać odporność na uszkodzenia węzła, stąd każda z warstw powinna być uruchomiona na co najmniej dwóch węzłach,
- zapewniać realizację usługi w trakcie awarii w czasie obsługi,
- obsługiwać scenariusz próby ponownego wpięcia się przez węzeł dowolnej z warstw, z którym wcześniej utracono łączność, a zarazem być odporną na próbę zastąpienia nieosiągalnego sieciowo węzła (np. w wyniku ataku DoS, odmowa usługi) węzłem wrogim do tego nieuprawnionym,
- na bieżąco i transparentnie dla oprogramowania klienckiego zarządzać dostępnymi zasobami lokalizacyjnymi,
- zapewniać zadany (większy niż 1 ale dopuszczalny konfiguracją mniejszy niż liczba serwerów danych) poziom redundancji - z obsługą uzupełniania kopii na innych serwerach w przypadku awarii włącznie,
- cała część serwerowa usługi powinna być uruchamiana i zamykana jednokrotnym wywołaniem skryptu na jednym węźle serwera z jednym argumentem <start|stop|status> (wykorzystanie nieinteraktywnych metod automatycznego uwierzytelniania węzłów w SSH).

1.2. Wstępny opis rozwiązania

Nasze rozwiązanie to rozproszony system do rejestracji na przedmioty studentów uczelni wyższej. Warstwą wewnętrzną przechowującą dane wrażliwe, w tym przypadku będą węzły rejestrujące studentów, ideowo umieszczone w dziekanatach instytutów i obsługiwane przez osoby tam pracujące. Warstwa pośrednia przechowuje informacje o studentach, jednak bez informacji wrażliwej. Posiada jedynie dane o istnieniu studenta zapisanego na przedmioty RSO, SOI, ANA1 i SR, bez imienia, nazwiska i daty urodzenia. Owe informacje są przetwarzane na żądanie aplikacji klienckiej, dostępnej dla studentów, którzy chcą dowiedzieć się, ile osób zapisanych jest na dany przedmiot.

1.3. Wstępny projekt architektury

Wewnętrzna warstwa serwera wytwarza i przechowuje informacje na temat studentów uczelni wyższej - stworzone na podstawie danych wprowadzonych przez pracowników dziekanatu. Następnie owa warstwa przesyła informacje do warstwy zewnętrznej serwera (warstwy przetwarzającej) świadczącej usługi udostępniania informacji przetworzonej rozwiązaniom klienckim. Każda z warstw komunikować się będzie z użyciem protokołu TCP/IP przesyłając wiadomości zserializowane za pomocą Google Protocols Buffer.

Warstwa wewnętrzna serwera

Warstwa stworzona w technologii token ring. Węzły połączone są w pierścień i przesyłają sobie kolejno token zawierający informacje o potencjalnych zmianach. Tylko węzeł posiadający token może wykonywać operację aktualizacji danych. Każda zmiana musi być zatwierdzona przez pozostałe węzły. Zapewnia to synchronizację danych oraz bieżące sprawdzenie działania pozostałych węzłów.

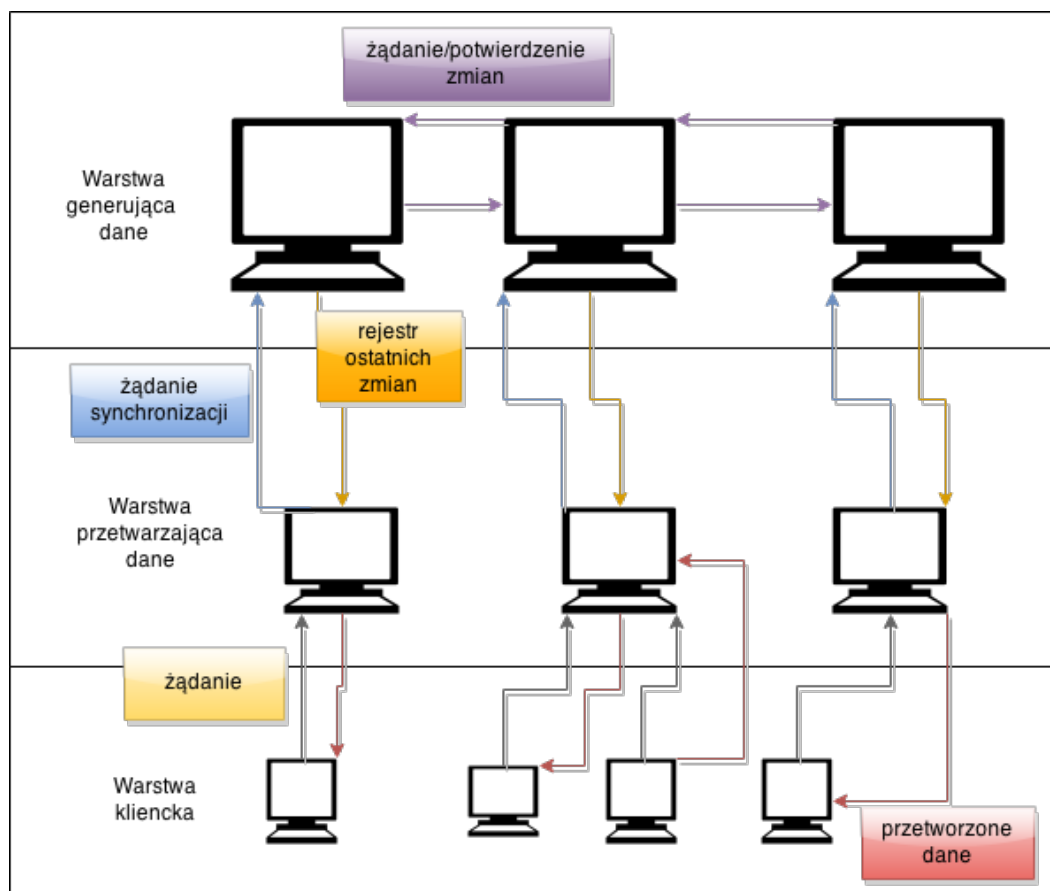
Dane wrażliwe przechowywane są w postaci rekordów zawierających:

- ID studenta
- imię studenta
- nazwisko studenta
- datę urodzenia
- listę przedmiotów na jakie jest zapisany

Warstwa zewnętrzna serwera (warstwa przetwarzająca)

Warstwa przeznaczona do komunikacji z aplikacją kliencką. Na żądanie klienta przetwarza wcześniej otrzymane od warstwy wewnętrznej dane i wynik przetwarzania zwraca klientowi. Moduł ten synchronizuje dane okresowo (częstość jest ustalana w pliku konfiguracyjnym), tak aby administrator sam był w stanie dostosować interwał do aktualnych potrzeb).

Dane o studentach przechowywane są w poniższej postaci:



Rysunek 1.1. Diagram ilustrujący komunikację pomiędzy poszczególnymi warstwami

- ID studenta
- listę przedmiotów na jakie jest zapisany

Aplikacja kliencka

Warstwa przeznaczona dla klienta to z założenia lekka i intuicyjna aplikacja konsolowa. Moduł ten bezpośrednio łączyć będzie się tylko z modułem zewnętrznym serwera. Użytkownik, za pomocą tej aplikacji będzie mógł dowiedzieć się ile osób w danym momencie jest zapisany na wybranie przez niego przedmiot. Aplikacja będzie działała synchronicznie, tzn. po wysłaniu żądania będzie się zawieszała w oczekiwaniu na odpowiedź.

1.4. Opis szczegółowy

1.4.1. Wymagania funkcjonalne

ID	Wymagania	Powody decyzji
WF1	Plik konfiguracyjny Każdy z węzłów danej warstwy będzie posiadał plik konfiguracyjny o tej samej strukturze. W zależności od typu dane z pliku zostaną odpowiednio skonfigurowane. Plik będzie zawierał: ID węzła, listę serwerów (zależnie od typu - zewnętrznych lub wewnętrznych), dodatkowe dane konfiguracyjne (np. odpowiadających za redundancję danych, okresową synchronizację, itd.)	Takie rozwiązanie jest wymagane w treści zadania. Ponadto ujednolici i ułatwia konfigurację projektu bez potrzeby ponownej kompilacji.
WF2	Zmiana jednego rekordu na raz Jeśli dany węzeł chce dodać/usunąć/zmodyfikować dane o studencie musi posiadać token, który przesyła do kolejnych serwerów na końcu otrzymując potwierdzenie, że każdy z węzłów wprowadził żadaną zmianę.	Ze względu na to, że w sieci warstwy wewnętrznej serwera istnieje jeden token jest sensowne przyjąć założenie, że na raz można wprowadzić jedną modyfikację w bazie. Ponadto pozwala to utrzymać nieduży rozmiar tokena, co sprzyja jego szybszemu przesyłaniu między węzłami. Nawet w przypadku wielu żądań (od każdego węzła jedno) powinno to działać na tyle szybko, że osoby wprowadzające dane nie odczują opóźnień. Dodatkowo wprowadzanie pojedynczych informacji pozwala na szybsze wykrycie awarii węzła serwera oraz zaoszczędza więcej czasu, niż w przypadku wpisywania dużej ilości danych na raz, a dopiero potem sprawdzania czy węzeł działa poprawnie.

ID	Wymagania	Powody decyzji
WF3	<p>Okresowa synchronizacja</p> <p>Synchronizacja z warstwą wewnętrzną następuje okresowo, w danych odstępach czasu, sterowanych za pomocą pliku konfiguracyjnego. Jest ona wykonywana na zarządzanie węzłów warstwy zewnętrznej. Synchronizacja polega na pobraniu jedynie zmienionych rekordów bazy, a nie wszystkich. Ostatnia funkcjonalność zostanie rozwiązana za pomocą, tzw. „znaczników czasu” (<i>ang. timestamp</i>).</p>	<p>Synchronizacja o konkretnych porach pozwala na odpowiednie rozporządzanie pracą systemu. W przypadku rzadkiej modyfikacji danych przez warstwę wewnętrzną serwera synchronizacja może być wykonywana, np. raz na dzień. Z drugiej strony można ustawić synchronizację co 1 sekundę, co spowoduje wrażenie ciągłej synchronizacji (oczywiście nie biorąc pod uwagę potencjalnych opóźnień). Takie rozwiązanie pozwala administratorowi na odpowiednią optymalizację działania systemu. Dzięki aktualizacji jedynie konkretnych danych, a nie bazy w całości jest to stosunkowo szybkie.</p>
WF4	<p>Redundancja danych (warstwa zewnętrzna serwera)</p> <p>W zależności od zmiennej ustawionej w pliku konfiguracyjnym warstwa przetwarzająca może posiadać pełną redundancję bądź nie. Przyjmuje się, że minimalna liczba serwerów posiadających w pełni aktualne dane to 3.</p>	<p>To założenie służy spełnieniu warunków zadania. Możliwość modyfikacji poziomu redundancji pozwala dopasować system do ilości węzłów w warstwie zewnętrznej serwera i do danego zapotrzebowania.</p>
WF5	<p>Udostępnianie usług</p> <p>Warstwa zewnętrzna obsługuje żądania klientów realizując usługę zwracania ilości studentów zapisanych na dany przedmiot. Serwer nasłuchuje czekając na żądania, a następnie (w ramach możliwości) je spełnia.</p>	<p>To założenie służy spełnieniu warunków zadania.</p>
WF6	<p>Obliczanie danych statystycznych</p> <p>Serwery zewnętrzne, zwane też przetwarzającymi, otrzymują dane od warstwy wewnętrznej, jednak zapisują z nich jedynie dane niewrażliwe używając ich do wyliczania ilości studentów zapisanych na dany przedmiot. Do tej operacji pełne informacje o studentach nie są potrzebne.</p>	<p>To założenie służy spełnieniu warunków zadania.</p>

ID	Wymagania	Powody decyzji
WF7	Dostęp do usługi pobrania danych statystycznych Aplikacja kliencka po połączeniu z serwerem zewnętrznym może pobrać informacje na temat ilości studentów zapisanych na dany przedmiot. Użytkownik otrzymuje jedynie nazwę przedmiotu i liczbę zapisanych studentów, bez dostępu do danych wrażliwych. Może być wysyłane jedno żądanie na raz od danego użytkownika. Nie ma funkcji agregacji wiele żądań w jedno.	Jest to dobry przykład funkcjonalności, która zabezpiecza nam dane wrażliwe przed zwyczajnym użytkownikiem, udostępniając jedynie informacje, które mogą być publiczne. Agregacja bądź wysyłanie wielu żądań na raz nie jest konieczne do sprawdzenia poprawnego działania systemu i nie zmienia w żaden sposób wizji jego funkcjonowania.

1.4.2. Wymagania niefunkcjonalne

ID	Wymaganie	Priorytet
WNF1	Czas między awariami Średni czas między awariami systemu powinien wynosić przynajmniej trzy miesiące, przy czym awarię definiujemy jako ciągłą niedostępność usługi przez co najmniej 10 minut bądź konieczność wykorzystania kopii zapasowej	bardzo wysoki
WNF2	Spójność danych po awarii Odzyskane dane nie mogą być starsze niż 6 godzin przed awarią.	bardzo wysoki
WNF4	Ochrona danych System powinien gwarantować ochronę danych osobowych studentów.	bardzo wysoki
WNF5	Czas naprawy Maksymalny czas niesprawności systemu po awarii nie może przekraczać 2 godzin.	bardzo wysoki
WNF6	Czas odpowiedzi klientowi Średni czas odpowiedzi klientowi na zapytanie, nie powinien trwać więcej niż 2 sekundy.	średni
WNF7	Maksymalny napływ studentów Maksymalny przypływ nowych studentów do rejestracji nie powinien być większy niż 10 na sekundę.	średni
WNF8	Średni napływ studentów Średni przypływ nowych studentów do rejestracji nie powinien być większy niż 5 na sekundę.	średni
WNF9	Dostępność Średni czas wyłączenia systemu w roku nie powinien przekraczać dwóch dni roboczych.	wysoki
WNF10	Ergonomia Użytkowanie systemu powinno być intuicyjne i przejrzyste dla użytkownika.	wysoki
WNF11	Praca po awarii węzła Awaria pojedynczego węzła nie powinna powodować awarii całej sieci (przy założeniu, że zachowana zostaje minimalna ilość węzłów).	bardzo wysoki
WNF12	Minimalna ilość węzłów do pracy warstwy wewnętrznej Warstwa wewnętrzna serwera przechowująca dane powinna poprawnie działać przy minimalnej liczbie węzłów trzy. Jeśli liczba węzłów po awarii spadnie poniżej tej liczby, poprawne działanie nie jest gwarantowane.	bardzo wysoki
WNF13	Minimalna ilość węzłów do pracy warstwy zewnętrznej Warstwa zewnętrzna serwera przetwarzające dane powinna poprawnie działać jeśli działa przynajmniej jeden węzeł tej warstwy.	bardzo wysoki
WNF13	Pomoc System powinien posiadać dobrze przygotowaną pomoc dla użytkownika.	niski

1.4.3. Dodatkowe założenia

ID	Wymagania	Powody decyzji
ZO1	Protokół TCP/IP Komunikacja między warstwami (wewnętrzna serwera - zewnętrzna serwera jak i zewnętrzna serwera - klienci) odbywa się za pomocą protokołów TCP/IP, wykorzystując Google Protocol Buffer (więcej w podrozdziale 3.2).	Protokół został przez nas wybrany ze względu na elastyczność dla różnych konfiguracji serwerów, tzn. bez względu na to czy warstwy serwera będą działały w obrębie jednego węzła czy nie protokół pozwoli na poprawną komunikację. W przypadku działania obu warstw serwera na jednym komputerze czy nawet w jednym procesie wydajność będzie mniejsza (w porównaniu np. do pamięci współdzielonej), jednak ze względu na ograniczony czas projektu i mniejsze skomplikowanie zostało wybrane to rozwiązanie.
ZO2	Proste komendy serwera Warstwy serwera będą uruchamiane i zamykane jednokrotnym wywołaniem skryptu na danym węźle serwera z jednym argumentem: start, stop oraz status, które będą kolejno: uruchamiać usługi serwera, wyłączać usługi serwera oraz wyświetlać status serwera.	Takie rozwiązanie jest wymagane w treści zadania. Ponadto pozwoli na łatwą obsługę i testowanie części serwerowej systemu.
ZO3	Technologia token ring (warstwa wewnętrzna serwera) Serwery warstwy wewnętrznej połączone są ze sobą w pierścień (każdy ma z góry ustalonego poprzednika i następnika w pliku konfiguracyjnym). Jeśli dany następnik nie odpowiada - brany jest następny adres węzła z listy. W ten sposób bez względu na ilość węzłów mamy zawsze do czynienia z taką samą strukturą (w szczególności - mamy tylko jeden węzeł). Każdy z węzłów wysyła swojemu następnikowi token zawierający informacje o zmianach potencjalnych zmian.	Taka architektura warstwy wewnętrznej pozwala na bieżące sprawdzanie stanu kolejnych węzłów. Jeśli token gdzieś się „zagubi” oznacza to, że dany węzeł uległ awarii. Ponadto taka architektura pozwala na szybką i łatwą synchronizację danych.

ID	Wymagania	Powody decyzji
ZO4	Pełna redundancja danych (warstwa wewnętrzna serwera) Każdy z węzłów wewnętrznych ma pełną i aktualną bazę studentów. Dane są synchronizowane na bieżąco.	Natychmiastowa synchronizacja danych i pełna redundancja pozwala uniknąć wielu problemów w trakcie działania serwera, między innymi: dodatkowej synchronizacji po awarii jednego z węzłów, która jest dość problematyczna i skomplikowana. Rozwiązanie to zostało zatwierdzone przez Prowadzącego.
ZO5	Wysyłanie „heartbeatów” (warstwa zewnętrzna serwera) Węzły zewnętrznej warstwy będą sprawdzały swoją dostępność za pomocą tzw. „uderzeń serce” (<i>ang. heartbeat</i>) wysyłanych w danych odstępach czasu (każdy z serwerów musi mieć ustalony ten sam interwał). Dzięki temu będą miały wiedzę na temat aktywności pozostałych serwerów. Brak „heartbeata” od danego serwera po upływie zadanego czasu będzie równoważny z jego brakiem dostępności.	Takie rozwiązanie pozwala na bieżąco monitorować stan poszczególnych węzłów warstwy zewnętrznej.
ZO6	Brak funkcji logowania użytkownika Klient łączy się z systemem identyfikując się jedynie swoim unikatowym ID, które wpisane jest pliku konfiguracyjnym.	Decyzja została zatwierdzona przez Prowadzącego. Wynika ona z braku konieczności stworzenia tej funkcjonalności (w związku z minimalnym wpływem na główny cel projektu) oraz na ograniczony czas i mniejszą ilość osób w grupie projektowej niż przewidywana.
ZO7	Brak interfejsu graficznego (aplikacja kliencka) -	Decyzja została zatwierdzona przez Prowadzącego. Wynika ona z braku konieczności stworzenia tej funkcjonalności (w związku z minimalnym wpływem na główny cel projektu) oraz na ograniczony czas i mniejszą ilość osób w grupie projektowej niż przewidywana.
ZO8	Komunikacja synchroniczna (aplikacja kliencka) Aplikacja kliencka będzie się komunikowała z serwerem w sposób synchroniczny.	Ze względu na ograniczoną funkcjonalność aplikacji klienckiej, sprowadzoną jedynie do wysyłania żądań i uzyskiwania odpowiedzi, ten sposób komunikacji wydaje się najbardziej trafny.

ID	Wymagania	Powody decyzji
ZO9	Brak funkcji „heartbeatu” (aplikacja kliencka) Aplikacja kliencka nie będzie na bieżąco sprawdzała stanu węzła do którego jest podłączona. Status serwera będzie przez nią sprawdzany jedynie przy logowaniu i wysyłaniu żądania. W przypadku awarii węzła aplikacja będzie przełączała się na kolejny węzeł z listy, aż do skutku.	Ze względu na naszą koncepcję aplikacji klienckiej ta funkcjonalność nie jest potrzebna. Z punktu widzenia użytkownika stan serwera interesuje go tylko przy wysyłaniu żądania. System heartbeatów dodaje duży narzut i komplikuje samego klienta nie dając w zamian znacznego wzrostu szybkości działania. W sytuacji awarii i tak musi nastąpić przełączenie na inny węzeł. Różnica polega jedynie na tym, że w przypadku systemu heartbeatów następuje to w momencie upływu określonego czasu bez odpowiedzi serwera, a w przeciwnym - jedynie wtedy kiedy serwer jest potrzebny, tzn. przy logowaniu bądź wysłaniu żądania.
ZO10	Brak interaktywności (aplikacja kliencka) Aplikacja kliencka będzie ograniczona jedynie do wywoływania określonych funkcji, według określonego z góry scenariusza.	W związku z ograniczonym czasem i głównym celem projektu interaktywna aplikacja kliencka nie jest koniecznością. Funkcjonalność systemu można poprawnie przetestować inicjując konkretne scenariusze (między innymi symulacje awarii czy błędów systemu). Decyzja ta została zatwierdzona przez Prowadzącego.

1.5. Potencjalne problematyczne scenariusze

- Podłączenie nowego serwera warstwy wewnętrznej
 - Brak innych aktywnych serwerów wewnętrznych
 - Jest aktywny przynajmniej jeden serwer wewnętrzny
- Podłączenie nowego serwera warstwy zewnętrznej
 - Brak aktywnego serwera wewnętrznego oraz brak innych aktywnych serwerów zewnętrznych
 - Brak aktywnego serwera wewnętrznego oraz aktywny przynajmniej jeden serwer zewnętrzny
 - Aktywny przynajmniej jeden serwer wewnętrzny oraz brak innych aktywnych serwerów zewnętrznych
 - Aktywny przynajmniej jeden serwer wewnętrzny oraz aktywny przynajmniej jeden serwer zewnętrzny

- Awaria węzłów warstwy wewnętrznej serwera
 - Nie posiada tokena, brak zadań w toku, brak oczekujących zadań
 - Nie posiada tokena, brak zadań w toku, oczekujące zadanie od klienta
 - Nie posiada tokena, zadanie w toku
 - Posiada token, brak zadań w toku, brak oczekujących zadań
 - Posiada token, zadanie w toku, brak zadań w toku, oczekujące zadanie od klienta
 - Posiada token, zadanie w toku
- Awaria węzłów warstwy zewnętrznej serwera
- Kwestia równomiernego podziału klientów między serwery oraz serwerów zewnętrznych między wewnętrzne
- Kwestia rozmieszczenia warstw serwera w obrębie jednego komputera bądź wielu odrębnych

1.6. Organizacja środowiska programistycznego

1.6.1. Język programowania Java

Zdecydowaliśmy się napisać kod źródłowy systemu w Javie. Wybór ten wynika z wielu zalet tego języka programowania. Funkcje sieciowe są dostępne w podstawowych bibliotekach. Wliczone w to są TCP/IP, UDP/IP, wydajne nieblokujące I/O, HTTP, wsparcie REST, XML, JSON oraz SSL.

Kolejną zaletą jest duży zbiór struktur danych (również w podstawowych bibliotekach). Budowa rozproszonego systemu wymaga wielu, często skomplikowanych i różnorodnych, struktur, które Java posiada. Ważna jest tu także kwestia wątków i synchronizacji między nimi. Na tym polu Java również dostarcza odpowiednie funkcje, struktury danych oraz narzędzie realizujące współdzielony dostęp. Ponadto bez problemu współpracuje z Google Protocol Buffers i jest multiplatformowa.

1.6.2. IntelliJ IDE

Jest to środowisko programistyczne dla Javy. Owe środowisko wyposażono w bardzo bogatą paletę narzędzi pozwalających na komfortowe tworzenie oraz edycję kodu. Mnogość narzędzi oraz znajomość środowiska przez zespół ostatecznie zaważyła na decyzji o wykorzystaniu IntelliJ.

2. Organizacja projektu

2.1. Podział zadań

Kierownik projektu, Dokumentalista - Domagała Bartosz

Odpowiedzialny za planowanie, realizację oraz zamykanie projektu. Ma zapewnić osiągnięcie założonych celów projektu i wytworzenie oprogramowania spełniającego określone wymagania jakościowe. Ponadto odpowiedzialny za całą dokumentację projektu, nadzorujący jej tworzenie i ostateczną formę.

Handlowiec, Specjalista ds. Docker'a, Programista - Kornata Jarosław

Odpowiedzialny za prezentację projektu przed prowadzącym. Ponadto główny specjalista rozwiązania Docker, posiadający na ten temat największą wiedzę, przekazywaną w trakcie projektu pozostałym osobom. Bierze czynny udział w tworzeniu kodu źródłowego programu oraz sumiennie dokumentuje swoją pracę.

Tester, Programista - Jędrzej Modzelewski

Odpowiedzialny za wszelkie testy związane z oprogramowaniem - projektujący je, programujący oraz egzekwujący. Bierze czynny udział w tworzeniu kodu źródłowego programu oraz sumiennie dokumentuje swoją pracę.

Architekt, Osoba odp. za repozytorium, Programista - Stepnowski Marcin

Odpowiedzialny za stworzenie i dbanie o odpowiednie wdrożenie architektury projektu oraz dba o zgodność tworzonych rozwiązań informatycznych z obowiązującymi standardami, wzorcami i strategią. Ponadto odpowiedzialny za założenie, zarządzanie oraz tworzenie kopii zapasowych repozytorium projektu. Bierze czynny udział w tworzeniu kodu źródłowego programu oraz sumiennie dokumentuje swoją pracę.

2.2. Planowany harmonogram pracy

Spotkanie I - koniec marca

- Spotkanie organizacyjne
- Ustalenie ról w projekcie i podział zadań
- Rozmowa na temat wizji projektu
- Umówienie się co do organizacji pracy i spotkań projektowych

Spotkanie II - 1 tydzień kwietnia

- Nauka i zrozumienie zasad działania Docker'a

- Wspólne sprawdzenie testowych konfiguracji Dockera
- Ustalenie wstępnej architektury systemu
- Ustalenie wymagań projektu
- Ustalenie języka w jakim projekt ma być tworzony
- Rozmowa na temat dodatkowych narzędzi do realizacji projektu

Spotkanie III - 2 tydzień kwietnia

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Rozważanie na temat potencjalnych wad systemu i możliwych sytuacji awaryjnych
- Ostateczna definicja wymagań funkcjonalnych i нефункциональных programu

Spotkanie IV - 3 tydzień kwietnia

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Upewnienie się o poprawności dotychczasowych założeń, wymagań
- Wprowadzenie poprawek do dokumentacji
- Opracowanie ostatecznej dokumentacji na prezentację etapu I

Prezentacja etau I - 23 kwietnia**Spotkanie VI - 4 tydzień kwietnia**

- Wprowadzenie poprawek i sugestii Prowadzącego po prezentacji etapu I
- Podział pracy przy pisaniu kodu źródłowego programu
- Ustalenie i doprecyzowanie funkcjonalności na prezentację etapu II
- Rozpoczęcie pisania kodu źródłowego
- Dyskusja na temat testów systemu i ich definicja
- Szczegółowe doprecyzowanie rozwiązania

Spotkanie VII - 4 tydzień kwietnia

- Wykonanie czynności i kwestii, które nie zostały zakończone bądź poruszone podczas spotkania VI
- Praca nad kodem źródłowym programu i poprawa błędów
- Testy oprogramowania, przygotowanie do prezentacji funkcjonalności

Spotkanie VIII - 1 tydzień maja

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Zdefiniowanie pełnego planu testów i upewnienie się o jego poprawności
- Ostateczne poprawki do dokumentacji dla etapu II
- Ostateczne przygotowanie prezentacji funkcjonalności systemu dla etapu II

Prezentacja etau II - 7 maja**Spotkanie IX - 2 tydzień maja**

- Wprowadzenie poprawek i sugestii Prowadzącego po prezentacji etapu II
- Praca nad kodem źródłowym programu i poprawa błędów
- Ustalenie pozostałych zadań do zrobienia

Spotkanie X - 3 tydzień maja

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Dyskusja na temat dalszego planu tworzenia kodu źródłowego systemu
- Praca nad kodem źródłowym programu i poprawa błędów

Spotkanie XI - 4 tydzień maja

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Praca nad kodem źródłowym programu i poprawa błędów
- Opracowanie wstępnej prezentacji końcowej

Spotkanie XII - 1 tydzień czerwca

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Praca nad kodem źródłowym programu i poprawa błędów
- Poprawki do prezentacji końcowej

Spotkanie XIII - 1 tydzień czerwca

- Spotkanie awaryjne, pozwalające na wykonanie zaległych czy niedokończonych zadań z poprzednich spotkań

Spotkanie XIV - 2 tydzień czerwca

- Sprawdzenie i podsumowanie dotychczasowej pracy
- Ostateczne poprawki do dokumentacji dla etapu III
- Ostateczne przygotowanie prezentacji funkcjonalności systemu dla etapu III

Spotkanie XV - 2 tydzień czerwca

- Spotkanie awaryjne, pozwalające na wykonanie zaległych czy niedokończonych zadań z poprzednich spotkań

Prezentacja etapu III - 11 czerwca**Spotkanie XVI - 11 czerwca**

- Świątowanie zaliczenia projektu przy piwie

2.3. Sprawozdania ze spotkań

Sprawozdania z konkretnych spotkań zawierają listę ustaleń oraz wykonanych czynności w związku z projektem. Poza spotkaniami, każdy z uczestników projektu pracował nad przydzielonymi zadaniami indywidualnie, we własnym zakresie, trzymając się konkretnych terminów.

2.3.1. Spotkanie I - 31 marca 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *obecny*
- Stepnowski Marcin - *obecny*

Spotkanie organizacyjne, którego celem był podział ról w projekcie oraz przydzielenie zadań, a także wstępne omówienie samej wizji projektu.

Szczegółowy opis wykonanej pracy

- Kierownik projektu ustalił podział ról zgodnie z umiejętnościami i preferencjami osób z grupy projektowej:
 - Domagała Bartosz - Kierownik, Dokumentalista
 - Kornata Jarosław - Handlowiec, Specjalista ds. Docker'a, Programista
 - Modzelewski Jędrzej - Tester, Programista
 - Stepnowski Marcin - Architekt, Osoba odpowiedzialna za repozytorium, Programista
- Ustalony został termin cotygodniowych spotkań projektowych
- Omówiona została treść zadania, upewniono się, że każdy z uczestników projektu je rozumie
- Prowadzone były rozmowy na temat wizji projektu każdego z uczestników, dobre pomysły spisywane i komentowane
- Zastanowiono się nad sprecyzowaniem treści zadania
- Każdy z uczestników projektu otrzymał zadania do wykonania:
 - Domagała Bartosz - stworzenie szkieletu dokumentacji za pomocą LaTeXa, zapisanie sprawozdania z odbytego spotkania
 - Kornata Jarosław - dowiedzenie się jak najwięcej o rozwiązaniu Docker
 - Modzelewski Jędrzej - pomyślenie o narzędziach, bibliotekach i językach programowania jakich można użyć w projekcie
 - Stepnowski Marcin - stworzenie repozytorium Git oraz zastanowienie się nad architekturą projektu

2.3.2. Spotkanie II - 8 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *obecny*
- Stepnowski Marcin - *obecny*

Spotkanie podsumowujące ostatnio wykonane zadania oraz występujące problemy. Ponadto podczas niego ustalono wstępną treść zadania i omówiono, a także przetestowano rozwiązanie Docker.

Szczegółowy opis wykonanej pracy

- Przekazanie przez specjalistę ds. Dockera informacji o tym rozwiązaniu pozostałym uczestnikom
- Wspólne sprawdzenie testowych konfiguracji Dockera
- Wstępny projekt architektury projektu zaproponowany przez Architekta
- Zastanowienie się nad wymaganiami niefunkcjonalnymi projektu
- Ustalenie języka w jakim stworzony zostanie projekt - Java
- Rozmowa na temat przydatnych i potrzebnych bibliotek w projekcie
- Każdy z uczestników projektu otrzymał zadania do wykonania:
 - Domagała Bartosz - zapis sprawozdania ze spotkania
 - Kornata Jarosław - zapis informacji o rozwiązaniu Docker oraz opis przeprowadzonych testów jego konfiguracji
 - Modzelewski Jędrzej - zapis informacji o platformie, języku i bibliotekach używanych w projekcie
 - Stepnowski Marcin - zapis informacji o projekcie architektury

2.3.3. Spotkanie X1 - 14 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *nieobecny*
- Modzelewski Jędrzej - *nieobecny*
- Stepnowski Marcin - *obecny*

Pierwsze spotkanie z Prowadzącym projekt. Informacje ze spotkania:

- Zatwierdzenie przez Prowadzącego tematu
- Zgoda Prowadzącego na brak funkcji logowania w aplikacji klienckiej, ze względu na główny cel projektu, ograniczony czas i mniejszą liczbę osób w grupie niż przewidziana
- Zgoda Prowadzącego na brak graficznego interfejsu użytkownika, ze względu na główny cel projektu, ograniczony czas i mniejszą liczbę osób w grupie niż przewidziana
- Zgoda Prowadzącego na brak interaktywnej aplikacji klienckiej, ze względu na główny cel projektu, ograniczony czas i mniejszą liczbę osób w grupie niż przewidziana
- Doprecyzowanie szczegółów w związku z 3-warstwową naturą systemu
- Uzyskanie odpowiedzi na pytania odnośnie architektury

2.3.4. Spotkanie III - 14 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *obecny*
- Stepnowski Marcin - *obecny*

Spotkanie podsumowujące ostatnio wykonane zadania oraz występujące problemy. Zastanowiono się podczas niego nad dotychczasowym wyborem rozwiązań, sprecyzowano treść zadania, ustalono wymagania funkcjonalne i нефункционалне projektu.

Szczegółowy opis wykonanej pracy

- Dyskusja na temat założeń systemu i jego wstępnej architektury
- Wprowadzenie poprawek do projektu na podstawie wniosków z dyskusji
- Wspólne ustalenie dodatkowych założeń, przesłanek do ich egzekwowania i konsekwencji z tym związanych
- Definicja wymagań funkcjonalnych i нефункционалных programu
- Rozmowa na temat wad systemu, potencjalnych zagrożeń, spisanie pomysłów
- Każdy z uczestników projektu otrzymał zadania do wykonania:
 - Domagała Bartosz - zapis sprawozdania ze spotkania, poprawienie błędów w dokumentacji oraz edycja szkieletu stylu, upewnienie się, że naniesione przez resztę grupy poprawki są prawidłowe
 - Kornata Jarosław - wprowadzenie poprawek do opisu przykładowej konfiguracji Dockera
 - Modzelewski Jędrzej - wprowadzenie poprawek do opisu języka oraz IDE w jakim tworzony będzie projekt
 - Stepnowski Marcin - rozwinięcie i wyjaśnienie decyzji projektowych dotyczących architektury podjętych w trakcie spotkania oraz zapisanie ich w dokumentacji

2.3.5. Spotkanie X2 - 21 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *nieobecny*
- Modzelewski Jędrzej - *nieobecny*
- Stepnowski Marcin - *obecny*

Drugie spotkanie z Prowadzącym projekt. Informacje ze spotkania:

- Zgoda Prowadzącego na pełną redundancję w warstwie zewnętrznej serwera (tzn. każdy z węzłów ma posiadać wszystkie aktualne dane) w związku z ograniczonym czasem i mniejszą liczbą osób w grupie projektowej niż przewidziana

- Zdobycie informacji odnośnie warstw serwera, które mogą być na tym samym węźle jak i na dwóch niezależnych węzłach
- Zgoda Prowadzącego na użycie protokołu TCP/IP do komunikacji między warstwami serwera bez względu na ich lokalizację - w związku z ograniczonym czasem i mniejszym skomplikowaniem zagadnienia

2.3.6. Spotkanie IV (TeamSpeak) - 22 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *nieobecny*
- Stepnowski Marcin - *obecny*

Spotkanie za pośrednictwem programu TeamSpeak mające na celu podsumowanie dotychczasowej pracy oraz wprowadzenie ostatecznych poprawek do dokumentacji przed prezentacją etapu I projektu.

Szczegółowy opis wykonanej pracy

- Poprawienie błędów w dokumentacji
- Dopisanie brakujących rzeczy do dokumentacji (głównie - kwestii potencjalnych problemów z systemem)
- Ostateczna edycja i poprawki szablonu projektu
- Podsumowanie dotychczasowej pracy, umówienie się na kolejne spotkanie w celu zaczęcia pracy nad kodem źródłowym systemu
- Rozmowa i ustalenia na temat prezentacji etapu I

2.3.7. Spotkanie V - 27 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *obecny*
- Stepnowski Marcin - *obecny (TeamSpeak)*

Pierwsze spotkanie po oddaniu etapu I projektu. Omówienie niezbędnych poprawek do dotychczasowej dokumentacji oraz opracowanie planu dalszej pracy związanej z kodem źródłowym programu i prezentacją.

Szczegółowy opis wykonanej pracy

- Poprawki w dokumentacji związane z sugestiami od Prowadzącego
- Opracowanie wstępnego planu testów
- Podział obowiązków przy pisaniu kodu źródłowego programu:

- Kornata Jarosław - struktura wiadomości przesyłanych między warstwami (oraz klasy pomocnicze); obiekty i funkcje liczące dane statystyczne w warstwie zewnętrznej; klasy do komunikacji z bazą danych, program do generowania danych wrażliwych
- Modzelewski Jędrzej - komunikacja wewnętrzna serwerów wewnętrznych, komunikacja między warstwami
- Stepnowski Marcin - komunikacja wewnętrzna serwerów zewnętrznych, komunikacja między warstwami
- Opracowanie wstępnego scenariusza prezentacji etapu II projektu
- Ustalenie zastosowania wzorca czynnościowego wzorca projektowego przy komunikacji między warstwami zwanego łańcuchem zobowiązań

2.3.8. Spotkanie X3 - 28 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *nieobecny*
- Modzelewski Jędrzej - *nieobecny*
- Stepnowski Marcin - *obecny*

Trzecie spotkanie z Prowadzącym projekt. Informacje ze spotkania:

- Pytanie odnośnie prezentacji funkcjonalności podczas oddawania II etapu - zgoda Prowadzącego na pokazanie działającego systemu, ale bez zadbania o kwestie awarii systemu czy redundancję
- Pytanie odnośnie struktury testów i potwierdzenie przez Prowadzącego, że testy nie muszą być automatyczne, ale możliwe do ponownej symulacji
- Rozmowa na temat token ringu i redundancji danych w warstwie wewnętrznej

2.3.9. Spotkanie V (TeamSpeak) - 28 kwietnia 2015

- Domagała Bartosz - *obecny*
- Kornata Jarosław - *obecny*
- Modzelewski Jędrzej - *obecny*
- Stepnowski Marcin - *obecny*

Spotkanie za pośrednictwem programu TeamSpeak mające na celu ustalenie struktury wiadomości jakie będą służyły do komunikacji w systemie oraz dalszy podział pracy.

Szczegółowy opis wykonanej pracy

- Ustalenie struktury wiadomości do komunikacji wewnątrz warstw i pomiędzy między warstwami systemu

3. Narzędzia i zewnętrzne biblioteki

3.1. Docker

3.1.1. Czym jest docker?

Docker jest narzędziem przeznaczonym do tworzenia przenośnych, wirtualnych kontenerów pozwalających na prostą i szybką replikację środowiska.

Oprogramowanie Docker wprowadza standaryzację w środowisko uruchomieniowe. Generowane kontenery są spójne i takie same w różnych środowiskach. W chwili obecnej Docker wymaga jądra Linux do uruchomienia, jednak działa nie tylko w wersji natywnej, ale również poprzez wirtualizację specjalnych minimalistycznych dystrybucji Linux na systemach Mac OS X oraz Windows (na przykład w darmowym narzędziu do wirtualizacji – VirtualBox).

3.1.2. Spójność

Docker wprowadza spójność w środowisko deweloperskie. Często, kiedy nad projektem pracuje więcej niż jeden programista (co w dzisiejszych czasach jest normą) pojawia się problem z różnymi wersjami użytego oprogramowania. Zdarza się, że jedna wersja biblioteki zachowuje się inaczej od drugiej (na przykład z powodu błędu). Kontener z założenia posiada jedną, konkretną wersję każdej biblioteki koniecznej do uruchomienia aplikacji. Zwiększa to przewidywalność oprogramowania. Oczekujemy bowiem, że w tym samym środowisku, aplikacja będzie zachowywać się tak samo.

3.1.3. Cykl Dockera

1. Stworzenie kontenera wraz z wszelkimi narzędziami i bibliotekami koniecznymi do uruchomienia aplikacji.
2. Rozprowadzenie kontenera, wraz ze wszystkimi narzędziami i bibliotekami.
3. Uruchomienie identycznego kontenera na dowolnej liczbie węzłów.

Prowadzi to do znacznych ułatwień w tworzeniu oprogramowania. Ten sam kontener może zostać rozprowadzony pomiędzy deweloperami, testerami, serwerami ciągłej integracji i w końcu środowiskiem produkcyjnym.

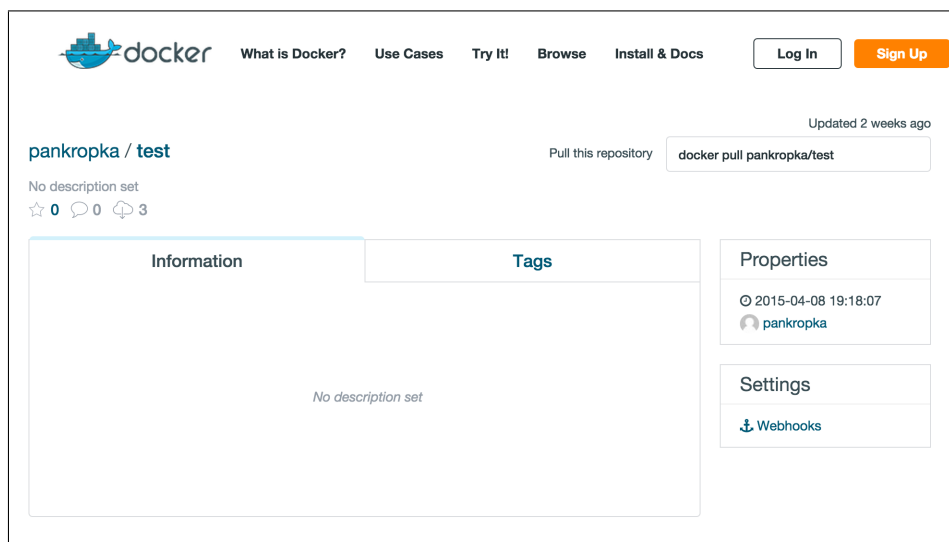
3.1.4. Centralne repozytorium

Każdy zarejestrowany użytkownik ma możliwość wgrywania własnych kontenerów do centralnego, ogólnie dostępnego repozytorium. Można tam znaleźć wiele różnych gotowych obrazów do pobrania. Są one przygotowane zarówno przez społeczność jak i przez twórców Dockera.

Centralne repozytorium pozwala użytkownikom na pobranie interesujących obrazów, szczególnie warte uwagi są repozytoria przygotowane pod konkretne rozwiązania (na przykład specjalnie pod serwer baz danych MySQL lub mongoDB).

Docker umożliwia również stworzenie prywatnych repozytoriów. Dzięki temu nie musimy upubliczniać prywatnych obrazów, ale jednocześnie możemy je rozprowadzać. Inną dostępną metodą jest zapis obrazu do pliku i przekazanie go w tradycyjny sposób.

3.1.5. Przeprowadzone testy



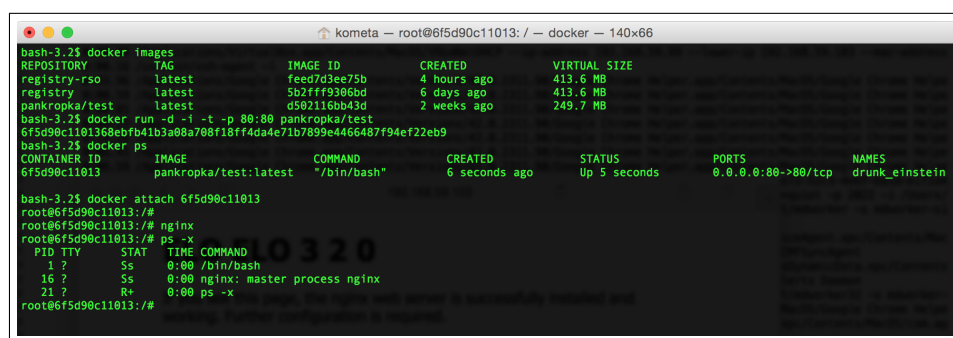
Rysunek 3.1. Repozytorium Docker

W ramach zapoznania się z platformą Docker wykonaliśmy następujące czynności:

- Zweryfikowaliśmy możliwość uruchomienia oprogramowania na systemach operacyjnych Ubuntu oraz Mac OS X.
- Na systemie Max OS X skorzystaliśmy z programu Boot2Docker, tworzącego maszynę wirtualną z minimalistyczną dystrybucją Linux konieczną do uruchomienia aplikacji Docker.
- Pobraliśmy obraz Systemu operacyjnego Ubuntu, na którym zainstalowaliśmy programy nginx oraz vim.
- Przy pomocy programu vim, zmodyfikowaliśmy treść domyślnego pliku index.html, który dostępny jest na porcie 80 po uruchomieniu programu nginx.
- Zweryfikowaliśmy zmianę w przeglądarce internetowej.



Rysunek 3.2. Pobranie obrazu i weryfikacja zmian



Rysunek 3.3. Działa!

- Zapisaliśmy zmiany w kontenerze oraz wgraliśmy zmieniony obraz do centralnego repozytorium (Rysunek 3.1).
- Pobraliśmy obraz na innym komputerze i po uruchomieniu zweryfikowaliśmy zmiany wprowadzone do podstawowej dystrybucji Ubuntu (Rysunek 3.2).
- W pobranym obrazie zainstalowane były programy nginx oraz vim. Domyślny plik programu nginx był również zmodyfikowany (Rysunek 3.3).

3.2. Google Protocol Buffers

Jest to elastyczna i wydajna metoda serializacji strukturyzowanych danych do wykorzystania między innymi w protokołach komunikacyjnych oraz przechowywaniu danych. Określana jest struktura informacji która będzie serializowana poprzez zdefiniowanie typu wiadomości w pliku „.proto”. Każda taka wiadomość jest logicznym rekordem informacji zawierającym pary nazwa-wartość. Format takiej wiadomości jest prosty, każdy typ składa się z jednego lub więcej unikatowych, ponumerowanych pól z których każdy ma nazwę i typ wartości. Typem wartości pola mogą być liczby, zmienne logiczne, łańcuchy znaków oraz inne wiadomości PBM. Pozwala to na hierarchiczne ułożenie struktury. Każde pole może być oznaczone na trzy sposoby: opcjonalne, wymagane, powtarzające się z zachowaniem kolejności.

Kiedy wiadomości są określone należy włączyć kompilator protocol buffer dla języka aplikacji. Generuje on klasy dostępu do danych na podstawie pliku „.proto”. Tworzone są podstawowe funkcje dostępu do każdego pola oraz metody do serializowania/parsowania całej struktury. Dla Javy generowane są pliki „.java” z klasami dla każdego typu wiadomości oraz specjalna klasa „Builder” do tworzenia instancji klas wiadomości.

Można dodać nowe pola do wiadomości nie zaburzając kompatybilności wstecznej. Stare pliki binarne ignorują nowe pola podczas parsowania. W przypadku protokołów komunikacyjnych które wykorzystują protocol buffers jako format danych, można rozszerzyć protokół bez obaw o wpływ na istniejący kod.

Wszystkie wyżej wymienione cechy zaważyły przy wyborze tego narzędzia do tworzenia systemu.

3.3. MySQL

Aplikacja serwerowa będzie działać w oparciu o relacyjną bazę danych MySQL.

Baza MySQL jest popularnym systemem relacyjnych baz danych z dobrą dokumentacją oraz szeroką i aktywną społecznością użytkowników. Ponadto system ten jest powszechnie znany wśród wszystkich członków projektu. Dodatkowo MySQL dobrze współpracuje z wybranym językiem programowania Java. Stąd nasz wybór padł właśnie na MySQL.