

**Zadaniem technologii RAID jest zwiększenie niezawodności i wydajności systemów dyskowych - w zastosowaniach profesjonalnych, ale i w komputerach ambitnych amatorów. Do wyboru jest osiem wariantów o specyficznych zaletach i wadach.**

Zadaniem technologii RAID jest zwiększenie niezawodności i wydajności systemów dyskowych - w zastosowaniach profesjonalnych, ale i w komputerach ambitnych amatorów. Do wyboru jest osiem wariantów o specyficznych zaletach i wadach.

Gdy przed 20 laty na Uniwersytecie Kalifornijskim w Berkeley obwieszczano światu powstanie RAID, przechowywanie dużych ilości danych wciąż było poważnym problemem. Stosowane powszechnie w centrach obliczeniowych dyski SLED (Single Large Expensive Disk) w formacie 14 cali miały wprawdzie dużą pojemność (2 do 3 GB) i zapewniały duże bezpieczeństwo danych, były jednak skrajnie drogie. Alternatywą były dużo tańsze (w przeliczeniu na pojemność), stosunkowo nowe, "małe" dyski w formacie 5,25 cala. Ich zastosowanie wiązało się jednak z pewnymi problemami.

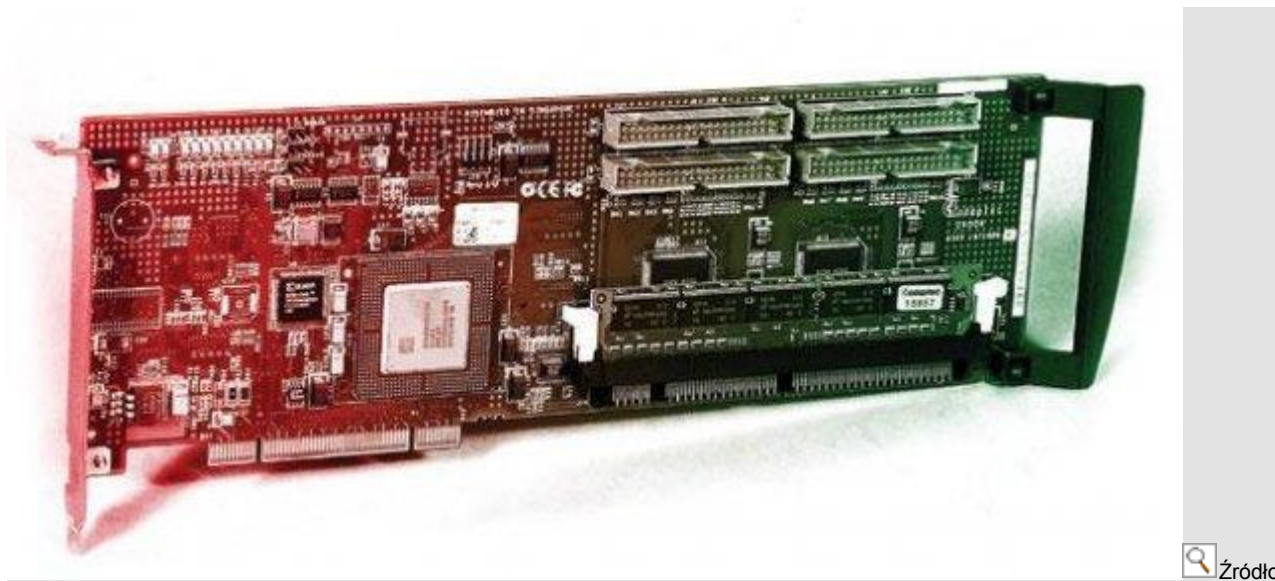
#### **Zobacz również:**

Składowanie danych na wielu małych zamiast na jednym dużym dysku utrudnia zarządzanie danymi. W przypadku połączenia dysków w pakiet niezależnych napędów (JBOD - Just a Bunch of Disks) znalezienie wolnego miejsca lub wcześniej zapisanych danych staje się skomplikowane.

Z drugiej strony, niezawodność małych dysków była znacznie mniejsza, niż napędów SLED. Dodatkowym utrudnieniem było to, że statystyczne prawdopodobieństwo utraty danych przechowywanych na wielu dyskach jest dużo większe, niż w przypadku pojedynczego dysku.

Rozwiązaniem problemu stała się propozycja, jaką zgłosiło trzech doktorantów Uniwersytetu w Berkeley - David Patterson, Garth Gibson i Randy Katz. Opracowali oni ideę połączenia wielu małych napędów w niezawodny zespół, wyposażony w mechanizmy wykrywania i korekcji błędów. W swojej pracy, opublikowanej w czerwcu 1988, nadali nowej technice niepowtarzalną nazwę. Tytuł pracy brzmiał A Case for Redundant Arrays of Inexpensive Disks (RAID). Tak narodził się RAID.

## Szybkie i niezawodne



Źródło

wydajności - typowy kontroler RAID, na zdjęciu - firmy Adaptec, z wieloma kanałami i własnym procesorem.

Patterson, Gibson i Katz zaproponowali w swojej pracy łącznie pięć różnych sposobów połączenia poszczególnych dysków w macierz. Nazwali je RAID Level 1 do 5 i ta terminologia obowiązuje do dzisiaj. Nazwa ta jest zresztą źródłem częstych nieporozumień. Słowo level sugeruje, że chodzi o kolejne poziomy pewnej procedury, tymczasem w rzeczywistości są to zupełnie różne, niezależne techniki. Techniki te nie tylko realizują założone cele - zapewniają tanie i niezawodne metody składowania dużej ilości danych - ale dają również dodatkowe korzyści. Po pierwsze, macierz RAID zachowuje się, z punktu widzenia użytkownika, jak pojedynczy napęd logiczny. Zarządzanie danymi jest tak proste, jak w przypadku pojedynczego dysku. Po drugie, niektóre z technik RAID zapewniają szybszy dostęp do danych w wyniku ich podziału na poszczególne napędy, z równoległym dostępem. Tę możliwość można jednak wykorzystać tylko wtedy, gdy jest wystarczająca liczba kanałów dostępowych do dysków. Połączenie tak wielu zalet sprawiło, że technologia RAID szybko znalazła szerokie zastosowanie, szczególnie w segmencie serwerowym. Obecnie jest osiem podstawowych poziomów RAID, od Level 0 do Level 7, oraz poziomy kombinowane, na przykład RAID 0+1, lub RAID 50.

## RAID programowy a sprzętowy

	RAID PROGRAMOWY	RAID SPRZĘTOWY
Koszty wdrożenia	niskie	wysokie
Wydajność	niska	wysoka
Obciążenie procesora komputera	wysokie	niskie
Zależność od platformy	tak	nie
Zależność od systemu operacyjnego	tak	tak

programowy a sprzętowy

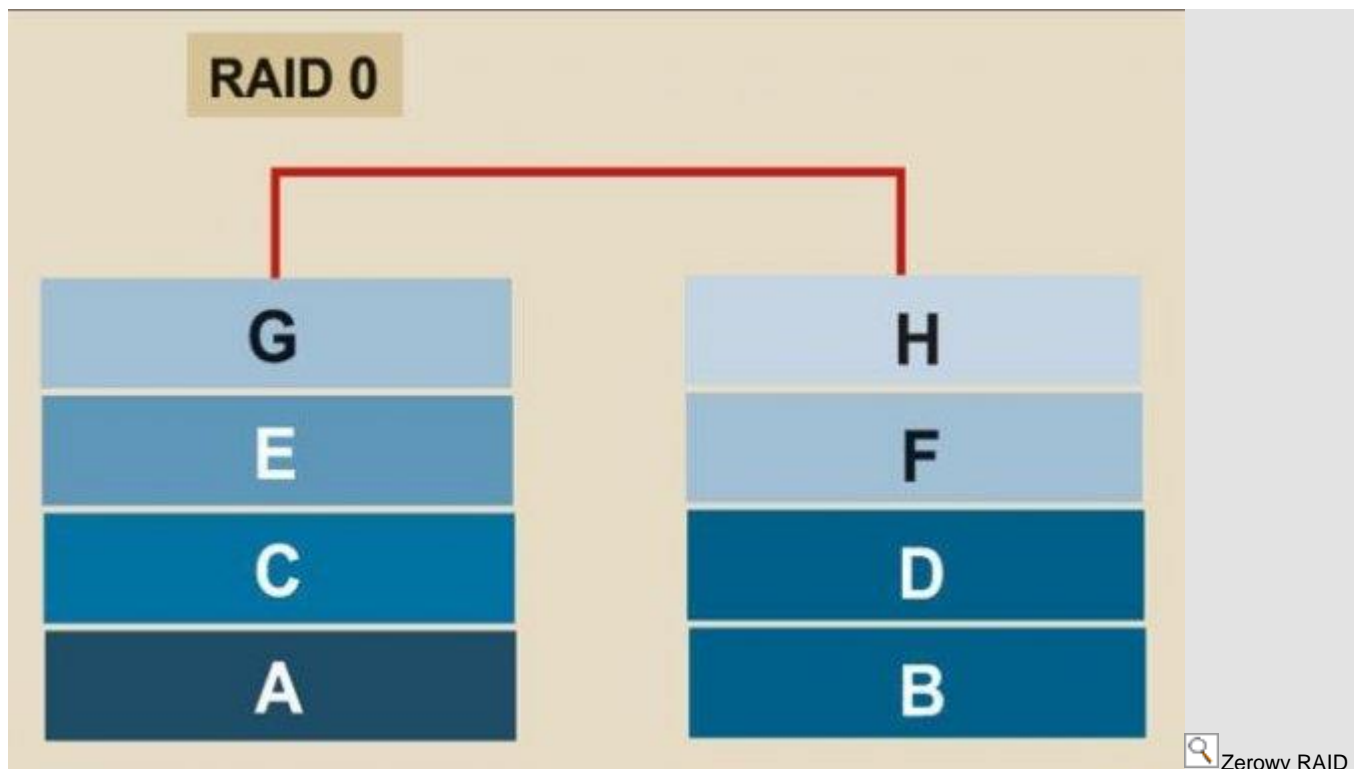
Nieco mylące określenia "programowy" i "sprzętowy" RAID - w końcu oba warianty potrzebują oprogramowania - odnosi się do sposobu realizacji.

W przypadku programowego RAID za sterowanie zespołem dysków odpowiada oprogramowanie zainstalowane na komputerze. Niektóre z systemów operacyjnych mają już niezbędne składniki. Na przykład Windows NT obsługuje RAID 0 oraz RAID 1 i 5 - ten ostatni tylko w wersji serwerowej. Linux obsługuje macierze poziomą 0, 1, 4 i 5.

RAID programowy jest więc w wielu przypadkach najtańszym i najprostszym rozwiązaniem. Natomiast oprogramowanie RAID bardzo obciąża procesor komputera, ponadto jest związane z konkretną platformą i systemem operacyjnym. Oprócz tego zwykle mamy tylko jedno lub dwa złącza do podłączenia napędów, co ogranicza możliwości równoległych odwołań do dysków, a zatem i wydajność.

W RAID sprzętowym macierzą steruje specjalny kontroler. Daje to w efekcie odciążenie procesora i wzrost wydajności. Ponadto kontroler RAID łączy się z dyskami przez wiele kanałów, co umożliwi równoległe odwołania i zapewnia szybki transfer danych. Sprzętowe macierze RAID pracują niezależnie od platformy, choć oczywiście wymagają oprogramowania dostosowanego do systemu operacyjnego.

## Mieszane zespoły dyskowe

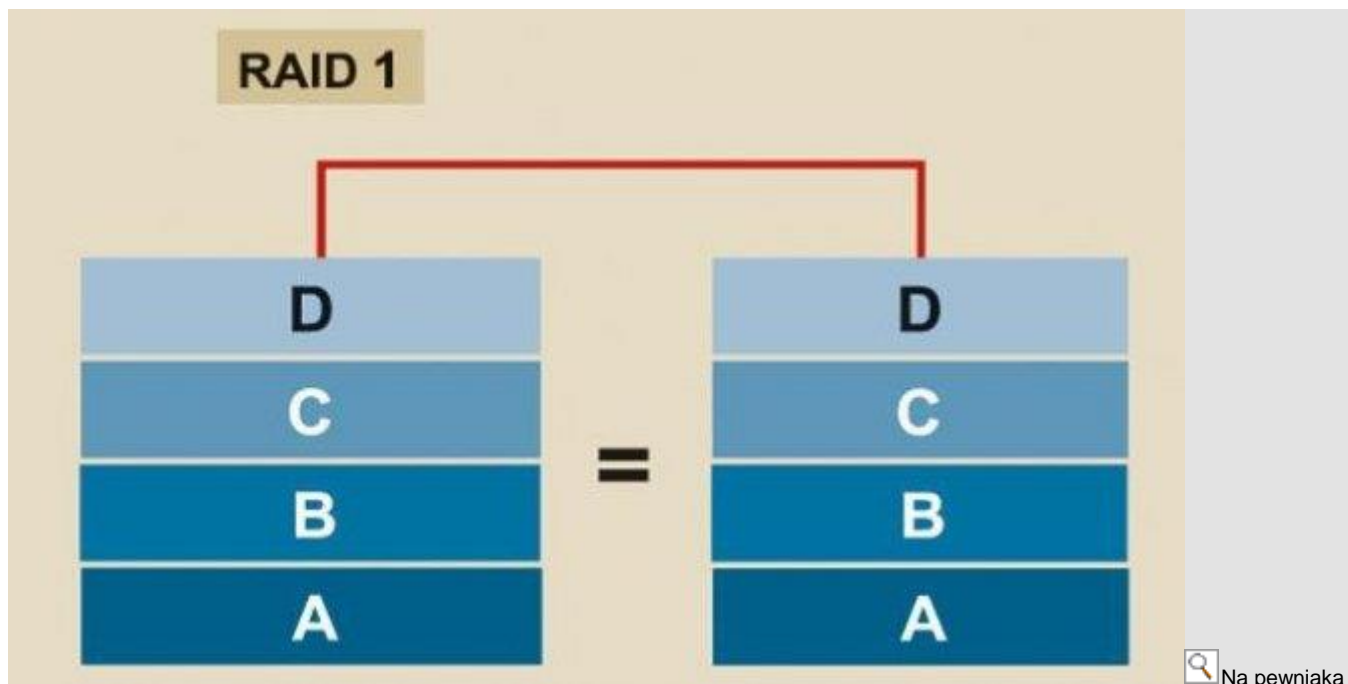


RAID Level 0 przyspiesza wprawdzie odwołania do dysków, jednak nie zapisuje danych w sposób nadmiarowy.

Jeszcze przed kilku laty kontrolery RAID stawiały bardzo wysokie wymagania co do stosowanych napędów. Obowiązkowy był interfejs SCSI, wszystkie napędy musiały mieć taką samą pojemność, a nawet pochodzić z tej samej serii produkcyjnej. Od pewnego czasu reżim nie jest już tak surowy. W przypadku serwerów nadal stosuje się SCSI lub Fibre Channel, przede wszystkim ze względu na wydajność. Do komputerów biurowych dostępne są też kontrolery z interfejsem Ultra ATA/66 i Ultra ATA/100.

Nowoczesne kontrolery i programowe rozwiązania RAID umożliwiają również stosowanie w zespole napędów różnej pojemności. Niemniej w takim przypadku nie można wykorzystać całej dostępnej pojemności netto dysków. Ponieważ technologia RAID zakłada stosowanie napędów o jednakowej pojemności, w konfiguracji mieszanej można wykorzystać na każdym dysku pojemność równą jedynie pojemności najmniejszego dysku. Na przykład w razie połączenia napędu 20 GB z dwoma napędami 30 GB, do dyspozycji będzie pojemność  $3 \times 20$  GB.

## RAID Level 0



RAID 1 tworzy na drugim dysku kopię danych.

RAID Level 0 to - jak wskazuje zero w nazwie - nie jest rozwiązaniem nadmiarowym. Służy jedynie do przyspieszenia dostępu do napędów. W tym celu RAID 0 łączy dwa lub więcej napędów w jeden napęd logiczny. Zasada działania polega na rozdzieleniu danych na kolejno po sobie następujące bloki (paski - stripes), równomiernie między wszystkie napędy. Dlatego też RAID 0 określa się często mianem systemu paskowego.

Równoległy odczyt względnie zapis na wielu dyskach przyspiesza wprowadzenie transfer, jednak wpływa ujemnie na poziom bezpieczeństwa danych. Jeżeli jeden z napędów ulegnie uszkodzeniu, wszystkie dane są stracone. Wzrost szybkości jest zauważalny głównie w przypadku dużych plików o podobnym charakterze. Mechanizm paskowania może wówczas działać równolegle na wszystkich napędach, zwielokrotniając transfer danych.

Podczas zapisu lub odczytu wielu małych plików czynnikiem ograniczającym jest czas dostępu poszczególnych dysków. W takim przypadku zestaw paskowy uzyskuje w najlepszym przypadku wydajność pojedynczego napędu. Ze względu na te właściwości systemu RAID 0 stosuje się przede wszystkim tam, gdzie przetwarzane są duże ilości danych, a więc w stacjach roboczych do obróbki plików audio/video lub CAD/CAM.

## RAID Level 1

Obliczanie informacji o parzystości	
	Zawartość dysku
Dysk A	11101100
Dysk B	10110011
Dysk C	01001101
Dysk parzystości	00010010

 Obliczanie informacji o parzystości

RAID Level 1 jest systemem lustrzanym (mirroring). To określenie wyjaśnia właściwie wszystko - wszystkie zapisy odbywają się równolegle na dwóch dyskach i w rezultacie każdy z dysków jest lustrzaną kopią drugiego.

Wszystkie dane są w dwóch kopiach, a więc poziom bezpieczeństwa jest wysoki. Nawet w przypadku całkowitej awarii jednego dysku dostępna jest kompletna kopia. Wadą rozwiązania jest to, że użyteczna, dostępna pojemność stanowi zaledwie 50 procent pojemności zespołu. Tak więc koszt składowania danych podwaja się.

Po podłączeniu dysków do własnych kanałów wydajność odczytu danych rośnie dwukrotnie. Z kolei zapis przebiega najwyżej tak szybko, jak na pojedynczym dysku. Dlatego też systemy lustrzane najlepiej sprawdzają się tam, gdzie przechowywane są ważne dane, przeznaczone głównie do odczytu.

## RAID Level 0+1

Połączenie systemu lustrzanego i paskowego przyspiesza transfer i podnosi poziom bezpieczeństwa danych. RAID 0 zapewnia szybszy transfer poprzez równoległy zapis i odczyt na wielu dyskach. Dodatkowe kopie lustrzane zestawów paskowych podnoszą poziom bezpieczeństwa danych. Różni producenci stosują różne oznaczenia - taki system może się nazywać RAID 0+1, RAID 0/1 lub RAID 10.

Dyskami można sterować na dwa sposoby. Załóżmy, że mamy do dyspozycji sześć dysków. Można więc najpierw połączyć trzy z nich w zestaw paskowy (RAID 0). Następnie oba tak powstałe dyski logiczne służą do utworzenia zestawu lustrzanego (RAID 1). Odwrotna konfiguracja daje pozornie taki sam wynik - najpierw tworzymy trzy zestawy lustrzane po dwa dyski, a następnie te trzy dyski logiczne łączymy w

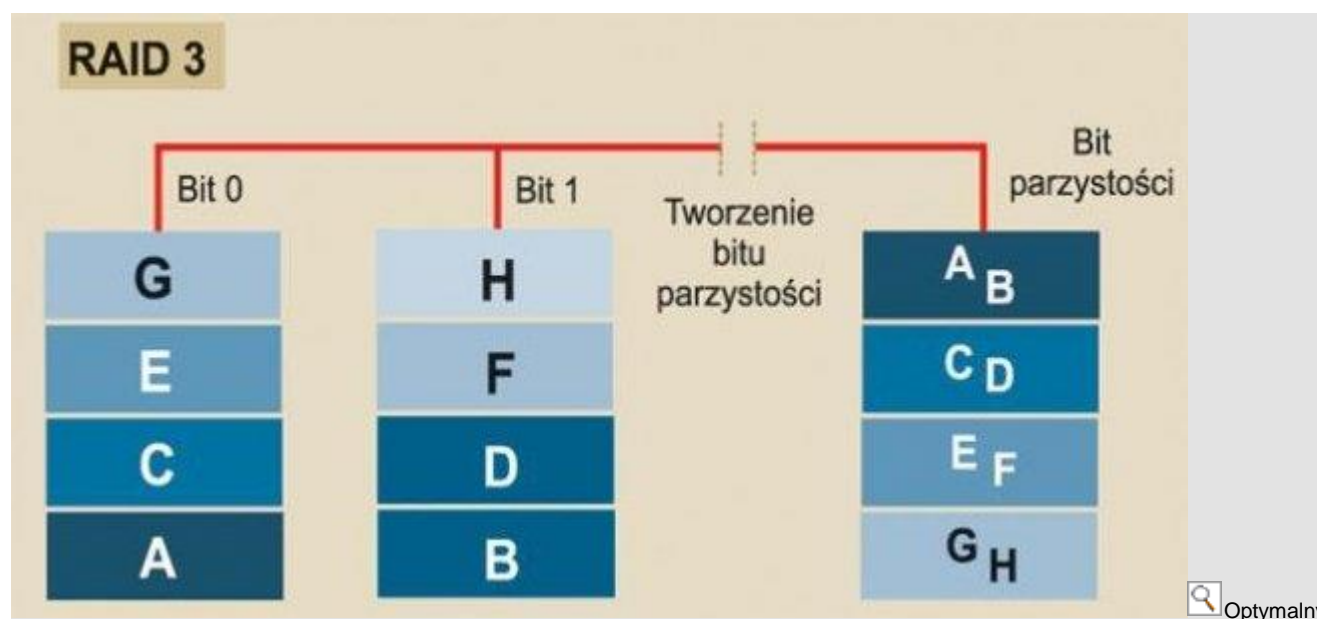
zestaw paskowy. W obu przypadkach ostatecznie mamy do dyspozycji pojemność trzech dysków.

### Zobacz również:

Jednak z punktu widzenia bezpieczeństwa danych pierwszeństwo należy przyznać drugiej metodzie. Łatwo zauważyć, że gdy najpierw tworzymy zestawy paskowe, to w razie awarii jednego z dysków cały zestaw jest bezużyteczny. Dane dostępne są wprawdzie w drugim zestawie paskowym, ale awaria któregoś z jego dysków spowoduje bezpowrotną utratę wszystkich danych.

W przeciwnym przypadku po awarii jednego z dysków zespół lustrzany wprawdzie utraci nadmiarowość, ale dopiero awaria drugiego dysku w tym samym zespole spowoduje, że dane przepadną bez-powrotnie. Prawdopodobieństwo całkowitej katastrofy jest w wypadku drugiej metody o jedną trzecią mniejsze w porównaniu z pierwszą metodą.

## Metody z korekcją błędów



tylko w przypadku dużych plików - RAID 3 wykorzystuje paskowanie bajtowe i dedykowany dysk parzystości.

System lustrzany gwarantuje wprawdzie wysoki poziom nadmiarowości, jednak jest rozrzutny w gospodarowaniu pojemnością dysku, a więc bardzo zwiększa koszty składowania danych. Rozwiązaniem tego problemu są macierze RAID Level 2 do 7 z korekcją błędów. Za pomocą metody paskowania dane użyteczne rozdzielane są najpierw na co najmniej dwa dyski z danymi użytecznymi. Na podstawie ich zawartości tworzona jest wartość kontrolna, która służy do rekonstrukcji danych w razie awarii dysku. Kod parzystości ECC zapisywany jest na oddzielnym dysku.

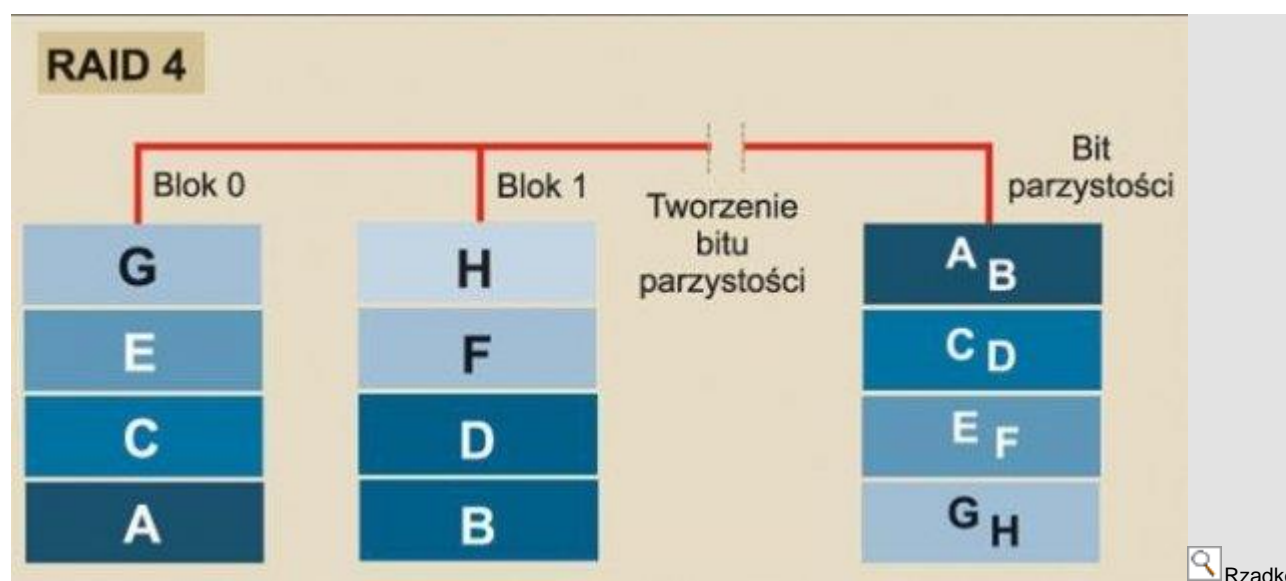
Kontrola parzystości to jedna z najstarszych metod korekcji błędów - dane użyteczne są wiązane ze sobą za pomocą operacji logicznej XOR, a wynik jest zapisywany na dysku parzystości. Rezultatem powiązania danych jest 1 w przypadku nieparzystej liczby bitów w ciągu danych. Liczba parzysta daje w wyniku 0.

Jeżeli dowolny dysk z danymi ulegnie awarii, ponowna operacja XOR umożliwia rekonstrukcję danych.

## Korekcja błędów to hamulcowy zapisu

Warianty RAID oparte na kontroli parzystości znacznie zmniejszają nadmiar danych niezbędnych do uzyskania odpowiedniego poziomu bezpieczeństwa. W naszym przykładzie dane nadmiarowe stanowią jedną trzecią danych użytecznych. Jeżeli dysków z danymi użytecznymi jest więcej, współczynnik ten jeszcze się zmniejszy. Jednakże nieustanne uaktualnianie informacji o parzystości wymaga dodatkowych operacji zapisu i odczytu.

## RAID Level2



wykorzystywany - RAID 4 przypomina RAID 3, jednak lepiej sobie radzi z odczytem małych plików.

RAID Level 2 zapewnia dodatkową ochronę przed błędami w obrębie dysków, jest jednak bardzo rzadko stosowany ze względu na skomplikowaną implementację. Z niewielkimi wyjątkami technologię tę stosowano jedynie w wielkich centrach obliczeniowych.

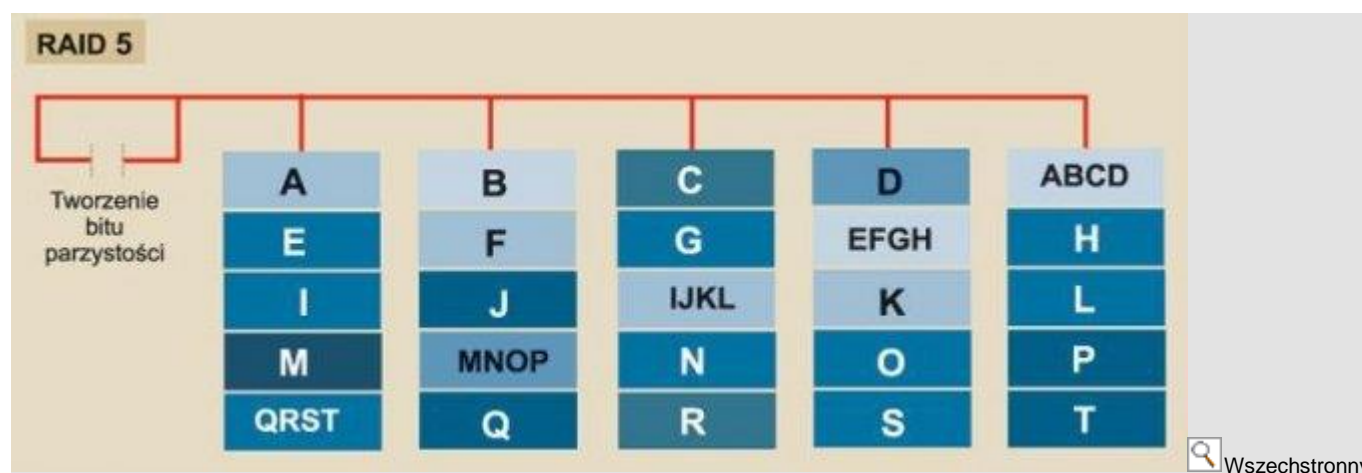
RAID 2 opiera się na bitowym podziale danych użytecznych. Stosowana w tym wypadku korekcja błędów zapewnia nie tylko zabezpieczenie przed awarią jednego



dysku. Wszystkie inne poziomy RAID zawiodą, gdy dane staną się na przykład niespójne w wyniku błędu zapisu. System rozpoznaje wprawdzie błąd, ale nie stwierdzi, który dysk dostarcza błędnych danych, a więc nie może ich skorygować.

Dlatego RAID 2 zapisuje na każde 8 bitów danych dodatkowo 2 bity kodu ECC. Dzięki temu można nie tylko odkryć błąd, lecz również ustalić miejsce, w którym wystąpił. RAID 2 przypomina pod tym względem pamięć ECC-RAM, która nie zatrzymuje komputera w razie wystąpienia błędu jednobitowego, lecz koryguje bit. Bitowy podział danych na napędy wymusza zastosowanie co najmniej dziesięciu napędów w zestawie. Możliwość równoległych odwołań podczas odczytu zwiększa szybkość odczytu ośmiokrotnie w stosunku do pojedynczego dysku. Podczas zapisu konieczność uwzględnienia dużego narzutu danych administracyjnych powoduje, że wydajność spada poniżej wydajności pojedynczego dysku.

## RAID Level 3



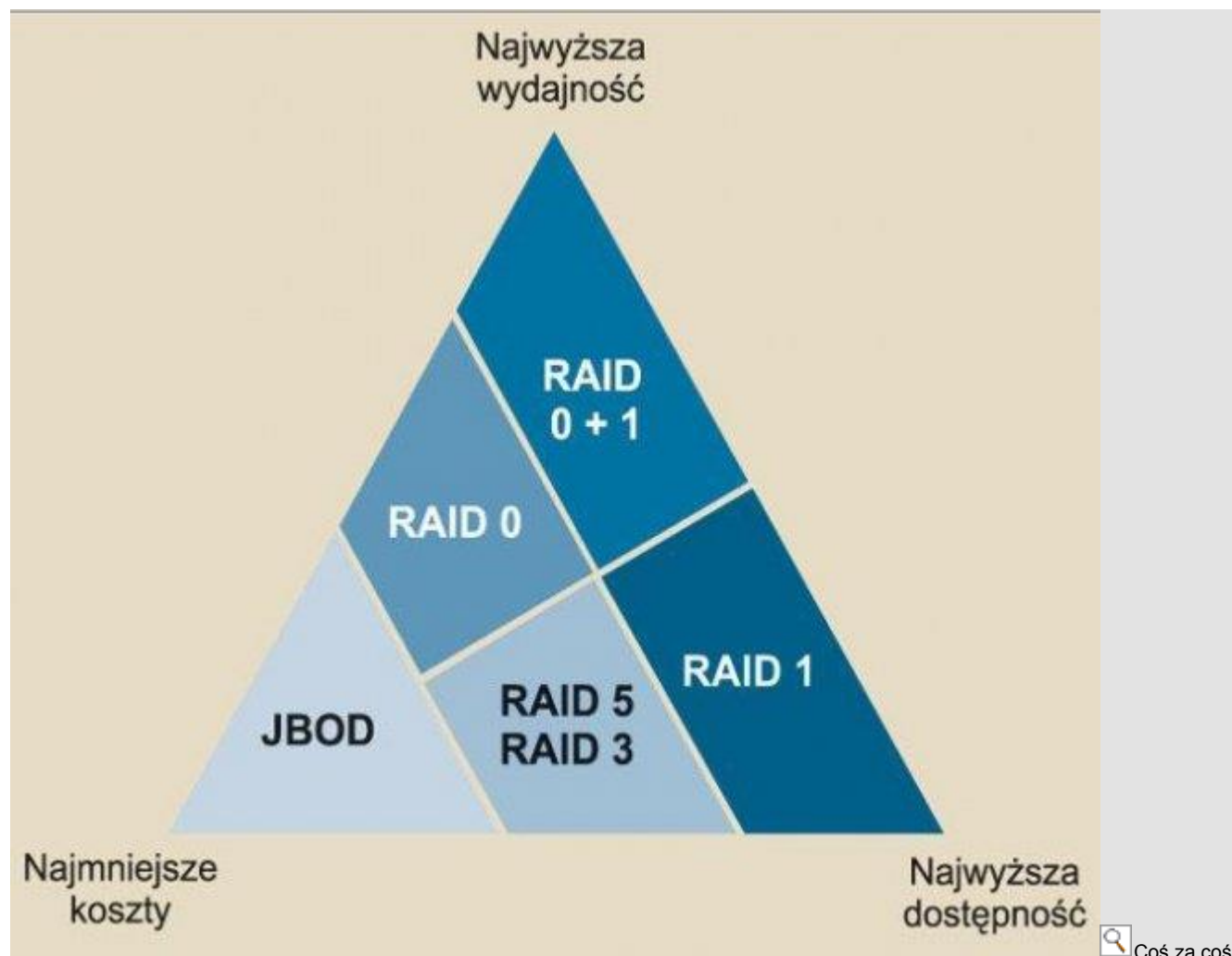
- RAID 5 jest szybszy w operacjach zapisu dzięki rezygnacji z oddzielnego dysku parzystości. RAID Level 3 polega na bajtowym paskowaniu danych. W przeciwieństwie do RAID 2, posługuje się zintegrowanymi funkcjami dysków do wykrywania błędów zapisu i odczytu. Dzięki temu wystarcza pojedynczy, dedykowany dysk parzystości.

W celu ułatwienia generowania danych ECC RAID 3 synchronizuje pozycje głowic w dyskach. Przyspiesza to zapis, gdyż dane parzystości i dane użyteczne zapisywane są na dyskach równolegle.

Liczne operacje odczytu małych, rozproszonych plików wymagają z kolei synchronicznego pozycjonowania głowic na wszystkich dyskach, a to musi trwać. RAID 3 zapewnia odczuwalne przyśpieszenie odczytu tylko w przypadku dużych plików, więc tę metodę stosuje się najczęściej podczas przetwarzania dużych,

jednolitych plików na pojedynczych komputerach. Typowe zastosowania to CAD/CAM i edycja multimedialnych.

## RAID Level 3



wybór odpowiedniego do potrzeb wariantu RAID to kompromis między kosztem, wydajnością i poziomem bezpieczeństwa.

RAID 4 polega na blokowym paskowaniu danych użytecznych. Aby uniknąć wad RAID 3 związanych z przetwarzaniem małych plików, zrezygnowano z synchronizacji głowic wszystkich dysków. Dane parzystości przechowywane są, podobnie jak w RAID 2 i 3, na dedykowanym dysku.

Połączenie paskowania blokowego i niezależnego dostępu do poszczególnych dysków pozwala na wydajny odczyt nawet małych plików. Niestety, nie ma nic za darmo - z powodu braku synchronizacji operacje zapisu są w RAID 4 dużo wolniejsze niż w RAID 3. Każda aktualizacja danych ECC wymaga uprzedniego odszukania na dysku parzystości odpowiedniej porcji danych. W ten sposób dysk ECC staje się wąskim gardłem systemu.

Ze względu na swoje właściwości RAID 4 nadaje się przede wszystkim do pracy w środowiskach, w których operacje odczytu są częstsze niż zapisu. Ogólnie rzecz biorąc, wersja ta jest rzadko stosowana, ponieważ daje niewielkie korzyści w porównaniu z RAID 3.

## RAID Level 5

Porównanie poziomów RAID							
	RAID 0	RAID 1	RAID 10	RAID 2	RAID 3	RAID 4	RAID 5
Liczba dysków	$n > 1$	$n > 2$	$n > 3$	$n > 10$	$n > 2$	$n > 2$	$n > 2$
Dyski nadmiarowe	0	1	1**	2	1	1	1
Narzut administracyjny pojemności (%)	0	50	50	20	100 /	100 / n	100 / n
Równoległe operacje odczytu	n	2	$n/2$	8	$n-1$	$n-1$	$n-1$
Równoległe operacje zapisu	n	1	1	1	1	1	$n/2$
Maksymalna przepustowość odczytu*	n	2	$n/2$	8	$n-1$	$n-1$	$n-1$
Maksymalna przepustowość zapisu**	n	1	1	1	1	1	$n/2$

\* Jako współczynnik w odniesieniu do pojedynczego dysku.

Porównanie

### poziomów RAID

Podobnie jak RAID 4, RAID 5 też dzieli dane użyteczne na bloki. Nie ma tu jednak oddzielnego dysku parzystości, a dane ECC zapisywane są wraz danymi użytecznymi równomiernie na poszczególnych dyskach. Dzięki temu maleje prawdopodobieństwo, że na jednym dysku wystąpią jednocześnie dwie operacje zapisu. Zapis na dyskach RAID 5 może w dużej części odbywać się równoległe. W dodatku wszystkie dyski są jednakowo obciążone - żaden nie pełni szczególnej funkcji dysku parzystości.

Rozdzielenie danych na wszystkie dyski pozwala też uzyskać dobrą wydajność odczytu. Ma to szczególne znaczenie przy odczycie wielu małych bloków danych. Z tego powodu RAID 5 stosuje się najczęściej w systemach bazodanowych i serwerach transakcyjnych.

Podobnie jak RAID 1, RAID 5 bardzo dobrze się łączy z RAID 0. Tak powstały system RAID 0+5 (alias RAID 50) zapewnia podobnie dobrą wydajność, jak RAID 10, przy niezawodności wyższej niż RAID 5.

## Egzotyczne - RAID 6 i RAID 7

W systemach RAID 3 do 5 dopuszczalna jest awaria tylko jednego dysku, gdyż w przeciwnym razie nie da się zrekonstruować danych za pomocą operacji XOR. RAID 6 obchodzi to ograniczenie, uzupełniając RAID 5 o dodatkowy dysk parzystości. W ten sposób dane można odzyskać nawet po awarii dwóch dysków, jednak dodatkowe bezpieczeństwo okupione jest spowolnieniem zapisu w porównaniu z RAID 3 do 5.

Również rzadko stosowana macierz RAID 7 ma budowę podobną do RAID 5. Jej producent, firma Storage Computer, stosuje w kontrolerze dodatkowy, lokalny system operacyjny, działający w czasie rzeczywistym. Szybkie magistrale danych i duże pamięci buforowe odciążają właściwą magistralę napędów. Technika ta podobno znacząco przyspiesza zapis i odczyt w porównaniu z innymi wariantami RAID. Ponadto, podobnie jak w RAID 6, można rozłożyć dane parzystości na wiele dysków.

**Zobacz również:**

## **RAID - przegląd wariantów**

Wybór właściwego wariantu macierzy RAID wymaga dokładnego rozważenia zależności między dostępnością, wydajnością i kosztem MB. Organizacja dysków w postaci prostego zestawu (JBOD) oznacza najmniejsze koszty, jednak pozostawia wiele do życzenia pod względem szybkości i niezawodności. System lustrzany, RAID 1, gwarantuje najwyższą dostępność. Ten sam system jednocześnie najmniej efektywnie wykorzystuje dostępną pojemność dysków, a więc generuje najwyższe koszty. Gdy zestawić te czynniki w postaci diagramu, powstanie typowy trójkąt RAID (ilustracja "Coś za coś").

Zależności między poziomami RAID, wydajnością i niezawodnością przedstawiamy też w tabeli "Porównanie poziomów RAID". Wynika z niej wyraźnie, że każdy wariant ma specyficzne wady i zalety.

Wymiana w locie

Niezależnie od tego, czy chodzi o wewnętrzną macierz dyskową, zewnętrzny podsystem pamięci masowych, macierz sprzętową czy programową, obowiązuje zasada - jeżeli jeden z napędów ulegnie uszkodzeniu, omawiane systemy RAID tracą nadmiarowość. Każda kolejna awaria prowadzi nieuchronnie do utraty danych. Stąd logiczny wniosek, że uszkodzony dysk trzeba bezzwłocznie wymienić, żeby zrekonstruować macierz. Najlepiej, gdy w macierzy znajduje się dodatkowy dysk, gotowy do podjęcia pracy w razie awarii. Taki dysk, nazywany hot fix, hot spare lub stand by, jest automatycznie uruchamiany przez kontroler i włączany do macierzy.

Jeżeli takiego dysku nie ma, uszkodzony trzeba wymienić ręcznie. Z reguły oznacza to jednak wyłączenie systemu, co może się wiązać wręcz z przerwą w działalności przedsiębiorstwa. Szczególnie trudno pogodzić się z przerwą w pracy serwerów. Rozwiązanie pośrednie stanowią macierze określane jako hot plug lub hot swap.

Dyski macierzy RAID są umieszczone w sankach (shuttle), które można wyjmować w trakcie pracy systemu. Po wymianie sanek z dyskiem należy włączyć nowy dysk do macierzy i zrekonstruować dane.

Gdy odbywa się to automatycznie, mamy do czynienia z funkcją auto rebuild. Taka automatyzacja wymaga, żeby kontroler mógł komunikować się z sankami dysków (stan napędu; nowe sanki). Jeżeli kontroler i sanki nie mogą się porozumieć, proces rebuild trzeba przeprowadzić ręcznie.

## Podsumowanie

Dostępność systemu komputerowego można znacznie zwiększyć, stosując macierze dyskowe. Macierze nie likwidują jednak niebezpieczeństwa utraty danych.

Niezawodność bliską całkowitej uzyskują podsystemy pamięci masowych, w których wszystkie komponenty, łącznie z kontrolerem, zasilaczem i wentylatorami, są wykonane nadmiarowo. Są takie systemy, jednak ich cena może przyprawić o zawrót głowy.

Jest też pewna zależność, która sprawia, że awarie dysków i innych elementów macierzy rzadko występują rozdzielnie. W praktyce ciągle zdarzają się sytuacje, w których prawdopodobieństwo awarii całej macierzy rośnie lawinowo, np. przepięcia spowodowane uderzeniem pioruna, powodzie i pożary. Wirusy, robaki i trojany atakują [macierze dyskowe](#) równie chętnie, jak pojedyncze komputery.

Nie można też wykluczyć skutków działania czynnika ryzyka numer jeden - człowieka. Nieodwracalna utrata danych zdarza się najczęściej nie w wyniku awarii sprzętu, lecz wskutek błędów personelu obsługującego. Nawet najbezpieczniejsza macierz nie pomoże odzyskać usuniętych lub uszkodzonych plików. Dlatego, nawet jeśli dysponujesz najbardziej wyrafinowaną macierzą

RAID, pamiętaj o podstawowej zasadzie - jedyną naprawdę skuteczną ochroną przed katastrofalną utratą danych to regularne i konsekwentne wykonywanie dobrze zaplanowanych kopii bezpieczeństwa.