

WYDAJNOŚCI ZŁĄCZEŃ ORAZ ZAPYTAŃ ZAGNIEŹDŹONYCH DLA SCHEMATÓW ZNORMALIZOWANYCH I ZDENORMALIZOWANYCH

1. Tabela stratygraficzna

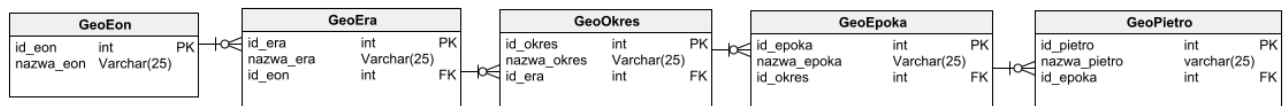
Tabela stratygraficzna – schemat obrazujący przebieg historii Ziemi na podstawie następstwa procesów geologicznych i układu warstw skalnych. Obecnie przyjęta tabela stratygraficzna została ustalona przez Międzynarodową Komisję Stratygrafii (ICS).

ERA	OKRES	EPOKA	WIEK (PIETRO)	SYMBOL	PIETRO LOKALNE	WIEK MLN	FAZA	ORO.		
K E N O Z O I K	CZWARTEDE Q	NEO PL MEZO PL EO PL	HOLOCEN	R188 R185 MINDEL-R185 MINDEL GUNZ-MINDEL GUNZ DONAU-GUNZ DONAU BIBER-DONAU BIBER	ZŁ WISŁY INT. EEMSKI ZŁ ODRY INT. MAZOWIECKI ZŁ SANU INT. MAŁOPOLSKI ZŁ NIDY INT. PODLASKI ZŁ NARWI PREGLACJAL	0,0125	WOŁOSKA ATTYCKA STYRYJSKA SAWSKA PIRENEJSKA LARAMUSKA SUBHERCYŃSKA ASTURIJSKA MŁODOGIMERYJSKA STAROGIMERYJSKA SAALSKA ASTURIJSKA SŁOUECKA BRETONSKA ERYJSKA ARDENSKA TAKOŃSKA BARDYŃSKA (SANDOMIERSKA) BAJKALSKA (GADOMUSKA) ASSYNTYJSKA DALSLANDZKA GOTYJSKA KARELSKA SAAMUSKA	A L P E J S K A		
						1,7				
						5				
						12,5				
	TRZECIODE TR	NEOGEN PALEOGEN PO	PLIOGEN	ROMAN DAK PANTON SARMAT BADEH KARPAT OTUNANG EGENBURG EODER	PL3 PL1 M3 M2 M1	MESSYN NEOT TORTON HELVET BURDYGAŁ AKWITAN EODER			17,5	
			OLIGOCEN	SZAT RUPEL	OL3 OL1	LUDYK LATTORF			22,5	
			EOCEN	PRIABON BARTON LUTET YPREZ	E3 E2 E1	SPARNAK LANDEN MONT			35	
			PALEOCEN	TANET DAN	PC3 PC1				55	
	M E Z O Z O I K	KREDA CR	GÓRNA DOLNA	MASTRICHT KAMPAN SANTON KONIAK TURON CEVOWAN	CRM CRCP CRCT ORCN CRT CRC	EMSZER			SENON CRIN	65
				ALB APT BARREM HOTERYW WALANZYN BERGAS	CRAL CRAP CRBA CRH CRV CRB	WRARON			NEOKOM	140
JURA J		GÓRNA (MALM)	TYTON KIMERID OKSFORD	JT JRM JO	INFRAWALANZYN PORTLAND, WOLG JP	210				
		SRODKOWA (DOOGER)	KELOWEJ BATON BAJOS AJALEN	JCL JET JBU JA	WEZUL	250				
		DOLNA (LIAS)	TOURNE PLENSBACH STENMUR HETANG	JTO JPL JIS JH		290				
TRIAS T		GÓRNY	RETYK NORYK KARNIK	T3	KAUPER TK	360				
		SRODKOWY	LADYN ANZYK	T2	WAPIEN MUSZLOWY	410				
		DOLNY	SCVTYK	T1	RET PSTRY PIASKOWIEC	440				
PERM P		GÓRNY	TATAR KAZAN KUNGUR	P3	TURYNG	CECHSZYN	450			
		DOLNY	KRYWICK SAKMAR ABELSK	P1	SAXON	CZERWONY SPĄGOWIEC	490			
P A L E O Z O I K	KARBON C	GÓRNY	STEFAN WESTFAL NARM	CS CW CN	PT LURALSKE PT MOSKIEWSKIE PT SASZURSKIE	SILEZ	590			
		DOLNY	WIZEN TURNERJ FAMEN FRAN	CW CT DFA DFR	SERPUCHOW ETRENT, STRUN	DINANT	650			
	DEWON D	GÓRNY	ZWYET EFEL	DOT DE	KUWIN		700			
		DOLNY	EMS ZIGEN ZEDYN	DEM DZ DGD	KOBLENC		750			
		GÓRNY	POSTLUDLOW LUDLOW	SP SLD	PODLASIE		800			
	SILUR S	DOLNY	WENLOK LANDOWER	SW SLA	SALOP WALENT		850			
		GÓRNY	ASZOL KARADOK	OC			900			
		SRODKOWY	LARDIEL LAWREN	OL OLN			950			
	ORDOWIK O	DOLNY	ARENDY TREMADOK	OAR OT			1000			
		GÓRNY	POTSDAM	CM3			1050			
SRODKOWY		AKAD	CM2			1100				
P R E K A M B R	KAMBR CM	DOLNY	GEORG	CM1			1150			
		PROTEROZOIK	WEND	EDICARIAN		PTV	1200			
			GÓRNY	RYFEJ GÓRNY			1250			
			RYFEJ SRODKOWY		PT3	1300				
			RYFEJ DOLNY			1350				
	ARCHAIK	SRODKOWY			PT2	1400				
		DOLNY			PT1	1450				
		ARCHAIK MŁODSZY			A3	1500				
		ARCHAIK STARSZY			A2	1550				
		HADAK			A1	1600				
					1650					

© BOHUN (WŁ. E. SZCZWCZYK)
WSZELKIE PRAWA ZASTRZEŻONE

2. Konstrukcja wymiaru geochronologicznego

Aby zbadać wydajności złączeń oraz zapytań zagnieżdżonych wykonany został schemat tabeli stratygraficznej w postaci znormalizowanej przedstawiony poniżej:



Rys. 1. Schemat tabeli stratygraficznej w postaci znormalizowanej

Formę zdenormalizowaną tabeli geochronologicznej osiągnięto tworząc jedną tabelę za pomocą złączenia naturalnego obejmującego wszystkie tabele zawarte w schemacie formy znormalizowanej tabeli stratygraficznej.

```
CREATE TABLE GeoTabela AS (SELECT * FROM GeoPietro NATURAL JOIN GeoEpoka NATURAL JOIN GeoOkres
NATURAL JOIN GeoEra NATURAL JOIN GeoEon);
```

GeoTabela		
id_pietro	int	PK
nazwa_pietro	Varchar(25)	
id_eon	int	
nazwa_eon	Varchar(25)	
id_era	int	
nazwa_era	Varchar(25)	
id_okres	int	
nazwa_okres	Varchar(25)	
id_epoka	int	
nazwa_epoka	Varchar(25)	



Rys. 2. Schemat tabeli stratygraficznej w postaci zdenormalizowanej

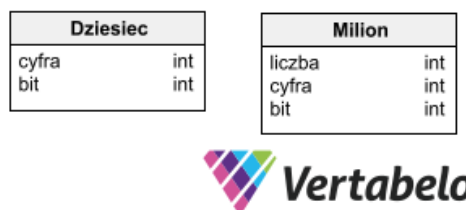
3. Testy wydajności

W testach skupiono się na porównaniu wydajności złączeń oraz zapytań zagnieżdżonych, wykonywanych na tabelach o dużej liczbie danych. W zapytaniach testowych łączono dane z tabeli geochronologicznej z danymi z tabeli Milion, wypełnionej kolejnymi liczbami naturalnymi od 0 do 999 999.

Tabela Milion została utworzona na podstawie odpowiedniego autozłączenia tabeli Dziesiec wypełnionej liczbami od 0 do 9 :

```
CREATE TABLE Milion (liczba int, cyfra int, bit int);
```

```
INSERT INTO Milion SELECT a1.cyfra +10* a2.cyfra +100*a3.cyfra + 1000*a4.cyfra + 10000*a5.cyfra +
10000*a6.cyfra AS liczba , a1.cyfra AS cyfra, a1.bit AS bit FROM Dziesiec a1, Dziesiec a2, Dziesiec a3, Dziesiec
a4, Dziesiec a5, Dziesiec a6 ;
```



Rys. 3. Schemat tabel Dziesiec i Milion

3.1. Konfiguracja sprzętowa

Testy zostały wykonane dla każdego zapytania w czterech ogólnodostępnych systemach zarządzania bazą danych na komputerze o następujących parametrach:

CPU : Intel(R) Core(TM) i5-8600K 3.60GHz,

RAM: Pamięć 12GB DDR4 (2633 MHZ),

Dysk: Samsung SSD 850 EVO 250GB

S.O.: Windows 10.

Jako systemy zarządzania bazami danych wybrano ogólnodostępne oprogramowanie:

MySQL, wersja Community Server 8.0,

PostgreSQL, wersja 14.2,

SQLite, wersja 3.38.5,

SQL Server 2019

3.2. Kryteria testów

W teście wykonano szereg zapytań sprawdzających wydajność złączeń i zagnieżdżeń z tabelą geochronologiczną w wersji zdenormalizowanej i znormalizowanej. Testy zostały wykonane w dwóch etapach:

1. Bez nałożonych indeksów z wyjątkiem kluczy głównych tabel
2. Po nałożeniu indeksów na wszystkie kolumny biorące udział w złączeniu

W tym celu zaproponowano cztery zapytania:

Zapytanie 1 (1 ZL), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej, przy czym do warunku złączenia dodano operację modulo, dopasowującą zakresy wartości złączanych kolumn:

```
SELECT COUNT(*) FROM Milion INNER JOIN GeoTabela ON (mod(Milion.liczba,68)=(GeoTabela.id_pietro));
```

Zapytanie 2 (2 ZL), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, reprezentowaną przez złączenia pięciu tabel:

```
SELECT COUNT(*) FROM Milion INNER JOIN GeoPietro ON (mod(Milion.liczba,68)=GeoPietro.id_pietro)
NATURAL JOIN GeoEpoka NATURAL JOIN GeoOkres NATURAL JOIN GeoEra NATURAL JOIN GeoEon;
```

Zapytanie 3 (3 ZG), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej, przy czym złączenie jest wykonywane poprzez zagnieżdżenie skorelowane:

```
SELECT COUNT(*) FROM Milion WHERE mod(Milion.liczba,68)= (SELECT id_pietro FROM GeoTabela WHERE
mod(Milion.liczba,68)=(id_pietro));
```

Zapytanie 4 (4 ZG), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, przy czym złączenie jest wykonywane poprzez zagnieżdżenie skorelowane, a zapytanie wewnętrzne jest złączeniem tabel poszczególnych jednostek geochronologicznych:

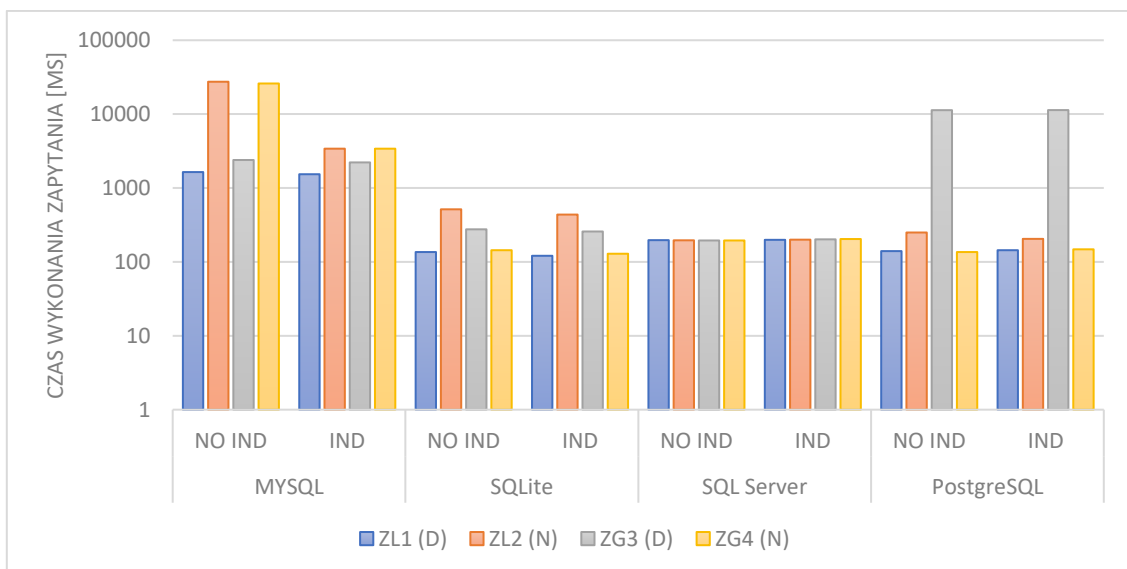
```
SELECT COUNT(*) FROM Milion WHERE mod(Milion.liczba,68) IN (SELECT GeoPietro.id_pietro FROM GeoPietro
NATURAL JOIN GeoEpoka NATURAL JOIN GeoOkres NATURAL JOIN GeoEra NATURAL JOIN GeoEon);
```

4. Wyniki

Wszystkie testy zostały wykonane po piętnaście razy z pominięciem skrajnych wyników. Następnie wyciągnięte zostały średnia oraz minimum z każdego testu. Wyniki przedstawione są w tabeli poniżej:

Zapytanie	MYSQL				SQLite				SQL Server				PostgreSQL			
	BEZ INDEKSÓW		Z INDEKSAMI		BEZ INDEKSÓW		Z INDEKSAMI		BEZ INDEKSÓW		Z INDEKSAMI		BEZ INDEKSÓW		Z INDEKSAMI	
	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG
ZL1	1609	1644	1516	1537	134	136	120	121	191	197	196	199	134	140	136	144
ZL2	27093	27451	3375	3409	508	515	433	437	189	196	196	200	235	250	195	205
ZG3	2343	2395	2203	2221	274	276	256	258	191	195	195	202	11238	11313	11246	11348
ZG4	25750	25981	3375	3408	142	144	128	129	188	195	195	204	121	136	138	148

Tab. 2. Tabela z wynikami dla poszczególnych systemów



Rys. 4. Wykres wyników z uwzględnieniem normalizacji i indeksowania

5. Wnioski

- W znacznej większości przypadków postać zdenormalizowana jest szybsza niż postać znormalizowana. Wyjątkami są: zapytanie zagnieżdżone ZG3 w systemie PostgreSQL oraz zapytanie zagnieżdżone ZG4 w systemie SQLite
- W MySQL nałożenie indeksów znacznie przyspieszyło wykonanie zapytań w postaci znormalizowanej (ZL2 i ZG4), natomiast wykonanie zapytań w postaci zdenormalizowanej zostało przyspieszone nieznacznie.
- W SQL Server nie zauważono znacznej różnicy w czasie wykonania zapytań po nałożeniu indeksów, wykonania wszystkich zapytań zostały nieznacznie spowolnione.
- Najbardziej wydajnymi systemami zarządzania baz danych dla wykonanych zapytań były SQL Server i SQLite, przy czym SQL Server zanotował największą regularność w wynikach pomiędzy zapytaniami.
- W PostgreSQL zauważono nieznaczne spowolnienie wykonania zapytań po nałożeniu indeksów, z wyjątkiem złączenia ZL2, gdzie tabela stratygraficzna występowała w postaci znormalizowanej.
- Najwolniej wykonywane są zapytania w postaci znormalizowanej (ZL2 i ZG4) w systemie MySQL oraz ZG3 w systemie PostgreSQL
- MySQL okazał się najmniej wydajnym systemem zarządzania bazami danych w przeprowadzonych testach.
- Pomędzy systemami PostgreSQL, MySQL i SQLite kod był jednolity, w SQL Server natomiast, kod musiał zostać znacznie zmieniony, aby wykonać wyżej pokazane testy.
- MySQL charakteryzuje się dużą zgodnością kolejnych prób

Podsumowując, w większości przypadków postać zdenormalizowana jest wydajniejsza niż postać znormalizowana. Nie można jednak zapominać o zaletach normalizacji, która wprowadza porządek do analizowanego schematu. W znacznej większości nałożenie indeksów skraca czas wykonania zapytań. Najbardziej wydajnym z testowanych systemów zarządzania bazami danych okazał się SQL Server, który charakteryzował się największą regularnością w czasach wykonania zapytań, bez względu na normalizację czy typ zapytania. Minusem tego systemu jest jednak różniaca się od pozostałych systemów składnia.

6. Bibliografia

- Adam Piórkowski, Łukasz Jajeńska – „Wydajności złączeń oraz zapytań zagnieżdżonych dla schematów znormalizowanych i zdenormalizowanych”
- <https://vertabelo.com>
- https://geotyda.pl/tabela_stratygraficzna.php
- https://pl.wikipedia.org/wiki/Tabela_stratygraficzna