

Artur WILCZEK  
Politechnika Wrocławska, Instytut Informatyki

## SKŁADOWANIE DANYCH W PAMIĘCI FLASH - STUDIUM PRZYPADKU

**Streszczenie.** Klasyczne, magnetyczne nośniki pamięci masowej stosowane w systemach komputerowych są coraz częściej zastępowane układami pamięci flash. Dyski SSD (ang. Solid State Disc) zbudowane w oparciu o pamięci flash zastępują dyski twarde HDD (ang. Hard Disk Drive). Dyski SSD osiągają zdecydowanie krótsze czasy dostępu, kilkudziesięciokrotnie niższe w porównaniu do dysków HDD przy odczycie losowym. Niestety, odczyt sekwencyjny, a szczególnie zapis danych na dyskach SSD nie jest już równie szybki. Praca zawiera analizę zastosowania pamięci flash do składowania danych w przykładowej konfiguracji systemu zarządzania bazą danych.

**Słowa kluczowe.** Wydajność baz danych, składowanie danych, pamięci flash

## FLASH MEMORY DATA STORAGE – CASE STUDY

**Summary.** Flash memory replaces classic, magnetic mass memory storage used in computer systems. SDD (Solid State Disc) disks based on flash memory replace hard discs drives (HDD). SDD discs offer much shorter access times, tens times shorter in respect to HDD discs where random access is considered. Unfortunately sequential access and write operations are not so effective. This work aims to analyze profits of using flash memory storage in sample database system configuration.

**Keywords:** database efficiency, data storage, flash memories

### 1. Wprowadzenie

W aplikacjach systemów zarządzania bazami danych duże znaczenie ma efektywność zapisu danych, jak i zewnętrznej, rozumiana zarówno jako krótki czas dostępu do danych ale także jako niski koszt składowania danych. Dyski twarde (HDD, ang. *Hard Disk Drives*) są podstawowym rodzajem urządzeń stosowanych do składowania danych. Przez ponad 40 lat (pierwsze

dyski twarde zaczęto wprowadzać w końcu lat 50. ubiegłego stulecia) nie pojawiły się urządzenia, które mogłyby zagrozić pozycji dysków twardych. Od połowy lat 90. można zaobserwować istotny wzrost popularności urządzeń opartych na pamięci flash – dysków SDD (ang. *Solid State Disc*). Obniżenie kosztu wytworzenia tego rodzaju rządzeń, zwiększenie ich pojemności oraz poprawa niezawodności sprawiają, że aktualnie dyski SDD stosowane są także w popularnych systemach komputerowych.

Efektywne składowanie danych ma szczególnie duże znaczenia w aplikacjach systemów zarządzania bazami danych (SZBD). Skrócenie czasu odczytu i zapisu danych charakteryzujące pamięci flash pozwala oczekiwać, że ich zastosowanie, w warstwie składowania danych systemu zarządzania bazą danych, przyczyni się do istotnej poprawy wydajności systemu. Pamiętać należy jednak, że współczesne SZBD powstawały z uwzględnieniem charakterystyki nośników magnetycznych. Stosowane są mechanizmy buforowania danych w pamięci operacyjnej, dane są odczytywane i zapisywane sekwencyjnie, pozwala to na optymalne wykorzystanie nośników magnetycznych.

W niniejszej pracy omówiono wybrane własności pamięci flash, mające decydujący wpływ na ewentualne korzyści wynikające z jej zastosowania w warstwie składowania danych SZBD. Następnie przedstawiono wyniki badań przeprowadzonych w oparciu o przykładową konfigurację SZBD.

## 2. Charakterystyka pamięci flash

Termin „pamięć flash” jest powszechnie używany jako nazwa pamięci stałej zbudowanej z układów EEPROM (ang. *Electrically Erasable Programmable Read Only Memory*). Układy flash wywodzą się od programowalnych pamięci stałych ROM wykorzystywanych np. do zapisywania systemów BIOS. Aktualnie pamięci flash kojarzone są przede wszystkim z miniaturowymi, przenośnymi nośnikami danych, stosowanymi np. do zapisu plików multimedialnych w telefonach komórkowych i aparatach fotograficznych.

Pamięci flash charakteryzuje odmienny niż w przypadku pamięci RAM sposób organizacji dostępu do poszczególnych komórek pamięci. Aby zachować nowe dane w zapisanej wcześniej komórce pamięci flash, należy ją najpierw wykasować. O ile operacje odczytu i zapisu danych mogą dotyczyć dowolnej komórki pamięci, operacja kasowania danych dotyczy całych bloków pamięci – blok pamięci flash stanowi jednostkę kasowania danych. Komórki pamięci flash charakteryzuje ograniczona liczba operacji kasowania danych. Przekroczenie tej liczby powoduje uszkodzenie pamięci.

## 2.1. Pamięci NOR i NAND

Pierwsze pamięci flash pojawiły się około 1980 roku w laboratoriach firmy Toshiba. W 1984 roku wynalazca pamięci flash dr Fujio Masuoka zaprezentował wynik swoich prac na IEEE International Electro Devices Meeting (IEDM) w San Francisco [1]. Zaprezentowano wówczas pierwszy rodzaj pamięci flash - pamięci NOR. Trzy lata później, także w trakcie konferencji IEDM, zaprezentowano układy NAND [2]. Nazwy rodzajów pamięci flash pochodzą od nazw bramek logicznych, które przypominają układ i wzajemne połączenia komórek pamięci.

W pamięciach NOR każda komórka pamięci (pojedynczy bit) może być odczytywana, zapisywana i kasowana pojedynczo. W nowszych pamięciach typu NOR komórki pamięci tworzą sektory lub bloki, które kasowane są w całości, natomiast zapis bądź odczyt danych dotyczy bajtów lub słów.

W pamięciach typu NAND grupy komórek pamięci są połączone szeregowo, co znacznie obniża liczbę połączeń pomiędzy nimi. Połączenia pomiędzy komórkami pamięci zajmowały znaczną część powierzchni układu pamięci, dlatego pamięci NAND z mniejszą liczbą połączeń cechuje niższy koszt wytworzenia układu i większa pojemność [3]. Wadą pamięci NAND jest wolniejszy odczyt niż w przypadku pamięci NOR.

Bloki pamięci NAND składają się ze stron, które są jednostką odczytu i zapisu danych. Nadal, podobnie jak w przypadku pamięci NOR, blok jest jednostką kasowania danych. Dodatkowo w przypadku pamięci NAND zapis stron w bloku musi być wykonywany sekwencyjnie.

## 2.2. Pamięci SLC i MLC

Kolejną innowacją w konstrukcji pamięci flash było wprowadzenie układów MLC (ang. *Multi Level Cell*). W układach MLC, zarówno NOR jak i NAND, pojedyncza komórka pamięci może przechowywać więcej niż jeden bit. Układy SLC (ang. *Single Level Cell*) przechowują w komórce pamięci wyłącznie jeden bit. Układy pamięci wykorzystujące technologię SLC są szybsze, a przy tym bardziej niezawodne od układów MLC. Ponadto liczba cykli kasowania pamięci zbudowanych w technologii MLC jest zdecydowanie mniejsza od tych, które zostały wykonane w technologii SLC [4].

## 2.3. Funkcje sterowników pamięci

Producenci nośników pamięci zbudowanych w oparciu o pamięci flash wprowadzają w sterownikach tych urządzeń rozwiązania zwiększające ich niezawodność. Sumy kontrolne pozwalają na detekcję błędów zapisu lub odczytu danych. Błędne bloki danych są oznaczane i nie wykorzystywane ponownie. Algorytmy równoważenia zużycia komórek pamięci (ang. *wear-leveling*) mają za zadanie równomiernie zapisywać dane we wszystkich komórkach pamięci urządzenia, tak aby uniknąć szybszego zużycia pewnych komórek pamięci (najczęściej kasowa-

nych). Jak podają producenci [5], trwałość pojedynczej komórki pamięci osiąga w przypadku pamięci NAND SLC 100 000 cykli kasowania danych, dla układów NAND MLC około 10 000 cykli.

## 2.4. Dyski SDD

Dyski SDD (ang. *Solid State Disks*) to urządzenia pamięci masowej zbudowane w oparciu o pamięci flash. Popularne urządzenia tego typu wytwarzane są przy wykorzystaniu układów NAND MLC, co pozwala obniżyć koszty ich wytwarzania. Podstawową zaletą dysków SSD, w porównaniu do klasycznych dysków twardych, jest brak ruchomych części. Sprawia to, że dyski te charakteryzują się cichą pracą oraz o wiele większą odpornością na ewentualne uszkodzenia mechaniczne. W celu zachowania kompatybilności z dyskami HDD, dyski SSD posiadają identyczne interfejsy jak dyski twarde (najczęściej SATA), i można je montować w systemach komputerowych zamiennie z dyskami HDD.

## 3. Składowanie danych na dysku SDD - Studium przypadku

Nośniki pamięci zewnętrznej oparte na pamięciach flash mają istotnie większą wydajność w porównaniu do nośników magnetycznych. Możemy zatem oczekiwać, że wykorzystanie dysku SDD do składowania danych, w miejsce dysku HDD, powinno poprawić wydajność SZBD. Takie ogólne sformułowania mogą być wyłącznie wskazówkami przy podejmowaniu konkretnych decyzji projektowych dotyczących architektury sprzętowej systemu bazodanowego. Aby poprzeć te przypuszczenia, należałoby przeprowadzić szersze badania dotyczące wpływu zastosowania urządzeń SDD na wydajność SZBD. W ramach niniejszej pracy przedstawiono wyniki badań przeprowadzonych na konkretnej konfiguracji wybranego SZBD, w której zastąpiono dysk HDD urządzeniem SDD. Zaproponowaną metodę badania wydajności operacji wejścia-wyjścia można z powodzeniem wykorzystać przy badaniu innych urządzeń. Zważywszy na to, że aktualnie dostępne są różne typy urządzeń SDD produkowane przez wielu dostawców, a oferta dysków HDD jest jeszcze bogatsza, wyniki przeprowadzonych testów nie mogą być podstawą formułowania ogólnych ocen, stanowią studium konkretnego przypadku.

### 3.1. Konfiguracja środowiska testowego

Badania przeprowadzono na SZBD Oracle 10g w wersji 10.2.0 zainstalowanym na systemie komputerowym o następujących parametrach:

- Procesor: Intel Core 2, T5600,
- Pamięć operacyjna: 2 GB.

Badania polegały na wykonywaniu zapytań SQL przygotowanych tak, aby powodowały wywoływanie operacji wejścia-wyjścia (w skrócie we-wy), czyli operacji odczytu, zapisu i kasowania danych w pamięci zewnętrznej. Badano dwie konfiguracje systemu różniące się wyłącznie w warstwie składowania danych:

- Konfiguracja A: Składowanie bazy danych na dysku HDD,
- Konfiguracja B: Składowanie bazy danych na dysku SDD.

W celu zachowania obiektywnych warunków przeprowadzenia badań wykorzystano w obu konfiguracjach urządzenia wykorzystujące połączenia do systemu komputerowego poprzez interfejs USB 2.0:

- W konfiguracji A – przenośny dysk HDD, 200 GB, bufor 16 MB, 7200 obr/min, SATA II, cena – 220 zł,
- W konfiguracji B – przenośny dysk SDD, 64 GB, SATA II, MLC, NAND deklarowana przez producenta dysku maksymalna prędkość zapisu – 80 Mb/s, odczytu – 170 Mb/s, cena – 450 zł.

### 3.2. Porównanie wydajności urządzeń

Tabela 1

Wydajność operacji we-wy SDD

Test	IOPS
1. Losowy odczyt danych o rozmiarze $\leq 64$ KB	1050
2. Losowy odczyt danych o rozmiarze $\leq 8$ MB	16
3. Sekwencyjny odczyt danych	530

Tabela 2

Wydajność operacji we-wy HDD

Test	IOPS
1. Losowy odczyt danych o rozmiarze $\leq 64$ KB	53
2. Losowy odczyt danych o rozmiarze $\leq 8$ MB	6
3. Sekwencyjny odczyt danych	450

Przed wykonaniem badań realizowanych na SZBD przeprowadzono badania wydajności urządzeń przy realizacji operacji wejścia-wyjścia o różnej charakterystyce. Celem tych badań było porównanie wydajności urządzeń dla operacji odczytu losowego i sekwencyjnego. Wyniki pomiarów liczby operacji we-wy na sekundę (IOPS) dla urządzenia SDD zawiera tabela 1, dla dysku HDD tabela 2.

Wyniki pozwalają potwierdzić zdecydowanie większą wydajność dysku SDD w przypadku operacji losowego odczytu danych. Przy odczycie sekwencyjnym przewaga dysku SDD nie jest już równie znacząca.

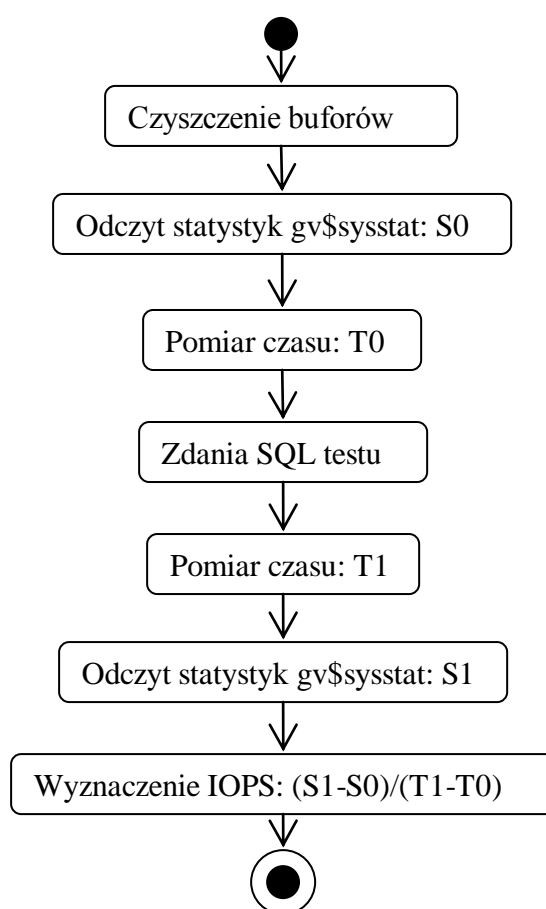
Warto także porównać, jak kształtuje się koszt składowania danych na porównywanych nośnikach. Zakup dysku HDD jest wydatkiem ponad dwukrotnie mniejszym niż zakup dysku

SDD, po uwzględnieniu pojemności urządzeń i wyznaczeniu kosztu składowania 1 GB danych, przewaga dysku HDD jeszcze się zwiększyła. Wyniki porównania zawiera tabela 3.

Tabela 3

Koszt składowania danych			
Urządzenie	Koszt [zł]	Pojemność [GB]	Koszt GB zł/GB
1. Dysk SDD: 450 zł, 64 GB	450	64	7.03
2. Dysk HDD: 220 zł, 200 GB	220	200	1.10

### 3.3. Procedura badania wydajności operacji we-wy SZBD



Rys. 1. Procedura testowa

Fig. 1. Test procedure

Współczesne systemy zarządzania bazami danych, takie jak Oracle 10g, uwzględniają charakterystykę magnetycznych nośników pamięci zewnętrznej. Wykorzystują mechanizmy buforowania danych pozwalające optymalizować wykonanie niezbędnych operacji we-wy. Obiektywne wyniki badań wydajności operacji we-wy wymagają pominięcia wpływu buforowania danych realizowanego przez SZBD. Z tego powodu przyjęto, że przed wykonaniem zdań SQL testu i rozpoczęciem pomiarów bufory SZBD będą czyszczone.

Badania wydajności operacji we-wy oparto na statystykach zbieranych przez SZBD Oracle, które są dostępne w perspektywie systemowej `gv$sysstat` [6]. Rysunek 1 prezentuje podstawowe operacje realizowane w ramach każdego z przeprowadzonych testów.

Dla każdego z przeprowadzonych testów wyznaczono liczbę operacji we-wy zrealizowanych przez SZBD w ciągu sekundy. Do przeprowadzenia obliczeń niezbędny był pomiar czasu realizacji zdań SQL wykonywanych przez SZBD oraz zebranie wartości statystyk systemu przed rozpoczęciem testu i bezpośrednio po jego zakończeniu.

### 3.4. Wyniki testów wydajności SZBD

Testy wydajności SZBD wykonano na przykładowym modelu danych. Zdania SQL przypadków testowych dotyczyły selekcji i usuwania pojedynczych wierszy tabeli oraz operacji masowego wstawiania i usuwania wierszy z tabeli testowego modelu danych. Tabela testowa zawierała 10 mln wierszy i wypełniała przestrzeń danych o rozmiarze 3,01 GB.

Zdania SQL, stanowiące testowe obciążenie systemu, dobrano następująco: 50% wierszy wstawianych w operacji insert, 50% wierszy odczytywanych w operacji select, 10% wierszy wyszukiwanych pojedynczo, 10% wierszy usuwanych pojedynczo, 50% wierszy usuwanych w operacji delete. Badania miały na celu pomiar wydajności operacji we-wy realizowanych przez SZBD, w związku z tym model danych oraz treść zapytania SQL nie mają tutaj istotnego znaczenia.

Tabela 4

Wydajność SZBD SDD	
Statystyka	IOPS
1. Odczyty blokowe	50
2. Zapisy blokowe	110
3. Odczyty wieloblokowe	1
4. Zapisy wieloblokowe	7

Statystyki dostępne w perspektywie systemowej `gv$sysstat` prezentują skumulowane liczby operacji we-wy w podziale na losowe dostępy do dysku (statystyka 1, 2) i dostępy odczytujące sekwencyjnie wiele bloków (statystyka 3, 4). Liczbę operacji we-wy na sekundę (IOPS) wyznaczono na podstawie ilorazu czasu potrzebnego na wykonanie zdań SQL testu i wartości statystyk. Wyniki badań dla urządzenia SDD zawiera tabela 4, dla dysku HDD tabela 5.

Tabela 5

Wydajność SZBD HDD	
Statystyka	IOPS
1. Odczyty blokowe	34
2. Zapisy blokowe	229
3. Odczyty wieloblokowe	4
4. Zapisy wieloblokowe	14

Wyniki nie potwierdzają oczekiwanej zdecydowanej poprawy wydajności operacji we-wy po zastosowaniu dysku SDD. Przy testowym obciążeniu bazy danych dysk SDD osiąga lepsze wyniki wyłącznie dla operacji powodujących odczyt pojedynczych bloków danych. Dla operacji wymagających masowego ładowania danych wyraźna jest przewaga dysku HDD. Podobnie dysk HDD jest bardziej wydajny w przypadku operacji zapisu danych.

Wyniki testów zdają się potwierdzać tezę, że istotna przewaga badanego dysku SDD w wydajności losowych odczytów danych (widoczna w wynikach zamieszczonych w tabeli 1 i 2) nie ma decydującego wpływu na wydajność operacji na pamięci zewnętrznej realizowanych przez SZBD. Zastosowanie dysku SDD pozwala oczekiwać poprawy wydajności dostępu do pamięci zewnętrznej jedynie przy odczycie danych.

Przeprowadzone badania nie wyczerpują możliwości wszystkich konfiguracji warstwy składowania danych SZBD nawet w badanej, prostej konfiguracji. Ciekawe wyniki może przynieść analiza wpływu składowania wyłącznie plików przeznaczonych na składowanie danych na urządzeniu SDD (szybki odczyt) przy pozostawieniu plików kontrolnych i dzienników transakcji SZBD na dysku HDD (szybki zapis). Równie interesujące może być przeniesienie segmentów tymczasowych wykorzystywanych np. podczas operacji złączeń na dysk HDD (szybki zapis) przy pozostawieniu dodatkowych plików danych na dyskach SDD (szybki odczyt). Badania w tym zakresie będą przedmiotem dalszych prac.

#### **4. Podsumowanie**

W pracy przedstawiono badania wydajności składowania danych przez SZBD na urządzeniu SDD w porównaniu do dysku twardego HDD. Badania przeprowadzono na popularnych urządzeniach dostępnych aktualnie na rynku.

Stwierdzono wielokrotną różnicę na korzyść dysków SDD w wydajności losowych odczytów danych w porównaniu do dysku HDD. Przewaga ta nie przekłada się jednak na równie znaczącą poprawę wydajności odczytu danych w SZBD. Operacje dotyczące masowego zapisywania i kasowania danych są realizowane przez SZBD bardziej wydajnie na dyskach HDD.

Istotnemu zwiększeniu kosztów składowania danych w przypadku zastosowania dysku SDD (dla badanej konfiguracji sprzętowej jest to różnica siedmiokrotna) nie towarzyszy równie zdecydowana poprawa wydajności.

Z pewnością wyników badań przeprowadzonych na jednej, wybranej konfiguracji sprzętowej nie można uogólniać. Wyniki pracy mogą być jednak cenną wskazówką dla projektantów rozwiązań małej i średniej skali opartych na popularnych nośnikach pamięci.



**BIBLIOGRAFIA**

1. Masuoka F., Asano M., Iwahashi H., Komuro T., Tanaka S.: A New Flash EPROM Cell Using Triple Polysilicon Technology. Proceedings of the International Electron Devices Meeting, Vol. 30, 1984.
2. Masuoka F., Momodomi M., Iwata Y., Shirota R.: New Ultra High Density EPROM and Flash EEPROM with NAND Structure Cell. IEDM Technical Digest, 1987.
3. Lai S. K.: Flash memories: successes and challenges. IBM Journal of Research and Development. Vol. 52, Issue 4, 2008.
4. Jung T.-S., Choi Y.-J., Suh K.-D., Suh B.-H., Kim J.-K., Lim Y.-H., Koh Y.-N., et al.: A 3.3 V 128 Mb Multi-Level NAND Flash Memory for Mass Storage Applications, Solid-State Circuits Conference. Digest of Technical Papers, San Francisco 1996.
5. Kingston Technology Corporation, Flash Memory Guide, 2009.
6. Niemiec R.J.: Oracle Database 10g Performance Tuning Tips & Techniques, McGraw-Hill, 2007.

Recenzent: Dr inż. Paweł Kasprowski

Wpłynęło do Redakcji 31 stycznia 2010 r.

**Abstract**

This work presents results of experiment conducted to analyze profits of using flash memory storage in sample database system configuration. Standard DBMS system were used with the same configuration used with hard disk drives. Results prove efficiency improvement where it goes to access times, but not so obvious to justify higher costs.

**Adres**

Artur Wilczek: Politechnika Wrocławska, Instytut Informatyki, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Polska, artur.wilczek@pwr.wroc.pl.