Ilia Solodov[1]

Zabezpieczenie danych cyfrowych przechowywanych na dyskach twardych typu HDD i SSD — możliwości i problemy

We współczesnym procesie karnym coraz częściej mamy do czynienia z dowodami cyfrowymi[2]. Podobnie jak w przypadku tzw. dowodów klasycznych istnieją określone reguły i zasady postępowania z takiego rodzaju śladami[3]. Przede wszystkim należy podkreślić, że dowód cyfrowy może ulec zniszczeniu zupełnie tak samo, jak każdy inny dowód. Trwałość dowodu cyfrowego zależy od wielu czynników, m.in. od miejsca, w którym się znajduje, oraz od charakteru jego nośnika. Niezabezpieczony lub zabezpieczony w sposób nieprawidłowy ślad cyfrowy może być zniszczony jak celowo, tak i przez przypadek.

W publikacjach anglojęzycznych odnoszących się do badań sprzętu komputerowego można odnaleźć termin „kolejność zmian" (ang. *order of volatility*), tj. katalog podzespołów komputera uwzględniający takie cechy jak stabilność oraz niezmienność znajdujących się w tych podzespołach danych cyfrowych. Elementy konstrukcyjne wymienione na liście zmian zostały w określony sposób uporządkowane. Na górze znajdują się te, które są uważane za najmniej stabilne, na dole zaś umieszczono te, które są uznawane za bardziej trwałe. Obecnie ta lista wygląda następująco:

— procesor komputera (CPU — ang. *central processing unit*) oraz pamięć podręczna (ang. *cache*),
— elementy zawierające informacje o sieciowych połączeniach użytkownika i uruchamianych programach (ang. *routing table*, tj. tablica trasowania,

[1] Dr Ilia Solodov — w 2016 r. obronił pracę doktorską *Etyka biegłego w procesie karnym* na Wydziale Prawa i Administracji Uniwersytetu Warmińsko-Mazurskiego w Olsztynie. W dorobku naukowym ma jedną monografię recenzowaną, kilkanaście publikacji, w tym w czasopismach punktowanych, oraz kilka rozdziałów w recenzowanych monografiach wieloautorskich. Brał udział łącznie w trzynastu konferencjach naukowych o zasięgu międzynarodowym w Polsce, Rosji, na Litwie i Ukrainie. Aktualnie pracuje w Polsce na stanowisku technika informatyki śledczej w UratujemyTwojeDane.pl w Olsztynie. Zachował również status adwokata w Rosji. Zainteresowania: rower, pływanie, muzyka.

Adres do korespondencji: <badanie.etyka@gmail.com>.

[2] Z badan wynika, że w ponad 70% spraw karnych występują ślady cyfrowe w tej lub innej postaci. Por. W.A. Kasprzak, *Ślady cyfrowe. Studium prawno-kryminalistyczne*, Warszawa 2015, s. 206.

[3] J. Sammons, *The Primer for Getting Started in Digital Forensics*, Waltham 2015, s. 1.

ARP cache, tj. pamięć podręczna protokołu ARP, *process table*, tj. lista uruchomionych procesów, *kernel statistics*),
— pamięć operacyjna (RAM — ang. *random-access memory*),
— pamięć wirtualna,
— dane na dysku twardym,
— dane przechowywane na zdalnych urządzeniach,
— dane przechowywane na urządzeniach archiwizujących[4].

Warto zwrócić uwagę na to, że katalog może służyć realizacji wymogów dotyczących identyfikowania, gromadzenia, przejmowania i przechowywania cyfrowego materiału dowodowego. Specjalista, który jako pierwszy podejmuje działania techniczne w sytuacji zagrożenia bezpieczeństwu informacji (ang. *Digital Evidence First Responder*), musi na miejscu przeprowadzenia czynności określić optymalną kolejność zabezpieczenia znalezionego tam sprzętu komputerowego. Zgodnie z przyjętymi wymogami należy najpierw zabezpieczyć najmniej trwałe dane cyfrowe, np. znajdujące się w pamięci operacyjnej RAM, plik wymiany, aktywne procesy[5].

Z listy zmian wynika, że podstawowy (w przypadku komputerów stacjonarnych i mobilnych) nośnik informacji, którym jest dysk twardy, znajduje się na ostatnim miejscu wśród podzespołów, które należy zabezpieczyć (zdalne urządzenia oraz urządzenia archiwizujące znajdują się fizycznie poza komputerem). Niemniej jednak jest to podstawowe źródło danych cyfrowych, które mogą mieć znaczenie dla wyjaśnienia istotnych okoliczności faktycznych sprawy. Obecnie na rynku najczęściej można spotkać dwa typy dysków twardych różniących się między sobą mechanizmem działania, tj. dyski magnetyczne typu HDD (ang. *Hard Disk Drive*) oraz dyski półprzewodnikowe SSD (ang. *Solid State Drive*).

Mechanizm działania dysku typu HDD można przedstawić przez analogię z kasetami magnetofonowymi i VHS (ang. *Video Home System*), które jako nośniki informacji jeszcze nie tak dawno były bardzo popularnymi rozwiązaniami. Element roboczy kasety stanowiła taśma magnetyczna, która w trakcie nagrywania czy odtwarzania za pomocą silnika elektrycznego była przewijana z jednej szpuli na drugą, przechodząc w międzyczasie przez głowicę odczytująco-zapisującą. Kasety magnetyczne okazały się rozwiązaniem uniwersalnym, jeżeli chodzi o ofiarowane możliwości. Taśmy magnetyczne przeznaczone początkowo do zapisu i przechowywania dźwięku można było wykorzystywać również do nagrywania danych w postaci binarnej. Kasety jednak ofiarowały wyłącznie sekwencyjny dostęp do zapisanych na nich danych. Żeby odtworzyć potrzebny fragment, trzeba było najpierw przewinąć taśmę do odpowiedniego punktu, co było jednym z głównych powodów wycofania kaset z rynku w epoce cyfrowej.

[4] Por. P. Henry, *Best Practices In Digital Evidence Collection*, <https://digital-forensics.sans.org/blog/2009/09/12/best-practices-in-digital-evidence-collection>, 5 listopada 2018 r.

[5] *PN-EN ISO/IEC 27040: 2016-12. Technika informatyczna — Techniki bezpieczeństwa — Bezpieczeństwo pamięci masowych*, <https://www.gov.pl/documents/31305/0/ekspertyza_il_pib_normy_bezpieczenstwa_2016.pdf/406fb99b-5c86-afb3-4658-93c3910cdcef>, 5 listopada 2018 r.

Współczesne dyski twarde typu HDD działają na tych samych zasadach, czyli wykorzystują zjawisko indukcji magnetycznej. Analogiem taśmy magnetycznej na dyskach są tzw. ścieżki (ang. *track*) ułożone na całej powierzchni talerza w sposób podobny do pierścieni. Sam talerz jest pokryty cienką warstwą materiału magnetycznego i wykonuje się go z aluminium (w dyskach twardych w formacie 3,5 cala) lub ze szkła (w nośnikach w formacie 2,5 cala, które możemy znaleźć w laptopach czy w dyskach przenośnych)[6]. Ta teoretycznie prosta konstrukcja była stosowana już w 1956 r. w pierwszych dyskach twardych RAMAC (ang. *Random Access Method of Accounting and Control*) firmy IBM i pozwalała na bezproblemowy odczyt danych z dowolnego miejsca na talerzu, który obracał się z prędkością 1200 RMP (ang. *revolutions per minute*, tj. obrotów na minutę). We współczesnych dyskach HDD prędkość obracania się talerzy wynosi od 5000 do 7200 RPM, jeśli mówimy tu o urządzeniach klasy podstawowej, i od 10 000 do nawet 15 000 RPM w profesjonalnych, stosowanych w serwerach dyskach. Dane zapisane na ścieżce odczytuje głowica odczytująco-zapisująca, która robi to, unosząc się nad powierzchnią talerza za pomocą efektu poduszki powietrznej na wysokość kilku nanometrów, dzięki wysokiej prędkości rotacji talerza, a także specjalnym elementom aerodynamicznym. Brak kontaktu fizycznego pomiędzy bezpośrednim nośnikiem danych a głowicą umożliwia przeprowadzenie operacji odczytu-zapisu informacji z większą prędkością oraz zwiększa okres służby tych elementów.

Powierzchnia talerza magnetycznego wygląda na pierwszy rzut oka na idealnie gładką, ze względu na wyjątkowo niskie unoszenie lecącej głowicy, aczkolwiek zawsze ma nierówności. Z tego powodu metalowe ramie, na którym mocuje się głowicę, ze względu na swoją sztywność nie nadaje się do jej utrzymywania nad powierzchnią dysku. W związku z tym pomiędzy ramieniem a głowicą odczytująco-nagrywającą został umieszczony niewielki, cienki kawałek metalu o specyficznym kształcie — zwany zawieszeniem. Na jednym ramieniu mocuje się wszystkie głowice odczytująco-zapisujące, więc poruszają się one niemal synchronicznie, ale ze względu na szczególnie wysoką gęstość zapisu stosowaną we współczesnych dyskach twardych, niezależnie od ilości zainstalowanych głowic nad powierzchnią talerza, w każdym konkretnym momencie może pozycjonować się tylko jedną z nich, a więc głowice pozycjonują się z dyskiem pojedynczo. Na przeciwnym końcu ramienia znajduje się VCM (ang. *Voice Coil Motor*, tj. pozycjoner). Jest to cewka, która pracuje w stałym polu magnetycznym wytwarzanym przez parę magnesów permanentnych i przymocowanych do obudowy dysku. Będąc sztywnie zamontowany na ramieniu, VCM może przesuwać go w określony sposób, umieszczając głowice nad określonym obszarem talerza magnetycznego. Opisane komponenty znajdują się w metalowej obudowie, której zadaniem jest ochrona delikatnych części nośnika przed negatywnym wpływem czynników zewnętrznych.

[6] A.A. Mamun, G. Guo, Chao Bi, *Hard Disk Drive Mechatronics and Control*, Boca Raton 2006, s. 3, 8.

Obudowa dysku składa się z reguły z dwóch elementów, tj. stosunkowo grubej, wykonanej ze stopu metalowego części dolnej oraz cieńszej, uformowanej w określony sposób metalowej pokrywy. Części dolna i górna obudowy HDD chronią nie tylko przed negatywnymi czynnikami fizycznymi, lecz także przed przypadkowym dotknięciem przez użytkownika talerzy czy głowic. Nawet najmniejsze cząsteczki kurzu, jeśli trafią wewnątrz obudowy, mogą znaleźć się pod pracującą głowicą, a zważywszy na wyjątkowo mały dystans pomiędzy nią a talerzem (grubość poduszki powietrznej pomiędzy głowicą a obracającym się talerzem jest o wiele mniejsza niż rozmiar cząsteczek bytowego kurzu) mogą je uszkodzić. Uszkodzone głowice mogą stracić właściwości aerodynamiczne oraz zacząć dotykać powierzchni talerza, co spowoduje dalsze uszkodzenia i w konsekwencji doprowadzi do awarii dysku.

Na obudowie dysku typu HDD znajduje się moduł PCB (ang. *Printed Circuit Board*, tj. płytka drukowana) zwany także sterownikiem dyskowym. Na powierzchni PCB umieszcza się takie podzespoły elektroniczne jak: procesor, kontroler silnika, pamięć podręczna, interfejs komunikacyjny i inne, a więc w jej współczesnym kształcie jest ona prawdziwym komputerem. PCB umożliwia komunikowanie się komputera z fizycznymi elementami dysku twardego. Istotne jest to, że PCB nie zapisuje przekazywanych przez system operacyjny danych bezpośrednio na dysk, lecz poddaje je pewnej obróbce składającej się z kilku etapów. Po pierwsze, aby zmniejszyć liczbę błędów przy zapisywaniu oraz późniejszym odczytywaniu danych, przed rozpoczęciem zapisu wykonuje się ich kodowanie. W procesie zapisu urządzenie stosuje się do pięciu różnych algorytmów kodowania danych, zaś przy odczytywaniu te same algorytmy stosuje się w odwróconej kolejności[7]. Po drugie nawet dopiero wyprodukowane talerze magnetyczne posiadają określone defekty powierzchni, które producenci ujawniają w trakcie wielogodzinnego testowania dysków. Znalezione defekty zostają opisane w specjalnym, mówiąc z pewnym uproszczeniem, pliku, którym posługuje się oprogramowanie dysku. Z upływem czasu obszary talerzy z defektami mogą się poszerzać, mogą również powstawać nowe obszary, na których występuje degradacja powierzchni. Na takich obszarach nie tylko nie można zapisywać nowych danych, ale również nie zawsze można odczytać tych, które były zapisane wcześniej. Takie miejsca na talerzu noszą potoczną nazwę bad sektory (od ang. *bad sector*). Współczesne dyski twarde są zbudowane w taki sposób, aby przez cały czas zapewniały bezproblemową pracę. W związku z tym dyski posiadają swego rodzaju strefy zapasowe zlokalizowane w odrębnych obszarach talerzy. Gdy oprogramowanie dysku odnajduje bad sektor, zamiast niego jest wyznaczany normalnie działający sektor w strefie zapasowej. To powoduje, że badacz nie wie np. jaki jest adres fizyczny określonego pliku lub jego części. W rzeczywistości pliki mogą znajdować się w różnych miejscach na powierzchni talerza magnetycznego. W związku z tym należy uznać

[7] V. Kanekal, *Data Reconstruction from a Hard Disk Drive using Magnetic Force Microscopy*, San Diego 2013, s. 5.

za błędny zaprezentowany w literaturze pogląd, że miejsce przypisywane danym przez system operacyjny faktycznie odpowiada fizycznej ich lokalizacji na powierzchni dysku twardego. Prawdą natomiast jest to, że system operacyjny posługuje się przy zapisywaniu danych adresacją LBA (ang. *Logical Block Addressing*, tj. adresacja logiczna), ale fizyczny adres zapisywanym danym nie nadaje on, a układ sterowania dysku, czyli PCB.

Ważną częścią dysku twardego jest jego oprogramowanie. We współczesnych dyskach twardych znajdziemy go w dwóch miejscach — w PCB w układzie scalonym ROM (ang. *read-only memory*, tj. pamięć komputerowa przeznaczona tylko do odczytu), a także w specjalnych strefach na talerzach SA (ang. *Service Area*, tj. Strefa Serwisowa). Coraz większa gęstość zapisu danych powoduje, że zmniejszają się zarówno poszczególne ścieżki na talerzu, jak i same głowice odczytująco-zapisujące. Efektem ubocznym miniaturyzacji jest to, że na pozycję głowicy, którą zajmuje w chwili odczytu ścieżki, ma bezpośredni wpływ cały szereg czynników, takich jak: wibracje obudowy, stan elementów ruchomych dysku (np. łożyska osi silnika czy pozycjonera VCM), temperatura wewnątrz dysku (wpływa na rozszerzanie czy zmniejszanie głowicy i właściwości jej zawieszenia), ciśnienie atmosferyczne czy wilgotność powietrza powodujące zmiany w wysokości „lotu" głowicy. Znaczenie mają również mikroskopijne wady oraz różnice w budowie pojedynczych głowic, których nie da się w pełni wyeliminować przy masowej produkcji. Aby poszczególne świeże komponenty wyprodukowanego dysku mogły wspólnie i efektywnie pracować w granicach deklarowanego przez producenta okresu służby (z reguły obejmuje on czas od 3 do 6 lat[8]), końcowy produkt poddaje się wielogodzinnemu testowaniu. W jego procesie są ustalane takie parametry jak: unikatowe mikroskopijne różnice w wymiarach konkretnego kompletu głowic, charakterystyki wibracyjne zainstalowanego ramienia, defekty powierzchni magnetycznej talerza itp. Na podstawie tych danych producent wylicza tzw. parametry adaptywne, dzięki którym sterownik dysku nie tylko będzie poprawnie kierował unikatowym kompletem komponentów fizycznych danego dysku, ale też będzie przewidywał i kompensował wszelkie fluktuacje w ich pracy, zapewniając stałe funkcjonowanie całego systemu. Ten element programowy dysku twardego, z powodu coraz większej miniaturyzacji parametrów fizycznych ścieżek zapisu, staje się coraz bardziej istotny nie tylko z perspektywy normalnej pracy całego urządzenia, ale też dla zwykłego, standardowego odczytu zapisanych danych. Sama zaś procedura testowania w jej obecnym stanie określana jest przez badaczy jako hiperregulacja (ang. *hyper-tuning*). W związku z tym, że część opisanych wyżej unikatowych (dla każdego konkretnego dysku) parametrów znajduje się w module ROM, uszkodzenie tego elementu dysku może skutkować brakiem możliwości dostępu do danych, jeśli chodzi o środki standardowe, czyli wymianę poszczególnych niedziałających komponentów. Zawartość ROM,

[8] S. Anthony, *How long do hard drives actually live for*, <https://www.extremetech.com/computing/170748-how-long-do-hard-drives-actually-live-for>, 5 listopada 2018 r.

w zależności od modelu dysku, można odtworzyć lub podmienić, ale pozostałą część parametrów zapisaną w odrąbie SA można stracić na zawsze — razem z jakąkolwiek szansą na odtworzenie znajdujących się na dysku danych.

Warto zwrócić uwagę na to, że chociaż współczesne HDD są skonstruowane w taki sposób, aby można było je używać w ekstremalnych, ze względu na obciążenia wibracyjne, warunkach (w komputerach samochodowych lub w przenośnych odtwarzaczach multimedialnych, np. iPod), wciąż zdarzają się przypadki awarii tego skomplikowanego mechanizmu.

Na liście pokazującej kolejność zmian (ang. *order of volatility*) zachodzących w podzespołach komputera brakuje jednego urządzenia, które także służy do przechowywania informacji oraz w ostatnich latach zyskało na popularności. Dyski twarde typu SSD coraz częściej wykorzystuje się jako podstawowy nośnik danych zarówno w urządzeniach klasy konsumenckiej, jak i specjalistycznej (serwery). Rynek dysków SSD rośnie ekspotencjalnie. Zgodnie z prognozami już w 2021 r. sprzedaż dysków SSD przewyższy sprzedaż dysków twardych typu HDD[9]. Aktualnie istnieje ponad 80 firm, które produkują podobne nośniki[10]. Dla porównania dyski typu HDD są produkowane obecnie tylko przez 3 firmy[11]. Kilka lat temu dyski SSD wyprzedziły dyski typu HDD pod względem maksymalnej pojemności. Przodownik wśród dysków typu HDD posiada deklarowaną pojemność 15 terabajtów[12], jednak już na tę chwilę ustępuje konkurentowi z rodziny dysków SSD amerykańskiej firmy Nimbus, który dysponuje zawartością 100 terabajtów[13].

Z punktu widzenia kryminalistyka istotne jest to, że w przypadku, gdy użytkownik ma do czynienia z dyskiem półprzewodnikowym, może on o tym nawet nie wiedzieć. Na poziomie użytkownika korzystanie z dysku SSD niczym nie różni się od korzystania ze zwykłego dysku twardego, ponieważ system operacyjny definiuje nośnik jako tzw. urządzenie blokujące (ang. *block device*), tj. ciąg adresów LBA od 0 do kilku miliardów[14]. Zwracając się do każdego z nich, system odczytuje lub zapisuje znajdujące się pod tym adresem dane. Tak samo nic się nie zmienia dla użytkownika.

---

[9] *SSD и HDD-диски: динамика продаж, рыночные доли вендоров*, <https://blog.colobridge.net/2017/12/ssd_hdd_sales_dynamics_market_share_of_vendors>, 5 listopada 2018 r.

[10] *List of solid-state drive manufacturers*, <https://en.wikipedia.org/wiki/List_of_solid-state_drive_manufacturers>, 5 listopada 2018 r.

[11] *Hard disk drive*, <https://en.wikipedia.org/wiki/Hard_disk_drive#Manufacturers_and_sales>, 5 listopada2018 r.

[12] E. Stój, *WD prezentuje Ultrastar DC HC620 — HDD o pojemności 15 TB*, <https://www.purepc.pl/pamieci_masowe/wd_prezentuje_ultrastar_dc_hc620_hdd_o_pojemnosci_15_tb>, 5 listopada 2018 r.

[13] J. Porter, *Nimbus' mammoth 100TB SSD is now the world's largest storage of its kind*, <https://www.trustedreviews.com/news/100tb-ssd-nimbus-data-3429280>, 5 listopada 2018 r.

[14] I. Sestanj, *NAND Flash Data Recovery Cookbook*, s. 52, <http://adreca.net/NAND-Flash-Data-Recovery-Cookbook.pdf>, 5 listopada 2018 r.

Posługuje się on bowiem środkami systemu operacyjnego, tj. tworzeniem, modyfikacją czy też przeglądaniem plików. Jedyną zaś różnicą w przypadku korzystania z dysku SSD z perspektywy użytkownika będzie znaczący wzrost prędkości zapisu/odczytu danych.

W celu zapewnienia kompatybilności dysk SSD jest fizycznie bardzo podobny do dysku typu HDD, jeśli chodzi o wygląd zewnętrzny, nazwę czy też lokalizację w komputerze. Niemniej jednak pod względem właściwości, szczególnie trwałości znajdujących na nim danych, dysk SSD jest bardziej podobny do umieszczonej na trzeciej pozycji na liście zmian pamięci operacyjnej RAM.

Swoją nazwę dysk SSD otrzymał od dysku typu HDD, lecz w odróżnieniu od tego drugiego dysk SSD nie ma poruszających się elementów, tj. talerzy, głowicy czy ich ramion[15], o czym użytkownik dysku SSD może przekonać się osobiście. Dysk półprzewodnikowy umożliwia większą ingerencję do swego wnętrza niż klasyczny dysk typu HDD, np. z dysku SSD można zdjąć górną pokrywę, zajrzeć do jego wnętrza, co nie spowoduje jego awarii czy też modyfikacji (uszkodzenia) zapisanych na nim danych. Taki dysk będzie dalej funkcjonował, co jest skutkiem zupełnie innej zasady jego działania. Urządzenie to składa się z trzech podstawowych elementów: kontrolera pamięci, układów scalonych NAND Flash oraz pamięci operacyjnej (tzw. bufora).

Pamięć typu NAND Flash pełni rolę nośnika danych. Pierwsza cześć tej nazwy jest odwołaniem do bramki logicznej NAND, czyli *NOT AND*, i odzwierciedla istotę działania poszczególnych komórek pamięci. Druga zaś cześć nazwy w tłumaczeniu z języka angielskiego oznacza błyskawicę, co odnosi się do strony fizycznej działania półprzewodników. Wyjątkowa, w porównaniu z dyskiem typu HDD, szybkość operacji odczytu i zapisu osiągana jest dzięki połączeniu tych półprzewodników w jeden zespół, co umożliwia zapis i odczyt danych jednocześnie z kilku miejsc, a więc ich prędkości się sumują. Przy tym pojedyncze komórki pamięci działają powolnie. Wspomniane pojedyncze komórki są tranzystorami zbudowanymi na podobieństwo elementów konstrukcyjnych pamięci operacyjnej RAM, jednak dysk SSD nie traci swojej zawartości w momencie wyłączenia zasilania. Ten efekt osiągnięto dzięki wyposażeniu półprzewodników w dodatkową tzw. bramkę pływającą (ang. *floating gate*), która pozwala utrzymywać ładunek elektryczny nawet po odłączeniu elementu od źródła zasilania. Każdy z tranzystorów ma więc możliwość autonomicznego przechowywania ładunku elektrycznego. W pamięciach starego typu SLC (ang. *Single Level Cell*, tj. komórka jednopoziomowa) obecność ładunku w tranzystorze interpretowała się jako „0" w kodzie binarnym, zaś jego brak był uznawany za „1". We współczesnych pamięciach typu MLC (ang. *Multi Level Cell*, tj. komórka wielopoziomowa) czy TLC (ang. *Triple Level Cell*, tj. komórka trójpoziomowa) liczy się poziom naładowania komórki, dzięki czemu można przechowywać potencjalnie większe ilości danych.

[15] J. Sammons, *The Primer for Getting Started in Digital Forensics*, Boca Raton 2015, s. 167.

W celu obniżenia kosztów produkcji pamięć typu flash jest zaprojektowana w taki sposób, aby jej komórki wykonywały tylko trzy operacje:
— programowanie (ang. *program*), czyli umożliwiały zapis ładunku elektrycznego,
— odczyt (ang. *read*), czyli umożliwiały sprawdzenie aktualnego stanu ładunku,
— kasowanie (ang. *erase*), czyli umożliwiały rozładowywanie komórki, tj. pozbycie się ładunku elektrycznego.

Warto zaznaczyć, że nie wszystkie te operacje można wykonywać na poszczególnych tranzystorach. Programować i odczytywać można tylko tzw. strony (ang. *page*) jednostki pamięci składające się w przypadku pamięci typu SLC z około 16 000 pojedynczych tranzystorów. Aby przeczytać stan naładowania poszczególnego tranzystora (bit), należy przeczytać całą tzw. stronę. Jeszcze gorzej wygląda operacja rozładowywania. Fizyczne ograniczenia przyczyniają się do tego, że skasować można tylko tzw. blok, tj. jednostkę rozmiarową pamięci, która w pamięciach SLC składa się z ok. 2 000 000 pojedynczych tranzystorów (bitów). We współczesnych pamięciach typu TLC rozmiar bloku wynosi od 1,5 do 3 megabajtów. To oznacza, że w pamięci typu flash nie ma możliwości zmiany zawartości pojedynczego tranzystora ze stanu naładowania (kod binarny „0") do stanu rozładowania (kod binarny „1"). Jednoczesne rozładowanie takiej ilości komórek powoduje powstanie dużego przepływu prądu, swego rodzaju „błyskawicy". Wyzwolony z komórek prąd po drodze uszkadza pływające bramki oraz pozostałe elementy dysku, przez co operacja rozładowywania jest najbardziej szkodliwa pod względem żywotności pamięci. Właśnie w liczbie cykli programowania/kasowania mierzy się czas żywota pamięci flash. Operacja kasowania jest również najwolniejszą wśród wymienionych trzech, dlatego jest ona wykonywana z reguły w tle, gdy tzw. głowa SSD, tj. kontroler, ma wolną chwilę.

Zużyte komórki przestają się programować lub rozładowywać, ponieważ, gdy tylko zostaną zidentyfikowane przez kontroler, są wpisywane na listę uszkodzonych (ang. *bad*), których oprogramowanie dysku SSD nie będzie już wykorzystywać[16]. Przy tym należy dodać, że im dłużej dane pozostają w tym samym miejscu, tj. na jednym fizycznym obszarze pamięci, tym trudniej jest ten obszar pamięci rozładować. Dane w pewnym sensie wypalają się w komórkach pamięci (ang. *data retention*) podobnie do tego, jak kiedyś wypalały się statyczne rysunki na monitorach kineskopowych. Aby zapewnić zrównoważone zużycie (ang. *wear leveling*), konieczne jest przenoszenie danych z jednego obszaru fizycznego pamięci do drugiego.

To wszystko powoduje, że dysk SSD musi mieć swój własny, wewnętrzny system plików FTL (ang. *Flash Translation Layer*, tj. warstwa oprogramowania flash), który uwzględnia opisaną specyfikę działania pamięci

[16] J. Cooke, *The Inconvenient Truths of NAND Flash Memory*, s. 23, <http://cushychicken.github.io/assets/cooke_inconvenient_truths.pdf>, 5 listopada 2018 r.

flash oraz pozwala na przedstawienie urządzenia na poziomie systemu operacyjnego jako wspomnianego wyżej urządzenia blokującego. Działanie FTL, w odróżnieniu od systemu plików w systemie operacyjnym NTFS (ang. *New Technology File System*), FAT (ang. *File Allocation Table*) czy EXT (ang. *Extended File System*), odbywa się w sposób zupełnie niewidoczny zarówno dla użytkownika, jak i dla sytemu operacyjnego. Wie o nim tylko kontroler dysku SSD.

Należy brać pod uwagę również i ten fakt, że w odróżnieniu od dysku typu HDD w przypadku pamięci flash nie istnieje coś takiego jak modyfikacja danych. Jeśli dla dysku typu HDD nadpisywanie danych to tylko jeden przelot nad odpowiednimi sektorami głowicy odczytująco-zapisującej, to dla pamięci flash jest to cały szereg czynności.

Rozpatrzymy niektóre istotne aspekty działania FTL na poniższym przykładzie, w którym mamy do czynienia z dwoma blokami (blok „1" i blok „2") pamięci typu flash. Przyjmijmy, że w ramach kroku pierwszego do bloku „1" użytkownik zapisał dane A, B, C oraz D. Dalej użytkownik chce zmodyfikować już zapisane dane, poczynając od A, kończąc na D, oraz zapisać obok nowe dane E, F, G i H. Z uwagi na to, że raz zapisanych danych zmienić nie można, FTL zapisuje nowe dane A–D do obszaru wolnego bloku „1". W rezultacie powstają dane A'–D', a kolejno FTL oznacza wcześniejsze dane A–D jako nieaktualne (ang. *invalid*). Nieaktualność powoduje, że dane stają się niewidoczne dla systemu operacyjnego oraz trafiają do kolejki na kasowanie. Żeby pozbyć się nieaktualnych danych i zwolnić miejsce pod nowe dane, kontroler tworzy kopię danych ważnych (ang. *valid*) w obszarze wolnym (blok „2"), po czym usuwa wszystkie dane z bloku „1", rozładowując go. Ta ostatnia operacja określana jest jako tzw. zbieranie śmieci (ang. *garbage collection).* Oczywiście wszystkie opisane manipulacje z danymi są zupełnie niewidoczne z poziomu użytkownika. FTL pilnuje, aby użytkownik, tj. system operacyjny, nie widział niczego poza spójnym ciągiem adresów LBA.

Dla prawidłowego działania FTL potrzebne jest wolne miejsce. Producenci stosują różne metody umożliwiające identyfikację miejsc, które oprogramowanie dysku SSD może bezpiecznie zwolnić. Najprostsza sytuacja pojawia się, gdy oprogramowanie SSD (czyli FTL) samo stwarza dane nieaktualne (jak w podanym wyżej przykładzie). Wtedy dysk na pewno wie, że stara kopia danych nie jest potrzebna. Inaczej wygląda sytuacja, gdy użytkownik opróżnia kosz czy w inny sposób kasuje niepotrzebne mu pliki. W przypadku dysku typu HDD operacja powoduje zmianę kilku bitów (tzw. flagów) w służbowych, systemowych plikach, np. w systemie plikowym NTFS jest to tabela plików MFT (ang. *Master File Table*) oraz plik $BitMap. W rezultacie miejsce, które dalej jest zajęte przez pliki, będzie oznaczone jako wolne, a system operacyjny przestanie je widzieć[17]. W innym stosowanym w środowisku Windows systemie plików FAT w przypadku usunięcia pliku pierwsza litera jego nazwy jest zastępowana przez symbol specjalny[18].

[17] B. Carrier, *File system forensic analysis,* Crawfordsville 2005, s. 315.
[18] Tamże, s. 231.

Kiedy system operacyjny chce zapisać na dysku nowe dane, sprawdzane są flagi, po czym ten obszar LBA, który jest oznaczony jako wolny, zostaje nadpisany. W odróżnieniu od tego klasycznego schematu, kontroler dysku SSD, przed zapisaniem nowych danych, musi mieć do dyspozycji wolne miejsce, uprzednio rozładowane bloki.

Kontroler dysku SSD identyfikuje bloki, które można zwolnić na dwa sposoby:

— stosując algorytm umożliwiający ciągłe analizowanie kluczowych struktur konkretnego zastosowanego systemu plików — odrębny dla każdego systemu plików, tj. NTFS, FAT wersji 16, 32 i 64, EXT wersji 2, 3 i 4 czy HFS (ang. *Hierarchical Filesystem*, tj. system plików używany przez system operacyjny Apple Macintosh). W tym przypadku kontroler SSD musi mieć dodatkową moc obliczeniową oraz zwiększoną objętość pamięci RAM,

— uzyskując informacje od systemu operacyjnego.

Z uwagi na koszty na rynku dominują urządzenia wykorzystujące drugie podejście, które nazwa się TRIM (ang. *trim*, tj. przyciąć), zatem nadbudowę protokołu, za pomocą którego system operacyjny komunikuje się z nośnikiem danych, czyli protokołem ATA (ang. *Advanced Technology Attachments*), którego współczesną implementacją jest SATA (ang. *Serial Advanced Technology Attachment*). Za pomocą TRIM system może przekazać oprogramowaniu SSD adresy niepotrzebnych więcej użytkownikowi danych, które wpisuje na listę adresów przeznaczonych do kasowania i rozładowania w wolnej chwili (ang. *idle time*) w ramach wspomnianej wyżej operacji *garbage collection*. Niestety w praktyce prawidłowe działanie TRIM czasami się zakłóca, bo potrzebuje współdziałania co najmniej trzech elementów komputera: systemu operacyjnego (TRIM jest nadbudową systemu operacyjnego i powinien być w nim prawidłowo zaimplementowany), interfejsu znajdującego się pomiędzy OS (ang. *operating system*, tj. system operacyjny) a dyskiem, który musi akceptować polecenia TRIM, oraz samego dysku SSD[19]. Błędy w przypadku dowolnego z wymienionych komponentów mogą powodować, że TRIM nie będzie funkcjonował. Niedziałający lub nieprawidłowo działający TRIM z jednej strony będzie ułatwiał odzyskiwanie usuniętych danych, które pozostaną w komórkach pamięci, z drugiej strony po jakimś czasie kontroler wykorzysta wolne miejsce i będzie rozładowywał bloki bezpośrednio przed każdą następną operacją zapisywania danych, co zasadniczo obniży prędkość działania nośnika. Wadą TRIM jest to, że od momentu otrzymania przez SSD polecenia kontroler dysku zaczyna rozładowywać odpowiednie obszary pamięci (wykonywać trwałe usuwanie danych). Kontroler będzie to robił zupełnie niezależnie, gdy tylko zostanie podłączony do zasilania.

W 2011 r. grupa badaczy z australijskiego Murdoch University opisała wpływ mechanizmów działania dysków SSD na możliwości odzyskania dowodów cyfrowych. W ramach eksperymentu wybrano popularny model

[19] Y. Gubanov, O. Afonin, *SSD and eMMC Forensics 2016*, <https://articles.forensicfocus.com/2016/04/20/ssd-and-emmc-forensics-2016>, 5 listopada 2018 r.

dysku SSD o pojemności 64 gigabajtów, który porównywano z dyskiem typu HDD o pojemności 80 gigabajtów. W rezultacie ujawniono istnienie zjawiska tzw. korozji (ang. *self-corrosion*), tj. samodzielnego, bez poleceń użytkownika czy też systemu operacyjnego komputera, trwałego niszczenia przez kontroler dysku zapisanych na nim danych cyfrowych. Badany SSD przeprowadzał tę operację również w sytuacji, gdy badacze stosowali standardowe metody zapobiegania uszkodzeniu danych, które używane są przez biegłych policyjnych (utworzenie obrazu dysku, blokada zapisu). Co ciekawe, dysk SSD zaczynał usuwać dane już po upływie 3 minut od momentu włączenia zasilania, a po kolejnych 3 minutach całkowicie usuwał dane, które uprzednio były zaznaczone jako usunięte. W konsekwencji nie udało się odzyskać ani jednego całego pliku. Z 25 000 plików tekstowych, którymi badacze wypełnili dysk SSD, tylko jeden udało się odzyskać na 50% jego oryginalnej zawartości — był to najlepszy rezultat. Większość plików zostało uszkodzonych nawet w 82%. Powyższe działania przetrwało zaledwie 0,03% zapisanych na dysku danych. Należy zaznaczyć, że autorzy nie zauważyli żadnej sygnalizacji, iż trwa wykonywanie algorytmu *garbage collection*. Dysk nie podawał jakichkolwiek sygnałów, nie powodował systemowych ostrzeżeń wizualnych czy dźwiękowych, nie było też wibracji, hałasu lub kliknięć, których w podobnej sytuacji można by oczekiwać od klasycznego dysku typu HDD. Pod tym względem włączony dysk SSD niczym nie różnił się od dysku SSD, który był odłączony od zasilania. Wniosek był taki, że korozja danych cyfrowych w przypadku dysków SSD odbywa się w sposób, który nie ma rejestrowanych objawów fizycznych[20].

Ciągłe działanie algorytmu *garbage collection*, którego nie da się zatrzymać, powoduje, że dane ulegają modyfikacji zanim zostanie utworzona ich dokładna kopia. W efekcie standardowa weryfikacja integralności danych w przypadku dysków SSD jest niemożliwa. Jedynym sposobem utworzenia pełnej kopii takiego dysku jest operacja polegająca na wylutowywaniu kostek pamięci oraz późniejszym odczytywaniu ich pojedynczo przy użyciu specjalnego sprzętu (ang. *chip-off*)[21].

Wniosek końcowy brzmi dość paradoksalnie. Żeby zachować jak największą ilość danych zapisanych na dysku SSD, po ustaleniu, jaki nośnik jest wykorzystywany w badanym sprzęcie, należy niezwłocznie odłączyć źródło zasilania. Problem jest jednak taki, że nawet jeśli przeszukujący ma do czynienia z włączonym komputerem z rozkręconą obudową, to bez wiedzy specjalnej nie potrafi on określić, czy komputer jest wyposażony w zwykły dysk HDD, czy w dysk SSD, ponieważ ich wygląd zewnętrzny nie rozstrzyga tej kwestii. Z drugiej strony wyłączenie działającego komputera (ang. *live-system*) uniemożliwi analizę zawartości jego komponentów znajdujących się na samym początku listy zmian, takich jak pamięć RAM. Powyższe może skutkować nie tylko utratą cennych informacji dowodowych,

[20] G.B. Bell, R. Boddington, *Solid State Drives: The Beginning Of The End For Current Practice In Digital Forensic Recovery?*, s. 17, <http://www.jdfsl.org/subscriptions/JDFSL-V5N3-Bell.pdf>, 5 listopada 2018 r.

[21] Tamże, s. 6.

ale też całkowitym wykluczeniem możliwości dalszego badania danych, np. jeśli użytkownik korzysta z oprogramowania szyfrującego odrębne pliki czy może nawet cały dysk, hasło do ich odszyfrowania będzie się znajdowało w pamięci operacyjnej RAM. Bez należytego zabezpieczenia dane zostaną utracone w momencie odłączenia komputera od zasilania. Jedyną metodą utrwalenia znajdujących się na dysku SSD danych jest obciążenie nośnika przez zadania w taki sposób, żeby kontroler dysku nie rozpoczął operacji *garbage collection*. W tej sytuacji badacze mogą bezpiecznie wykonać kopię (obraz) przedmiotowego dysku, nie ryzykując przy tym utratą istotnych informacji.

Słowa kluczowe: kryminalistyka, ekspertyza, ślady cyfrowe, nośniki danych cyfrowych, odzyskiwanie danych cyfrowych

Keywords: digital forensics, expertise, electronic evidence, hard disk, digital data recovery

Streszczenie: W artykule zostały opisane mechanizmy działania dysków typu HDD oraz SSD. Autor szczególną uwagę zwraca na wynikające z tego komplikacje dla biegłych, których zadanie polega na odzyskiwaniu danych cyfrowych mogących mieć znaczenie dla toczącego się postępowania.

Summary: The article describes the mechanism of action of HDD and SSD disks. The author pays special attention to the resulting complications for the experts, whose task is to recover digital data that may be relevant to the ongoing proceedings.