

Kurs DevOps

Lista 9

11 i 17 grudnia 2025

Jako że zazwyczaj w dużych grupach znajduje się co najmniej jeden użytkownik Windowsa, to informuje się, iż zadania należy wykonywać z linii poleceń, o ile nie powiedziano inaczej w treści.

Zadanie 1. (4 p.)

Przy pomocy Kuberntesa uruchom dwie repliki serwera/aplikacji używającej HTTP, a następnie:¹

- (1p.) a) Skonfiguruj load balancing z użyciem HAProxy, tak by zachowywało sesję, tzn. pakiety od jednego klienta powinny być wysyłane za każdym razem do tego samego serwera. Zademonstruj, że Twoja konfiguracja działa.²
- (1p.) b) Dodaj trzecią replikę i potraktuj ją jako kanarka. Zdecyduj jaki procent ruchu należy kopiować do wersji testowej, a następnie odpowiednio skonfiguruj load-balancer.³
- (1p.) c) Pokaż, w jaki sposób do pakietu otrzymanego od klienta dodać nagłówek HTTP przed wysłaniem go do serwera. Następnie zademonstruj jak usunąć nagłówek z odpowiedzi, która przychodzi od serwera przed jej wysłaniem do klienta.
- (1p.) d) Dodaj do swojego load balansera reguły weryfikujące przychodzący ruch. W szczególności i) odrzucaj żądania HTTP wysyłane przez wget (lub alternatywnie wybrany inny program do automatycznego pobierania stron); ii) ogranicz liczbę żądań HTTP, które można wysłać w ustalonym okresie czasu z danego adresu IP.

Zadanie 2.

Rozszerz poprzednie zadanie o demonstrację, jak skonfigurować HAProxy, tak by ukryć przed klientem problemy z serwerem, który czasami nie odpowiada. Pokaż także, jak w przypadku błędu, przekierować połączenie od klienta do innego serwera.

Zadanie 3. (1.5 p.)

Rozszerz zadanie 1 i zaprezentuj, w jaki sposób skonfigurować HAProxy jako offloader SSL. Niech klient łączy się z HAProxy po HTTPS, proxy niech weryfikuje certyfikat i deszyfruje ruch, a następnie niech prześle zdeszyfrowane pakiety po HTTP do zwykłego serwera. Jakie zalety i wady ma takie rozwiązanie?

Zadanie 4. (1.5 p.)

Pokaż jak skonfigurować ECMP (zarówno na warstwie 3, jak i na warstwie 4), które dzieli ruch na trzy różne ścieżki.

Zadanie 5.

Zapoznaj się z "Case Study 2: When Load Shedding Attacks" z [1] w rozdziale 11. Czy balansowanie obciążenia (ang. load balancing) zawsze pomaga na niezawodność? Czy może zamiast tego doprowadzić do problemów? Czym jest load scheduling?

¹Każdy podpunkt można deklarować niezależnie.

²Z użyciem miniserve można np. wysyłać duży (większy niż kilka kB) plik na serwer. Będzie on musiał być podzielony na mniejsze pakiety, więc jeśli nie wszystkie dojdą do serwera, powinien pojawić się błąd.

³Może przyda się polityka round-robin?

Bibliografia

- [1] Betsy Beyer i in. *The Site Reliability Workbook: Practical Ways to Implement SRE*. 1st. O'Reilly Media, Inc., 2018. ISBN: 1492029505. URL: <https://sre.google/workbook/table-of-contents/>.