



Wyższa Szkoła Technologii Informatycznych w Katowicach

REFERAT PRACY DYPLOMOWEJ

Temat pracy: Analiza porównawcza migracji lokalnej infrastruktury serwerowej do wybranych rozwiązań chmur publicznych

Autor: inż. Paweł Michalski

Promotor: dr Jarosław Karcewicz

CEL I PODSTAWOWE ZAŁOŻENIA

Celem niniejszej pracy magisterskiej jest przeprowadzenie analizy i porównania wydajności różnych maszyn wirtualnych oferowanych przez dwóch wiodących dostawców usług chmurowych typu Infrastructure-as-a-service (IaaS) - Amazon EC2 i Microsoft Azure. W ramach pracy przeprowadzone zostaną różne testy, obejmujące wydajność procesora, pamięci RAM, dysku, przepustowość sieci oraz opóźnienia, w celu dokładnego zbadania wydajności maszyn wirtualnych. Po zebraniu danych przeprowadzona zostanie szczegółowa analiza w celu zidentyfikowania przyczyn różnic w wydajności poszczególnych komponentów pomiędzy dostawcami chmurowymi. Kolejnym celem pracy jest przeprowadzenie analizy cenowej, która pozwoli porównać rzeczywistą wydajność podobnych ofert od obu dostawców chmurowych oraz ocenić, jak znacząca jest różnica w cenie między nimi dla podobnych typów maszyn wirtualnych.

REALIZACJA PRACY

Rozpoczęto od charakteryzacji chmury obliczeniowej, omówiono różnorodne rodzaje i typy tego rozwiązania, jak również wyróżnieni dostawcy określani jako "liderzy" przez



Wyższa Szkoła Technologii Informatycznych w Katowicach

międzynarodową firmę analityczno-doradcą Gartner. Na podstawie raportu przedstawionego przez wspomnianą firmę, dokonano wyboru oraz opisu trzech głównych liderów: Amazon AWS, Google GPC oraz Microsoft Azure.

W ramach tej pracy wyselekcjonowano dwie instancje od dostawców Amazon oraz Microsoft, które spełniały wytyczone kryteria i były możliwie najbardziej zbliżone pod względem kosztów i konfiguracji sprzętowej. W związku z tym wybrano maszyny Azure F2s_v2 oraz F4s_v2, a także maszyny z serii AWS c5.large oraz c5.xlarge. Po zebraniu i przygotowaniu testów porównawczych, opracowano metodologię analizy wydajności i kosztów w środowisku chmurowym.

W celu przeprowadzenia testów wydajnościowych, wybrano Phoronix Test Suite - wszechstronną platformę testową, która oferuje wiele mikro benchmarków do oceny wydajności poszczególnych komponentów sprzętowych. Platforma ta jest stale rozwijana, a nowe narzędzia mogą w łatwy sposób być dodawane.

Zostały wybrane następujące testy:

- CoreMark
- 7-zip Compression
- C-Ray
- Intel Memory Latency Checker
- IOzone
- SQLite
- PHPBench
- PyPerformance

Przeprowadzono testy w różnych wariantach, co zaowocowało przeprowadzeniem łącznie 25 testów w różnych konfiguracjach.



Wyższa Szkoła Technologii Informatycznych w Katowicach

Prócz narzędzia Phoronix Test Suite, skorzystano także z aplikacji Speedtest CLI do przeprowadzenia testów sieciowych. Speedtest CLI zostało stworzone przez firmę Ookla i umożliwia pomiar prędkości połączenia internetowego za pomocą wiersza poleceń. W kontekście niniejszej pracy opracowano skrypt w języku PowerShell, który miał za zadanie gromadzić informacje dotyczące wydajności połączenia internetowego i zapisywać je w pliku CSV.

Aby przeprowadzić rzetelną analizę wydajności, postanowiono przyjąć następujące założenia:

- Wszystkie testy wydajności zostały przeprowadzone w jednolitym środowisku testowym, które zostało starannie skonfigurowane i zoptymalizowane. Każdy dostawca chmurowy działał w tej samej konfiguracji, aby zapewnić porównywalność wyników.
- W celu zapewnienia obiektywnego środowiska testowego, wybrano region lokalizacji maszyn wirtualnych u obu dostawców chmurowych, który był jak najbardziej wspólny. Ma to na celu zminimalizowanie wpływu geograficznego na wyniki testów.
- Przeprowadzono testy wydajności na różnych typach instancji dostępnych u każdego dostawcy chmurowego. Wybrano instancje, które najlepiej pasowały do badanego przypadku użycia, biorąc pod uwagę wymagania dotyczące mocy obliczeniowej, pamięci, oraz sieci.
- Testy zostały wykonane w dwóch różnych okresach czasowych. Pierwszy z nich odbył się w godzinach pracy, podczas gdy drugi został przeprowadzony poza godzinami biznesowymi, po godzinie 23:00 czasu lokalnego. Taki podział pozwolił na ocenę wydajności systemu zarówno w normalnych warunkach użytkowania, jak i w okresach o niższym obciążeniu. Dzięki temu uzyskano kompleksową analizę wydajności systemu oraz możliwość oceny jego stabilności w różnych scenariuszach.



Wyższa Szkoła Technologii Informatycznych w Katowicach

- Aby zapewnić wiarygodność wyników testów wydajności, przyjęto założenie odczekania co najmniej 15 minut po uruchomieniu systemu operacyjnego przed rozpoczęciem testów. Ten czas pozwala na stabilizację maszyny wirtualnej i upewnienie się, że żadne procesy uruchomieniowe nie obciążają jej zasobów.
- Każdy test został powtórzony minimum trzykrotnie, aby uzyskać dokładne i spójne wyniki oraz uwzględnić zmienność warunków środowiskowych.

PRODUKT KOŃCOWY – ANALIZA PORÓWNAWCZA

Podsumowując wyniki testów porównawczych dotyczących maszyn wirtualnych, można wyróżnić następujące wnioski:

- AWS osiągnął najlepszą wydajność CPU i dysku.
- W przypadku wydajności RAM, wyniki były zbliżone pomiędzy AWS, a Azure dla mniejszych instancji, natomiast dla większych instancji AWS wykazał lepszą wydajność.
- AWS uzyskał lepszą prędkość pobierania, podczas gdy Azure osiągnął lepszą prędkość wysyłania.
- AWS c5.large wyróżniał się najmniejszą różnicą między wynikami testów przeprowadzonych w godzinach pracy, a poza nimi.
- Azure F2s_v2 był najtańszym wyborem wśród maszyn wirtualnych.
- W przypadku prędkości sieci w przeliczeniu na GB, Azure również oferował tańsze rozwiązanie.

Podsumowując wyniki wszystkich przeprowadzonych testów, można także zauważyć, że istnieje różnica w wydajności maszyn wirtualnych w zależności od tego, czy testy były wykonywane w godzinach pracy czy poza nimi. Najbardziej widoczne było to dla maszyn Azure F2s_v2 i F4S_v2, które wykazywały lepszą wydajność w godzinach pracy, oraz dla maszyny AWS c5.large, która miała lepsze wyniki poza godzinami biznesowymi. Różnice w wynikach oscylowały w granicach 1-2%, ale ponieważ testy



Wyższa Szkoła Technologii Informatycznych w Katowicach

były powtarzane, a wyniki były tożsame, można stwierdzić, że konkurencja o zasoby w chmurze jest obecna, choć ma ona niewielki wpływ na ogólną wydajność maszyn.

Badanie to dostarczyło klarownego obrazu dla przyszłych badaczy, pracowników w dziedzinie chmury oraz interesariuszy firm, pomagając im w podejmowaniu decyzji dotyczących wyboru dostawcy usług chmurowych z uwzględnieniem wydajności sprzętu, szybkości sieci oraz kosztów korzystania z tych usług. Takie informacje mogą przekładać się na znaczne oszczędności finansowe, zarówno na poziomie indywidualnych użytkowników, jak i przedsiębiorstw, które mogą zaoszczędzić setki, tysiące lub nawet miliony dolarów miesięcznie lub rocznie.

W celu dalszego rozwoju badań, warto rozszerzyć analizę kosztów na inne regiony geograficzne, ponieważ ceny maszyn wirtualnych mogą się różnić w zależności od lokalizacji. Możliwość wdrażania maszyn wirtualnych na całym świecie daje przedsiębiorstwom możliwość optymalizacji kosztów poprzez wybór regionów o bardziej atrakcyjnych cenach. Dzięki temu przedsiębiorstwa mogą zidentyfikować regiony, w których koszty są niższe i dostosować swoje strategie chmurowe w celu maksymalizacji oszczędności. Ponadto, rozszerzenie badań na różne regiony geograficzne pozwoliłoby na bardziej wszechstronne porównanie wydajności i kosztów między dostawcami usług chmurowych na skalę globalną.