

Temat Projektu:

"Porównanie środowisk bazodanowych MySQL pod kątem operacji wykonywanych na bazie danych"

Stachiewicz Dawid 173218, Sulisz Szymon 173221

Opiekun pracy:

Mgr. Inż. Dariusz Rączka

Spis treści

1.	Ws	tęp		1
	1.1	Wpi	rowadzenie do tematu	1
	1.2	Cel	i zakres projektu:	1
2.	Ro	zwinię	cie	2
	2.1	SQI	_ Server Management Studio	2
	2.2	Mys	SQL – MariaDB	3
	2.3	Sch	emat bazy danych wykorzystanej w projekcie	5
	2.3.1 Dane		ne przechowywane w poszczególnych tabelach	7
	2.4	Imp	ort danych	.11
	2.4	.1	Dla SSMS	.12
	2.4	1.2	Dla MariaDB	.14
	2.5	Wn	ioski z procesu importowania bazy danych	.15
	2.5	5.1	SSMS	.15
	2.5.2		MariaDB	.15
3.	Por	ównai	nie pracy obu serwerów podczas ich obciążania	.16
	3.1	Pole	ecenia jakim baza danych była obciążana	.16
	3.2	Por	ównanie wyników podczas obciążania serwera	.18
	3.2	2.1	Sposób zbierania danych na temat obciążenia	.18
	3.2	2.2	Porównanie wyników wydajnościowych	.19
	3.2	2.3	Podsumowanie	.23
4.	Bib	oliogra	fia	.24
5.	Spi	s rysu	nków	.25
6.	Spi	s wvk	resów	.26

1. Wstęp

1.1 Wprowadzenie do tematu.

W dzisiejszym dynamicznym świecie informatyki, w którym dane stanowią kluczowy zasób, wybór odpowiedniego środowiska bazodanowego jest niezwykle istotny dla efektywnego zarządzania informacjami oraz sprawnego działania aplikacji. W kontekście projektu dotyczącego administracji systemów rozproszonych, badanie i porównanie dwóch popularnych platform - SQL Server Management Studio oraz MySQL-MariaDB - staje się kluczowym krokiem w procesie zrozumienia ich charakterystyk, możliwości i ograniczeń. Przeprowadzenie takiego porównania pozwala zidentyfikować różnice między tymi rozwiązaniami oraz określić, które z nich lepiej odpowiadają na konkretne potrzeby i wymagania projektów rozproszonych.



Rysunek 1 Środowisko bazodanowe MySQL

1.2 Cel i zakres projektu:

Celem niniejszej pracy jest przeprowadzenie kompleksowego porównania obu tych środowisk z uwzględnieniem ich technicznych aspektów, funkcjonalności oraz wydajności w kontekście istniejącej bazy danych. Projekt rozpoczyna się od importowania istniejącej bazy danych do obu środowisk, zapewniając spójność danych i struktury. Następnie przeprowadzane są testy wydajnościowe, które obejmują różnorodne scenariusze obciążeniowe, zapytania i operacje na danych. Testy te mają na celu ocenę wydajności, skalowalności i stabilności obu platform w praktycznych warunkach. Ostatecznie, wyniki testów są analizowane i porównywane, co umożliwia wyciągnięcie wniosków dotyczących

efektywności i przydatności w kontekście konkretnej aplikacji lub projektu rozproszonego. Poprzez to podejście projektowe możliwe będzie lepsze zrozumienie różnic między MySQL Server Management Studio a MariaDB oraz wybór optymalnego rozwiązania dla konkretnych potrzeb projektowych.

2. Rozwinięcie

2.1 SQL Server Management Studio

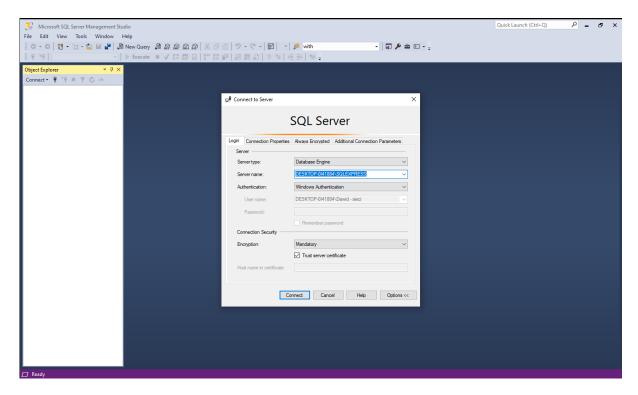
SQL Server Management Studio (SSMS) to narzędzie stworzone przez firmę Microsoft, zaprojektowane do zarządzania bazami danych działającymi na platformie Microsoft SQL Server. Jest to kompleksowe środowisko do zarządzania bazami danych, które umożliwia administratorom i programistom wykonywanie szeregu operacji administracyjnych, jak również rozwijanie, monitorowanie i optymalizację baz danych.

SSMS umożliwia wykonywanie różnorodnych zadań, takich jak:

- Tworzenie i edycja baz danych oraz tabel.
- Wykonywanie zapytań SQL do przeglądania, edycji i analizowania danych.
- Zarządzanie uprawnieniami dostępu do bazy danych.
- Tworzenie i zarządzanie procedurami składowanymi, funkcjami, wyzwalaczami i widokami.
- Monitorowanie wydajności bazy danych poprzez analizę planów wykonania zapytań, raporty o obciążeniu systemu itp.
- Wykonywanie operacji kopii zapasowych i przywracania danych.
- Konfigurowanie i monitorowanie replikacji danych.
- Rozwiązywanie problemów i diagnozowanie błędów w bazie danych.

My korzystamy z SSMS głównie w celu zarządzania bazą danych SQL Server, importowania danych, przeprowadzania zapytań SQL oraz monitorowania i optymalizowania wydajności bazy danych.

Aby rozpocząć korzystanie z SSMS, najpierw pobraliśmy aplikację odwiedzając oficjalną stronę internetową firmy Microsoft i wybierając odpowiednią wersję SSMS zgodną z naszym systemem operacyjnym jakim jest Windows 10 Pro. Po pobraniu pliku instalacyjnego, postępowaliśmy zgodnie z krokami instalacyjnymi.



Rysunek 2 Zrzut ekranu skonfigurowanego serwera MySQL

2.2 MySQL – MariaDB

MySQL i MariaDB to otwarte źródłowe systemy zarządzania relacyjnymi bazami danych (RDBMS), które cieszą się ogromną popularnością wśród deweloperów i przedsiębiorstw. MariaDB jest forkem MySQLa, co oznacza, że bazuje na kodzie źródłowym MySQLa, ale także wprowadza własne ulepszenia i nowe funkcje.

Te systemy bazodanowe są wykorzystywane do przechowywania, zarządzania i dostępu do danych w sposób zorganizowany, relacyjny i bezpieczny. Mogą być stosowane w różnych aplikacjach i scenariuszach, od małych witryn internetowych po duże aplikacje przedsiębiorstwowe.

MySQLa/MariaDB wykorzystujemy głównie do:

 Przechowywania danych: Dzięki relacyjnej strukturze danych, MySQL/MariaDB są idealnym rozwiązaniem do przechowywania danych w tabelach, zapewniając łatwy dostęp i wysoką wydajność.

- Zarządzania danymi: Oferują zaawansowane narzędzia do tworzenia, modyfikowania i usuwania danych, zapewniając pełną kontrolę nad informacjami przechowywanymi w bazie danych.
- Wykonywania zapytań: Pozwalają na wykonywanie złożonych zapytań SQL w celu przeglądania, aktualizacji, wstawiania i usuwania danych z bazy danych.
- Skalowania aplikacji: MySQL/MariaDB są elastyczne i łatwe w skalowaniu, co pozwala na obsługę rosnącej liczby użytkowników i danych w aplikacji.
- Zapewniania bezpieczeństwa: Oferują mechanizmy autoryzacji, uwierzytelniania i kontroli dostępu, aby chronić dane przed nieautoryzowanym dostępem i zagrożeniami bezpieczeństwa.
- Optymalizacji wydajności: Posiadają narzędzia do optymalizacji wydajności, takie jak indeksowanie danych, partycjonowanie tabel, analiza planów wykonania zapytań itp., co pozwala na uzyskanie szybkiego dostępu do danych nawet w przypadku dużych zbiorów danych.

Rysunek 3 Zrzut ekranu konsoli MySQL - MariaDB

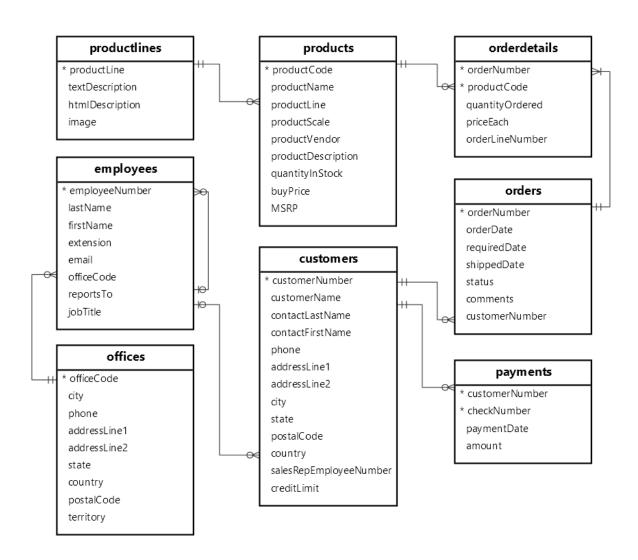
2.3 Schemat bazy danych wykorzystanej w projekcie.

W naszym projekcie używamy bazy danych "classicmodels" jako przykładowej bazy danych MySQL. Baza danych classicmodels opisuje sprzedaż detaliczną na skalę klasycznych samochodów. Zawiera typowe dane biznesowe, w tym informacje o klientach, produktach, zamówieniach sprzedaży, pozycjach zamówienia sprzedaży i wiele innych.

Schemat tej bazy danych składa się z następujących tabel:

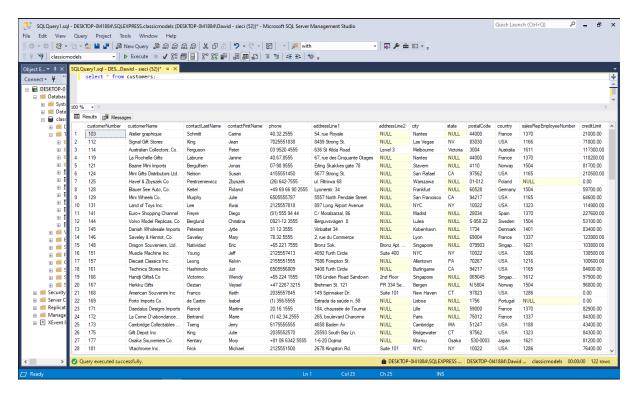
- 1) customers: przechowuje dane klientów.
- 2) **products:** przechowuje listę modeli samochodów w skali.
- 3) productlines: przechowuje listę linii produktów.
- 4) **orders**: przechowuje zamówienia sprzedaży składane przez klientów.
- 5) orderdetails: przechowuje pozycje zamówienia sprzedaży dla każdego zamówienia.
- 6) **payments:** przechowuje płatności dokonywane przez klientów na podstawie ich kont.
- 7) **employees**: przechowuje informacje o pracownikach oraz strukturę organizacyjną, np. kto jest podwładnym kogo.
- 8) offices: przechowuje dane o biurach sprzedaży.

Poniższy rysunek przedstawia diagram ER wykorzystywanej bazy:

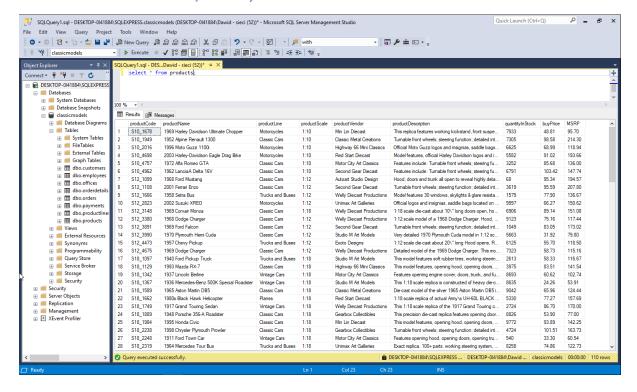


Rysunek 4 Diagram ER bazy danych "classicmodels"

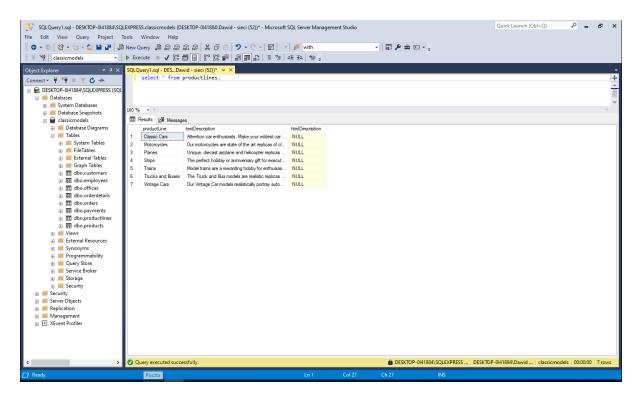
2.3.1 Dane przechowywane w poszczególnych tabelach



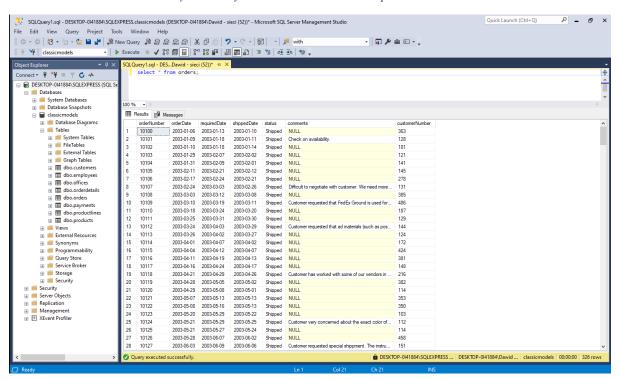
Rysunek 5 Wyświetlenie zawartości tabeli "customers"



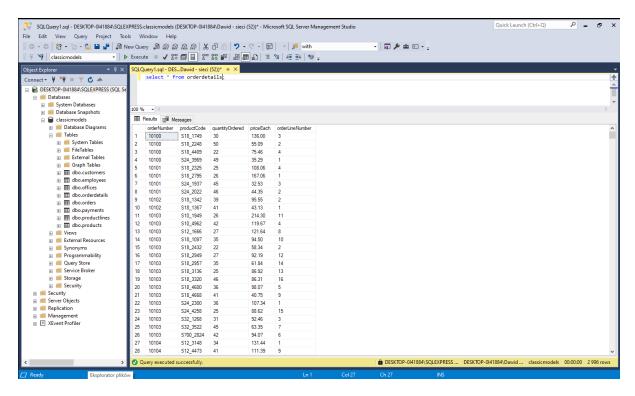
Rysunek 6 Wyświetlenie zawartości tabeli "products"



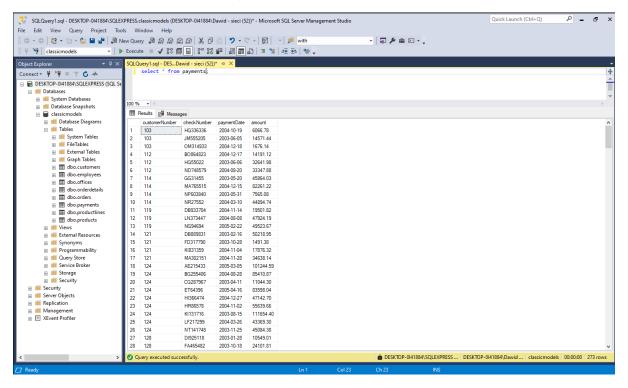
Rysunek 7 Wyświetlenie zawartości tabeli "productlines"



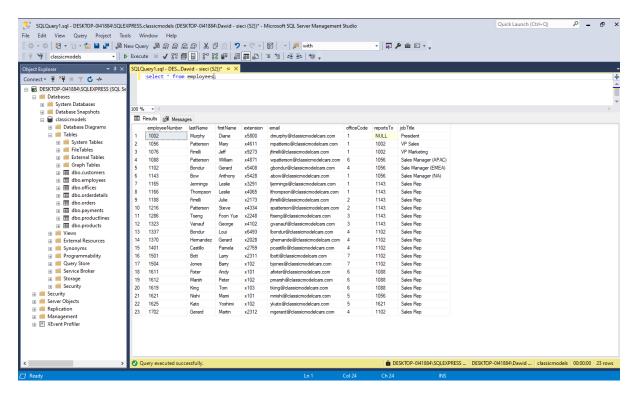
Rysunek 8 Wyświetlenie zawartości tabeli "orders"



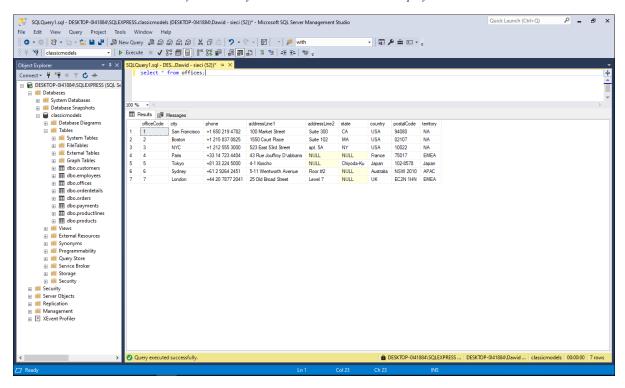
Rysunek 9 Wyświetlenie zawartości tabeli "orderdetails"



Rysunek 10 Wyświetlenie zawartości tabeli "payments"



Rysunek 11 Wyświetlenie zawartości tabeli "employees"



Rysunek 12 Wyświetlenie zawartości tabeli "offices"

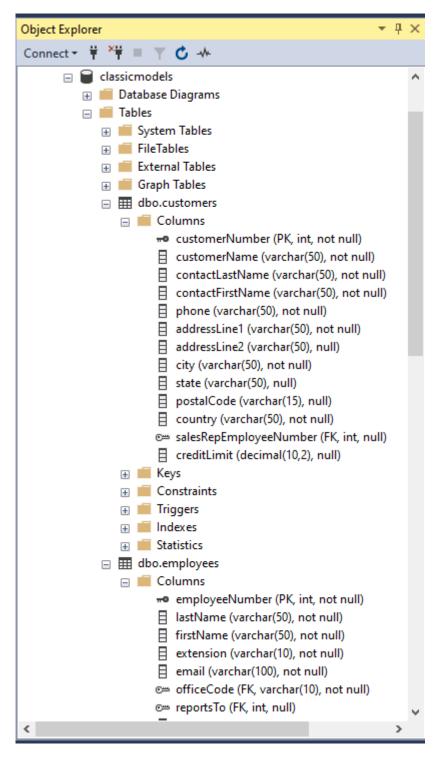
2.4 Import danych

Poniżej zaprezentowany jest fragment kody bazy "classicmodels", który zostanie zaimportowany do dwóch maszyn z przygotowanymi wcześniej środowiskami.

Rysunek 13 Fragment kodu SQL pobranej bazy danych

2.4.1 Dla SSMS

Do zaimportowania bazy danych "classicmodels" wykorzystałem wygenerowany kod SQL z pobranej bazy danych, a następnie go zmodyfikowałem, ponieważ SSMS nie rozpoznawał typu danych t.j "blob" czy "smallint".



Rysunek 14 Fragment typów zaimportowanych danych

Podczas importu danych do stworzonych wcześniej tabel również wystąpiły problemy np.

```
insert into productS(productCode, productName, productScale, productScale, productVendor, productDescription, quantityInStock, buyPrice, MSRP values
[1510_1678', '1989 Harley Davidson Ultimate Chopper', 'Notorcycles', '1110', 'Nin Lin Discast', 'Nis replica features working kickstand, front suspension, gear-shift lever, footbrake li
insert into offices(officeCode, city, phone, addressLine1, addressLine2, state, country, postalCode, territory) values
[171', 'San Francisco', '1450 82 129 4732', '100 Market Street', 'Suite 300', 'CA', 'USA', '34080', 'Nia'), '1215 837 0825', '1550 Court Place', 'Suite 102', 'MA', 'USA', '02107', 'No
insert into employees(employeeSumber, lastName, firstName, extension, email, officeCode, reportsTo_jobTitle) values
[10802, 'Marphy', 'Diane', '35080', 'dumrphy@classicondelcars.com', '1', NULL, 'President'), (1855, 'Pattersom', 'Nam', '3401', 'mpatterso@classicmodelcars.com', '1', 1002, 'VP Sales'), (1876, 'Fir
insert into customers(customerNamer, customerName, contactFirstName, phone, addressLine1, defressLine2, city, state, postalCode, country, salesRepEmployeeSumber, creditLimit _
insert into customers(customerNamer, customerName, contactFirstName, phone, addressLine1, defressLine2, city, state, postalCode, country, salesRepEmployeeSumber, creditLimit _
insert into orders(orderNumber, orderDate, requiredDate, shippedDate, status, comments, customerNumber) values
[18100, '2003-01-06', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-01-13', '2003-0
```

Rysunek 15 Problem przy poleceniu "insert"

Problem polegał na tym, że w tekście występował apostrof, który był traktowany jako koniec tekstu:

```
:ast','1:18 scale replica of actual Army\'s UH-60L BLACK HAWK Helicopter. 100% hand-as
```

Rysunek 16 Znalezienie źródła problemu

Należało usunąć zbędne apostrofy i wszystko pomyślnie zostało zaimportowane.

2.4.2 Dla MariaDB

W przypadku drugiej maszyny wystarczyło użyć polecenia source i wskazać ścieżkę do pliku, w którym znajdował się plik z bazą – w naszym przypadku było to:

Rysunek 17 Import bazy w środowisku MariaDB w wierszu poleceń

2.5 Wnioski z procesu importowania bazy danych

2.5.1 SSMS

Import danych w tym systemie był długi, problematyczny oraz niekomfortowy do codziennej pracy, ponieważ nie udało mi się znaleźć szybkiego rozwiązania istniejących problemów jak tylko ręcznie poprawiać błędy, co było bardzo czasochłonne. Nie oznacza to, że system ten jest zły, bo gdy mówimy o prezentacji zaimportowanych danych to SSMS idealnie się sprawdza, możemy czytelnie i przejrzyście odczytać co znajduje się w każdej tabeli oraz bardzo komfortowe jest okno nawigatora z lewej strony, w którym możemy podglądnąć strukturę tabel oraz ich typy, klucze czy też relacje.

2.5.2 MariaDB

Importowanie danych w środowisku MariaDB również miało swoje wyzwania, ale ogólnie proces ten był bardziej elastyczny i mniej czasochłonny niż w SSMS. Dzięki narzędziom takim jak phpMyAdmin czy HeidiSQL, możliwe było łatwiejsze zarządzanie błędami i szybkie ich rozwiązywanie. MariaDB dostarczała bardziej intuicyjne narzędzia do importu, które często automatyzowały wiele procesów.

Podczas importu za pomocą phpMyAdmin, interfejs użytkownika okazał się być bardzo przyjazny, a możliwości importu z różnych formatów plików (np. CSV, SQL) były dobrze wspierane. W przypadku napotkania błędów, system dostarczał szczegółowe komunikaty, które umożliwiały szybką diagnozę i poprawę problemów.

Z kolei HeidiSQL oferowała bogaty zestaw narzędzi, który ułatwiał importowanie dużych zbiorów danych. Funkcja podglądu i edycji danych na bieżąco była bardzo pomocna, podobnie jak możliwość łatwego definiowania kluczy głównych i obcych bez konieczności pisania złożonych zapytań SQL.

3. Porównanie pracy obu serwerów podczas ich obciążania

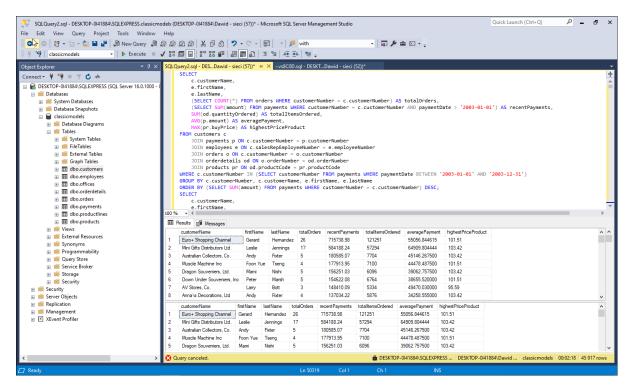
3.1 Polecenia jakim baza danych była obciążana

Poleceniem jakim została obciążona nasza baza danych była to rozbudowana kwerenda wybierająca dane z tabeli takie jak:

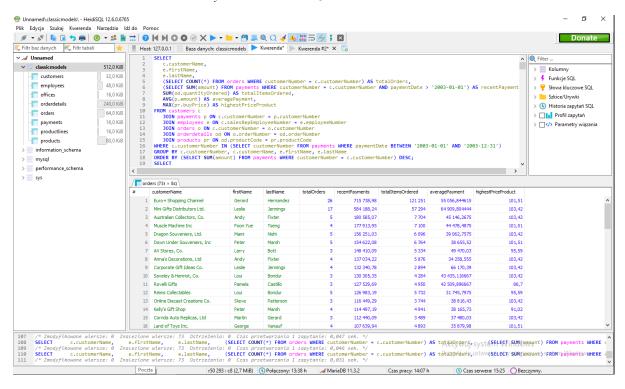
- CustomerName kolumna z tabeli
- FirstName kolumna z tabeli
- LastName kolumna z tabeli
- TotalOrders zagnieżdzona kwerenda zliczająca wszystkie zamówienia z tabeli orders dla których customersName bylo takie samo jak customersName z tabeli cusotmers
- RecentPayments zagnieżdzona kwerenda sumująca kwoty dla danego klienta zamówień opłaconych po dniu 01.01.2002
- TotalItemOrdered funkcja sumująca wszystkie ilości zamówień dla danego klienta
- AveragePayment funkcja licząca średnią kwotę za zamówienie dla danego klienta
- HighestPriceProduct funkcja wybierająca najdroższy zakupiony przez zamawiającego produkt

Kwerenda dodatkowo zawiera kilka połączeń tabel i warunek wybierający tylko tych klientów, którzy dokonali płatności pomiędzy dniem 01.01.2003, a 12.31.2003. Dane wyświetlane przez nią są grupowane poprzez CusotmerNumber, CustomerName, FirstName oraz LastName. Są one uporządkowane malejąco względem łącznych kwot zamówień.

Polecenie to zostało przedstawione na poniższych screenach zapisane już w konkretnych środowiskach bazodanowych



Rysunek 18 Kwerenda zastosowana w SSMS



Rysunek 19 Kwerenda zastosowana w HeidiSQL – środowisko MariaDB

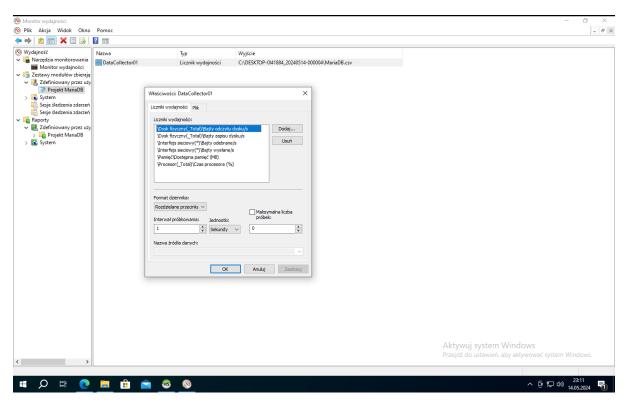
3.2 Porównanie wyników podczas obciążania serwera

3.2.1 Sposób zbierania danych na temat obciążenia

Dane na temat obciążeń kolejnych podzespołów w systemie zostały zbierane za pomocą skonfigurowanego monitora wydajności. Monitor ten co sekunde zbierał dane dotyczące taki informacji jak:

- Bajty odczytu dysku fizycznego / s
- Bajty zapisu dysku fizycznego / s
- Bajty odebrane przez interfejs sieciowy / s
- Bajty wysłane przez interfejs sieciowy / s
- Czas pracy procesora %

Poniżej pokazany jest skonfigurowany monitor wydajności dla obu maszyn wirtualnych, na których zainstalowano porównywane przez nas serwery bazodanowe

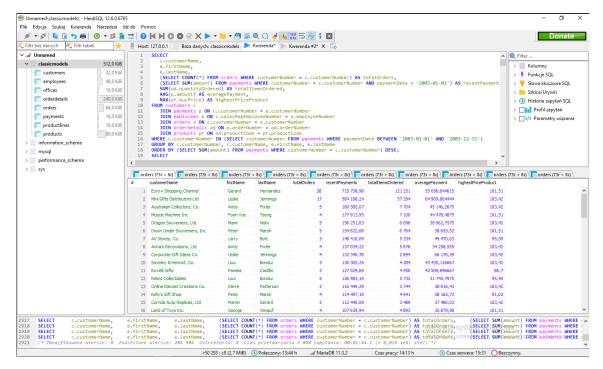


Rysunek 20 Skonfigurowany monitor wydajności

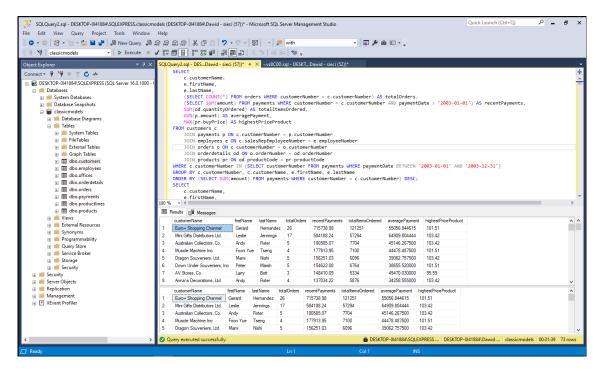
Podczas obciążania serwera polecenie zostało zapisane 2795 razy aby wywołać ciągłą pracę bazy podczas przeszukiwania danych i jak najbardziej obciążyć podzespoły aby można było sprawdzić, która aplikacja bazodanowa lepiej sobie radzi podczas pracy przy bardzo wymagających warunkach.

3.2.2 Porównanie wyników wydajnościowych

Zaczynając nasze porównanie największym parametrem, który nas zaskoczył podczas testów był czas wykonywania polecenia. Na serwerze MariaDB wynosił on nie całe 2 minuty co widać na screenie a dla drugiego serwera czas ten przekroczył ponad 30 minut. Testy były wykonywane przy takich samych warunkach pracy systemu co jeszcze bardziej zaskakuje taka wielka różnica pomiędzy oba serwerami



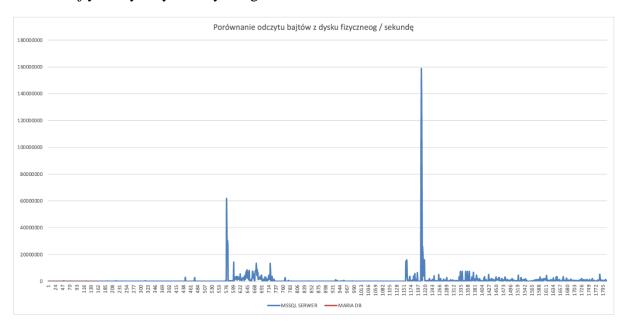
Rysunek 21 Kwerenda nr 2 w HeidiSQL



Rysunek 22 Kwerenda nr 2 w SSMS

Kolejnymi parametrami jakie porównywaliśmy były:

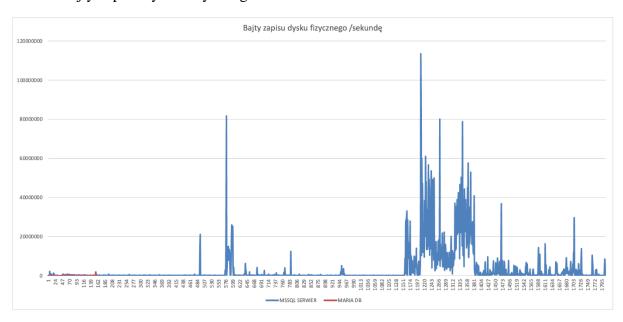
- Bajty odczytu dysku fizycznego



Wykres 1 Wykres przestawiający bajty odczytu dysku fizycznego

Widać dużo większe obciążenie dysku podczas wykonywania polecenia na MSSQL SERWER oraz widać 2 większe obciążenia w około 10 i 20 minucie.

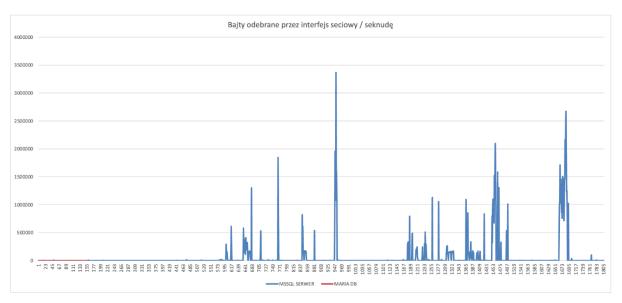
- Bajty zapisu dysku fizycznego



Wykres 2 Wykres przestawiający bajty zapisu dysku fizycznego

Możemy zaobserwować jeden wzrost wartości pod koniec wykonywania się polecenia na Maria DB oraz kilka małych i krótkich skoków i jeden duży i długi trwający około 3 minuty.

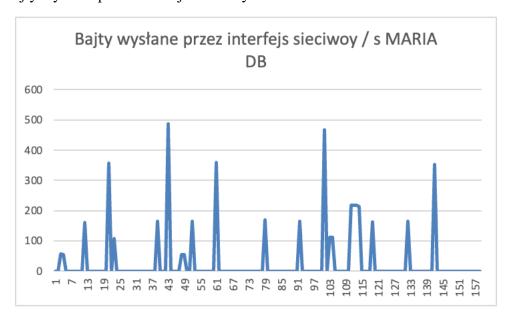
- Bajty odebrane przez interfejs sieciowy/ s



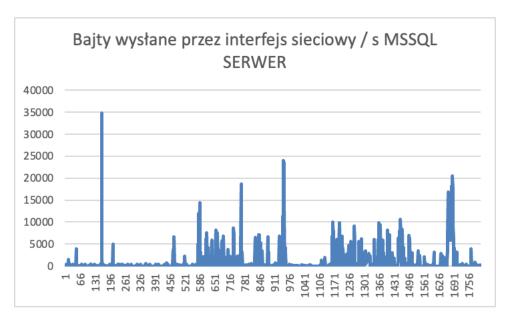
Wykres 3 -Wykres przestawiający bajty odebrane przez interfejs sieciowy/ s

Możemy zaobserwować jednostajny ruch odbierany podczas pracy Maria DB i bardzo wzrastający do wysokiej wartości i wracający do 0 w przypadku MS SQL SERWER.

- Bajty wysłane przez interfejs sieciowy /s



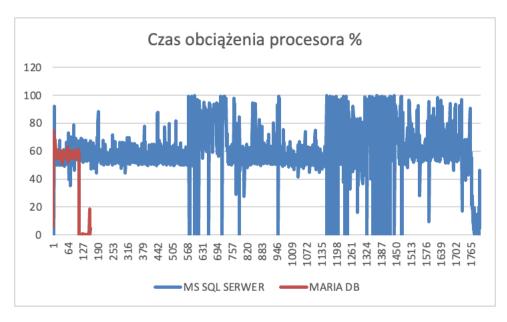
Wykres 4 Wykres przedstawiający bajty wysłane przez interfejs sieciowy /s w środowisku MariaDB



Wykres 5 Wykres przestawiający bajty wysłane przez interfejs sieciowy /s w środowisku SSMS

Na poniższych wykresach możemy zauważyć chwilowe wzrosty podczas wysyłania danych przez interfejs sieciowy jednak skala ilości bajtów podczas MS SQL SERWER jest dużo większa niż MARIA DB.

- Czas obciążenia procesora %



Wykres 6 Wykres przedstawiający czas obciążenia procesora (%)

Na tym wykresie mamy przedstawiony czas obciążenia procesora. Jak widać najwyższa wartość dla MARIA DB oscyluje w granicach 75% a średnia wartość wacha się około 50-60% więc procesor podczas tej pracy nie był bardzo obciążony. W przypadku MSSQL SERWER przez pierwsze 10 minut wartości miały trochę większą amplitudę lecz też wartości znajdowały się w granicach 50-70% dopiero od około 9 minuty wartości te zaczęły skakać od 0 do 100% zużycia procesora.

3.2.3 Podsumowanie

Podsumowując tą analizę na temat pracy działania serwerów podczas obciążania ich kwerendą wybierającą dużo lepiej wypadł serwer MariaDB. Czas pracy w jakim wykonało się zapytanie był 15 razy szybszy niż na MSSQL SERWER. Kolejnym aspektem jest obciążenie podzespołów komputerowych, na których postawiona jest nasza maszyna jest również z dużą korzyścią dla MariaDB. Nie są one używane do granic swoich możliwości dzięki czemu będą mogły dłużej bezawaryjnie cieszyć nas swoją pracą.

4. Bibliografia

- [1] Oficjalna strona MySQL MariaDB https://mariadb.org/
- [3] Dyer, Russell JT. Learning MySQL and MariaDB: Heading in the right direction with MySQL and MariaDB. "O'Reilly Media, Inc.", 2015.
- [4] Shears, Stephen. "Data Processing in a Database Management System Using Parallel Processing." (2022).
- [5] DuBois, P. (2013). MySQL. Addison-Wesley.

5. Spis rysunków

Rysunek 1 Środowisko bazodanowe MySQL	1
Rysunek 2 Zrzut ekranu skonfigurowanego serwera MySQL	3
Rysunek 3 Zrzut ekranu konsoli MySQL - MariaDB	4
Rysunek 4 Diagram ER bazy danych "classicmodels"	6
Rysunek 5 Wyświetlenie zawartości tabeli "customers"	7
Rysunek 6 Wyświetlenie zawartości tabeli "products"	7
Rysunek 7 Wyświetlenie zawartości tabeli "productlines"	8
Rysunek 8 Wyświetlenie zawartości tabeli "orders"	8
Rysunek 9 Wyświetlenie zawartości tabeli "orderdetails"	9
Rysunek 10 Wyświetlenie zawartości tabeli "payments"	
Rysunek 11 Wyświetlenie zawartości tabeli "employees"	10
Rysunek 12 Wyświetlenie zawartości tabeli "offices"	10
Rysunek 13 Fragment kodu SQL pobranej bazy danych	11
Rysunek 14 Fragment typów zaimportowanych danych	12
Rysunek 15 Problem przy poleceniu "insert"	13
Rysunek 16 Znalezienie źródła problemu	13
Rysunek 17 Import bazy w środowisku MariaDB w wierszu poleceń	14
Rysunek 18 Kwerenda zastosowana w SSMS	17
Rysunek 19 Kwerenda zastosowana w HeidiSQL – środowisko MariaDB	17
Rysunek 20 Skonfigurowany monitor wydajności	18
Rysunek 21 Kwerenda nr 2 w HeidiSQL	19
Rysunek 22 Kwerenda nr 2 w SSMS	20

6. Spis wykresów

Wykres 1 Wykres przestawiający bajty odczytu dysku fizycznego	20
Wykres 2 Wykres przestawiający bajty zapisu dysku fizycznego	21
Wykres 3 - Wykres przestawiający bajty odebrane przez interfejs sieciowy/ s	21
Wykres 4 Wykres przedstawiający bajty wysłane przez interfejs sieciowy /s w środow	isku
MariaDB	22
Wykres 5 Wykres przestawiający bajty wysłane przez interfejs sieciowy /s w środowis	sku
SSMS	22
Wykres 6 Wykres przedstawiający czas obciążenia procesora (%)	23