

**Studium Informatyki  
i Multimediiów**

**urządzenia techniki  
komputerowej**

*Semestr II*

**Opracował:**  
**mgr inż. Konrad Sobolewski**

## SPIS TREŚCI

1. Organizacja stanowiska pracy, przepisy BHP przy pracach związanych z urządzeniami komputerowymi. ....	3
1.1. Wstęp. ....	3
1.2. Zagrożenia. ....	3
1.3. Dolegliwości. ....	3
1.4. Zagrożenia zdrowia. ....	5
1.5. Odpowiednie stanowisko pracy ....	9
1.6. Podsumowanie. ....	10
1.7. Praktyka użytkowania sprzętu. ....	12
2. Zapis magnetyczny i dyskietki. ....	15
2.1. Techniki zapisu magnetycznego. ....	15
2.2. Metoda FM. ....	16
2.3. Metoda MFM. ....	16
2.4. Metoda RLL. ....	16
2.5. Fizyczna organizacja danych na dyskietce. ....	17
2.6. Uwzględnianie mechanicznych własności napędu. ....	20
2.7. Zabezpieczanie danych – kod CRC. ....	23
3. Dyski twarde. ....	26
3.1. Historia. ....	26
3.2. Budowa. ....	27
3.3. Parametry. ....	41
3.4. Montaż. ....	48
3.5. Uszkodzenia. ....	55
4. Karty graficzne. ....	62
4.1. Monitory. ....	62
4.2. Zasada działania karty graficznej. ....	68
4.3. Przegląd kart graficznych. ....	69
4.4. Karta VGA. ....	71
4.5. Inne rodzaje kart. ....	73
4.6. Akceleratory graficzne. ....	76
4.7. Karty grafiki 3D. ....	78
4.8. Accelerated Graphics Port. ....	79
4.9. VESA Feature connector. ....	80

# **1. ORGANIZACJA STANOWISKA PRACY, PRZEPISY BHP PRZY PRACACH ZWIĄZANYCH Z URZADZENIAMI KOMPUTEROWYMI.**

## **1.1. Wstęp.**

Komputer stał się narzędziem: pracy, poszukiwania informacji, zabawy. Trudno znaleźć obszar ludzkiej działalności gdzie nie ma komputerów, czy jest brak możliwości ich zastosowania. Człowiek w większości przypadków prowadzi „dialog” z komputerem wydając mu polecenia i oczekując na ich realizację przez maszynę, która najczęściej wykonanie sygnalizuje na ekranie. Praca przy komputerze w większości przypadków wykonywana jest w pozycji siedzącej i dla większości osób wydaje się lekką i przyjemną. Jednakże nie jest ona taką. Wymaga ona od użytkownika wzmoczonego wysiłku nie tylko intelektualnego, ale przede wszystkim fizycznego. Wielogodzinna pozycja siedząca, utrzymanie jej, brak ruchu powoduje, niedokrwienie organizmu, zaangażowanie tych samych grupy mięśni przez długi okres czasu powodując ich przemęczenie.

## **1.2. Zagrożenia.**

Czy praca przy komputerze rodzi jakieś zagrożenia dla naszego zdrowia? Zaczynamy odpowiadać na to pytanie zastanówmy się nad innymi pytaniami:

- Czy jedzenie wyłącznie pomidorów stanowi zagrożenie?
- Czy oglądanie telewizji w wygodnym fotelu (z odległości 2 metrów od ekranu) przez 8 godzin dziennie szkodzi?
- Wracając do naszego pytania: Czy praca przy komputerze rodzi jakieś zagrożenia dla naszego zdrowia?

Odpowiedź jest twierdząca: **TAK**. Ale możemy wyeliminować zagrożenia dla zdrowia lub obniżyć do minimum ryzyko powstania dolegliwości zdrowotnych.

Gdzie leżą źródła potencjalnych zagrożeń?:

- w nieergonomicznym stanowisku pracy – w niedostosowaniu stanowiska komputerowego do użytkownika,
- w niepoprawnej statycznej sylwetce przyjętej przez użytkownika podczas pracy z komputerem,
- w powtarzaniu tych samych czynności,
- zbyt długim, nieprzerwanym czasie pracy,
- w braku mikroprzerw i ćwiczeń relaksacyjnych,
- w niepoprawnej organizacji pracy.

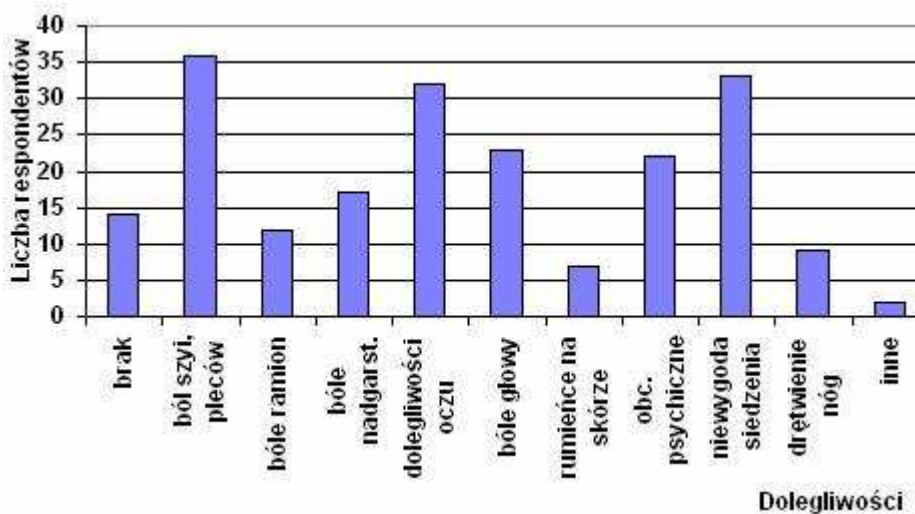
## **1.3. Dolegliwości.**

Długi okres pracy (wielogodzinny, kilkumiesięczny, kilkuletni) z komputerem wraz z niekorzystnymi czynnikami może powodować dolegliwości zdrowotne. Pojawienie się dolegliwości zdrowotnych zależy przede wszystkim od czynników będących potencjalnymi zagrożeniami,

czasu pracy z komputerem, indywidualnych cech użytkownika (w tym podatności). Dolegliwości zdrowotne mogą wystąpić m. in. w postaci:

- ból (głowy, pleców, karku, rąk, nadgarstków itd.),
- zapalenie (ścięgien, stawów, itd.),
- obrzęk, opuchlizna (palców, dłoni, stawów),
- mrowienie, drętwienie (ramion, nóg, itd.),
- osłabienie koordynacji ruchów,
- zmęczenie,
- osłabienie mięśni,
- pieczenie oczu, rozmazywanie się obrazu,
- rozdrażnienie,

Wyniki badań (rys.1.1) przeprowadzone wśród 75 respondentów pracujących na stanowiskach pracy z komputerem wykazały, że około 17% badanych nie stwierdziło żadnych dolegliwości związanych z pracą za biurkiem. Natomiast u 83% respondentów stwierdzono występowanie szeregu dolegliwości zdrowotnych. Ponad 35 respondentów (około 50% badanych) stwierdziło dolegliwości dotyczące bólu: szyi, pleców; dolegliwości oczu, niewygodna siedzenia. Ponad 27% cierpi na bóle głowy i nadmierne obciążenia psychiczne.



Rys. 1.1. Dolegliwości zdrowotne.

Ogólnie dolegliwości zdrowotne są spowodowane przez:

- niepoprawne ustawienie stanowiska komputerowego i jego wyposażenia (krzesło, stół, monitor, klawiatura, itd.) – które wymusza przyjęcie przez użytkownika niekorzystnej sylwetki podczas pracy, nadmierne obciążenie określonej grupy mięśni i układu szkieletowego (kostnego);
- nieodpowiednia organizacja stanowiska pracy i techniki pracy;
- przyjęcie niepoprawnej sylwetki podczas pracy - która wymusza na organizmie dodatkowy wysiłek, nadmierne obciążenie określonej grupy mięśni i układu szkieletowego (kostnego);

- powtarzanie tych samych czynności: np. wprowadzanie informacji do bazy danych, fakturowanie. Operacje tego typu charakteryzują się cyklicznym powtarzaniem tych samych czynności
- nie korzystanie z mikroprzerw i ćwiczeń relaksacyjnych, które umożliwiają regenerację zmęczonych partii ciała.

Aby zapobiec powstawaniu potencjalnych dolegliwości zdrowotnych, lub obniżyć ryzyko ich powstania należy przede wszystkim:

- poprawnie ustawić stanowisko komputerowe:

*Poprawne ustawienie stanowiska pracy – przeciwdziała przyjęciu złej i niekorzystnej sylwetki podczas pracy. Zapobiega nadmiernemu przeciążeniu określonych części układu mięśniowo-szkieletowego.*

- odpowiednio organizować pracę:

*Organizacja stanowiska pracy i techniki pracy – umożliwiają realizację zadań bez nadmiernego (niepotrzebnego) przeciążenia określonych części ciała np. palców dłoni.*

- przyjmować poprawną sylwetkę podczas pracy:

*Poprawna sylwetka podczas pracy – umożliwia wydatkowanie minimum energii oraz minimalne obciążenie ciała. Każda inna sylwetka wymaga od organizmu dodatkowego wysiłku na jej utrzymanie.*

- regularnie korzystać z mikroprzerw i ćwiczeń relaksacyjnych:

*Regularne przerwy i ćwiczenia – pozwalają na „złapanie oddechu”, zaangażowanie innych mięśni do pracy, rozluźnienie zmęczonych mięśni, polepszają przepływ krwi co prowadzi do regeneracji zmęczonych partii ciała.*

- obniżyć czas pracy z komputerem

*Obniżenie czasu pracy z komputerem - eksperci zalecają obniżenie czasu pracy z komputerem do 4 godzin dziennie.*

Specjaliści zalecają, aby zadania realizowane przez pracownika, były przeplatane innymi zadaniami, które angażują odmienne grupy mięśni i dostarczają ruchu.

#### **1.4. Zagrożenia zdrowia.**

„*Repetitive Strain Injury*” - **RSI** jest „zespołem urazów wynikających z chronicznego przeciążenia” objawiających się przewlekłymi bólami ramion, przedramion, przegubów i dłoni, których przyczyny mogą kumulować się, nie dając o sobie znać przez wiele lat. U źródeł RSI leżą nieergonomiczne warunki miejsca pracy. Podaje się, że po sześciu godzinach pracy poziom aktywności lewego palca wskazującego i związanych z tym obciążeń jest porównywalny z pracą nóg podczas czterdziesto kilometrowego marszu. Organizm człowieka nie jest w stanie sprostać tego rodzaju obciążeniom. Dochodzi w związku z tym do usztywnienia mięśni karku, rąk, pleców. Rozwój tych objawów pozostaje utajony w związku z tym, że mózg silnie koncentruje się w czasie pracy na tym, co dzieje się na ekranie i nie potrafi w odpowiednim momencie interpretować sygnałów ostrzegawczych. Są one zagłuszane aż do momentu, gdy stają się nie do

wytrzymania. Choroby wywołane przez komputer są często mylnie diagnozowane. Wielu lekarzy mimo ostrzeżeń specjalistów z zakresu medycyny pracy nigdy o RSI nie słyszało, a także nie uznano RSI za chorobę zawodową. Regularne przerwy podczas pracy oraz organizacja stanowiska pracy zgodnie z zasadami ergonomii pozwala zapobiec tym schorzeniom.

**Narażenie narządu wzroku.** Podczas długich nocy spędzonych podczas surfowania w Internecie możemy nabawić się takich dolegliwości jak przekrwione, łzawiące oczy, zaczerwienione spojówki, uczucie pieczenia i szczypania, bóle głowy, senność i apatia. Większość uskarża się na objawy zaburzenia ostrości widzenia, oraz wrażenie „wysychania” oczu. Nie wszyscy zdają sobie sprawę z tego, że podczas ośmiu godzin pracy polegającej na wpatrywaniu się w szczególności na ekranie nasze oczy wykonują około 30 000 ruchów. Tak przeciążone oczy nie potrafią dostosować się do kolejnych wymagań. Do tego dokładają się często brak ostrości, nadmierna jaskrawość i migotanie ekranu monitora. Jeżeli dochodzi do zmniejszenia częstotliwości mrużania powiekami powoduje to rzadsze zwilżanie powierzchni oka, a to już prawie syndrom Sicca, objawiający się wysychaniem i zmętnieniem rogówki oraz stopniową utratą wzroku. Winę ponoszą też sami użytkownicy. Nie wykorzystują oni zalet dużego ekranu, poprzez dobór odpowiedniej wielkości czcionki, a umieszczają na ekranie jak najwięcej informacji, poprzez stosowanie małych czcionek, kosztem obciążania własnego wzroku. Należy pamiętać o oszczędzaniu wzroku. Stosowanie czarnego pisma na jasnym tle działa najłagodniej. Wyświetlanie monochromatyczne ma zwykle lepszą rozdzielczość niż monitory kolorowe. Jeżeli praca nie wymaga koloru lepiej użyć monitora monochromatycznego. Jednakże, większość programów wymaga zastosowania wielu kolorów. Poza tym do odbioru obrazu kolorowego oko musi się dostosowywać bardziej, ponieważ z uwagi na różną długość fal światła kolorowe niejednakowo załamuje się na soczewce oka. Pisząc tekst nie powinno się także używać czcionek mniejszych niż siedmiopunktowe.

**Nieprawidłowa postawa.** Przeciętny pracownik zatrudniony przy komputerze eksploatuje ten sam komplet mebli przez 80 000 godzin (8 lat!). Jeżeli fotel niedostatecznie wspiera kręgosłup w odcinku lędźwiowym, to plecy są zmuszane do "dopasowywania się" do nieergonomicznego oparcia, a nie odwrotnie jak powinno być. W rezultacie tej nie naturalnej postawy sztywnieje kark i łopatki.

**Promieniowanie.** Praca przy komputerze ma duży wpływ na nasze zdrowie, gdyż wywołuje on różne działania, a eksperci ostrzegają między innymi przed wytwarzanym przez monitor polem elektrostatycznym i promieniowaniem elektromagnetycznym. Istnienie pól elektrostatycznych jest dostrzegalne w postaci znanego każdemu iskrzenia, które towarzyszy kontaktowi skóry z naładowaną elektrycznie powierzchnią ekranu monitora. Pole elektromagnetyczne powoduje polaryzację napięcia elektrostatycznego pomiędzy człowiekiem a ekranem. W rezultacie, gdy pracujemy przy monitorze, nasza twarz przyciąga naładowane cząsteczki kurzu w takim samym stopniu, jak naładowany elektrycznie ekran. Może to być powodem reakcji alergicznych i podrażnienia oczu. Najintensywniejsze pole elektrostatyczne wy-

stępuje przez pierwszy kwadrans pracy monitora. Promieniowanie elektromagnetyczne, doprowadza do impotencji, zakłócenia snu i systemu nerwowego przejawiające się drażliwością oraz zakłócenia rytmu dobowego organizmu, czy wpędza w depresję. Wystarczy przypomnieć sobie jak czuliśmy się po długiej pracy przy komputerze. Często były to objawy jak po ciężkiej pracy fizycznej lub zupełne przemęczenie psychiczne.

**Bezwonne gazy.** Komputer wydziela, o czym nie wszyscy wiedzą, bezwonne gazy - tlenki i furany, które nie są obojętne dla naszego zdrowia. Związki te, zaliczane do jednych z groźniejszych trucizn środowiska naturalnego, wchodzi w skład emulsji, którą pokrywa się obudowy monitorów i jednostek centralnych. Renomowani producenci twierdzą, że środki nieprzyjazne dla środowiska zostały wyeliminowane, ale... Przecież nie każdy kupuje markowy komputer. Należy także pamiętać, że nowy, z tak ogromną radością oczekiwany sprzęt komputerowy, powinien być "wygrzewany" w dobrze wietrzonym pomieszczeniu, by ulotniły się szkodliwe substancje chemiczne zawarte w podzespołach komputera. Jest to szczególnie ważne w salach, w których znajduje się kilka czy nawet kilkanaście komputerów.

**Uzależnienie od Internetu.** Odkryte na początku lat 90 zjawisko uzależnienia od Internetu staje się realnym zagrożeniem także w Polsce. W chwili obecnej nie ma oficjalnej i uznanej nazwy tego zjawiska. Nazywa się go internetoholizmem, internetozależnością, lub infoholizmem - uzależnieniem od informacji. Są jednak naukowcy negujący infoholizm, przyrównując to uzależnienie do hobby. Internetoholik zaniedbuje rodzinę, przyjaciół obowiązków, nie dba o własne zdrowie, rezygnuje z rozrywek i przyjemności. Redukuje czas przeznaczony na sen, a także je byle co i byle szybko. Pierwsi polscy internetoholicy pojawili się w Polsce już w 1993r. Jedynym w kraju ośrodkiem, który od 10 lat zajmuje się leczeniem tego typu uzależnień jest stołeczne Centrum Odwykowe. Z danych Centrum wynika, że większość uzależnionych stanowią osoby młode poniżej 35 roku życia. Zaskakujące może okazać się to, iż wśród najmłodszych przeważają dziewczynki. Coraz częściej zgłaszają się też osoby w wieku około 30 lat, które od kilku lat wyłączyły się całkowicie z utrzymywania kontaktów międzyludzkich i obecnie nie mogą znaleźć partnerów życiowych. Kwestie ochrony zdrowia nabierają jeszcze większego znaczenia, gdy użytkownikami komputerów są dzieci i rozwijająca się młodzież. Do tego dochodzi jeszcze specyfika gier – podczas sesji, będąc w 100% zaabsorbowani grą często świadomie ignorujemy zalecenia dotyczące bezpiecznej pracy i potrafimy w niewygodnej pozycji spędzić kilka godzin. Trzeba więc o siebie dbać.

**RSI - Repetitive Strain Injuries** – obrażenia powstałe na skutek powtarzalnych czynności. Pojawiają się w wyniku powtarzania tych samych czynności fizycznych, ruchów powodując uszkodzenia w ścięgnach, nerwach, mięśniach i innych miękkich tkankach ciała. Takie ruchy wykonywane są również podczas pracy z komputerem i dotyczą: pisania na klawiaturze czy manipu-

lacji myszką. Oczywiście ten problem nie dotyczy wyłącznie osób pracujących przy komputerze, dotyczy on również innych zawodów m.in.: muzyków, fryzjerów. Zagrożenia:

- pogorszenie sprawności fizycznej
- pogorszenie samopoczucia
- dyskomfort fizyczny i psychiczny

Określenie RSI związane jest z następującym światowym nazewnictwem:

- Occupational Overuse Syndrome
- Repetitive Strain Injury
- Work Related Upper Limb Disorder
- Cumulative Trauma Disorder
- Carpel Tunnel Syndrome
- Occupational Cervico-Brachial Disorder
- Muscle Tendon Syndrome
- Musculo-skeletal Injury

Twój organizm emituje sygnały ostrzegając Cię przed „niebezpieczeństwem”. Sygnałami ostrzegawczymi są:

- Ból
- Zapalenie
- Obrzęk, opuchlizna
- Mrowienie, drętwienie
- Osłabienie koordynacji ruchów
- Zmęczenie
- Osłabienie mięśni

Jeżeli odczuwasz dolegliwości zdrowotne powinieneś skonsultować się z lekarzem.

Zapobiec RSI możesz przez:

- Poprawne ustawienie stanowiska komputerowego:

*Poprawne ustawienie stanowiska pracy – przeciwdziała przyjęciu złej i niekorzystnej sylwetki podczas pracy. Przeciwdziała nadmiernemu przeciążeniu określonych elementów układu kostno-szkieletowego.*

- Odpowiednią organizację stanowiska pracy i techniki pracy:

*Organizacja stanowiska pracy i techniki pracy – umożliwiają realizację zadań bez nadmiernego przeciążenia określonych części ciała np. palców dłoni.*

- Przyjęcia poprawnej sylwetki podczas pracy:

*Poprawna sylwetka podczas pracy – umożliwia wydatkowanie minimum energii oraz minimalne obciążenie ciała. Każda inna sylwetka wymaga od organizmu dodatkowego wysiłku na jej utrzymanie*

- Regularne korzystanie z przerwy i ćwiczeń:



*Regularne przerwy i ćwiczenia – pozwalają na „złapanie oddechu”, zaangażowanie innych mięśni do pracy, rozluźnienie zmęczonych mięśni, polepszają przepływ krwi co prowadzi do regeneracji zmęczonych części ciała. Praca przy komputerze to praca siedząca. W dodatku w nie najzdrowszej dla naszego kręgosłupa pozycji (wychylenie torsu nieco do przodu, a więc długotrwałe napięcie wszystkich mięśni). Stwarza to kolejne zagrożenia zdrowotne, którymi są schorzenia kręgosłupa. Łatwo temu zapobiec - wystarczy co 30-60 minut robić krótkie przerwy: wstać, przejść się, chociaż przeciągnąć jak kot. Kilka sekund takich ćwiczeń i już jest lepiej.*

### **1.5. Odpowiednie stanowisko pracy**

Według zaleceń specjalistów, idealnie ustawione stanowisko komputerowe powinno znajdować się na bezokiennej ścianie, bokiem do okna, by światło padało z lewej strony, i w odległości nie mniejszej niż 1 metr. Ogólna zasada jest taka, że na ekranie monitora nie może się odbijać światło (refleksy świetlne mają negatywny wpływ na wzrok, oczy się szybciej i bardziej dotkliwie męczą). Oprawy oświetlenia ogólnego powinny być skierowane tak, by światło padało na podłogę lub blat stołu (nigdy na ekran monitora!). Ustawiając monitor na biurku należy zwrócić uwagę, by górny brzeg ekranu znajdował się nieco poniżej poziomu oczu, nigdy powyżej. Człowiek przy komputerze powinien siedzieć tak, by mniej więcej środek monitora znajdował się na wysokości jego wzroku. Pomieszczenie, w którym pracują komputery powinno być często wietrzone lub być wyposażone w klimatyzację. Należy zadbać także o obecność roślin doniczkowych, najlepiej paproci.

**Biurko.** Wysokość powinna być dopasowana do naszego wzrostu oraz przestrzeń pod nogi powinna wynosić minimum 50 cm głębokości i 70 cm szerokości.

**Krzesło.** Krzesło musi być stabilne, obrotowe, wyposażone w kółka, z możliwością regulacji wysokości siedziska i ustawiania oparcia. Musi ono zapewniać maksymalną swobodę ruchów. Dla pełnej ergonomii powinno posiadać poręcz, aby można było wygodnie opierać łokcie.

**Monitor.** Starsze monitory powinny posiadać filtr chroniący przed promieniowaniem elektromagnetycznym, oraz dobrze jeżeli posiadają filtry polaryzacyjne powodujące mniejsze zmęczenie wzroku w niekorzystnych warunkach zewnętrznego oświetlenia. Wskazane jest także, w monitorach posiadających odpowiednią funkcję, stosować rozmagnetyzowanie powierzchni ekranu. Należy też zwracać uwagę na wartość częstotliwości odświeżania ekranu. Według zaleceń amerykańskich odległość użytkownika od ekranu monitora komputerowego powinna wynosić ok. 70 cm, a z tyłu i boków - najmniej 130 cm. Górny brzeg ekranu powinien być nieco poniżej poziomu oczu, a w żadnym wypadku powyżej tego poziomu. Pamiętając o oszczędzaniu wzroku należy wykorzystywać zalety dużego ekranu i wiedzieć, że stosowanie czarnego pisma na jasnym tle działa najłagodniej (niektórzy uczniowie uwielbiają odwrotne zestawienie, a widoczne jest to podczas projektowania stron internetowych). Wyświetlanie monochromatyczne ma zwykle lepszą rozdzielczość niż kolorowe. Pisząc tekst dostosowujemy wielkość czcionki do własnych upodobań i wygody - zalecana bezszeryfowa Arial lub szeryfowa Times New Roman w

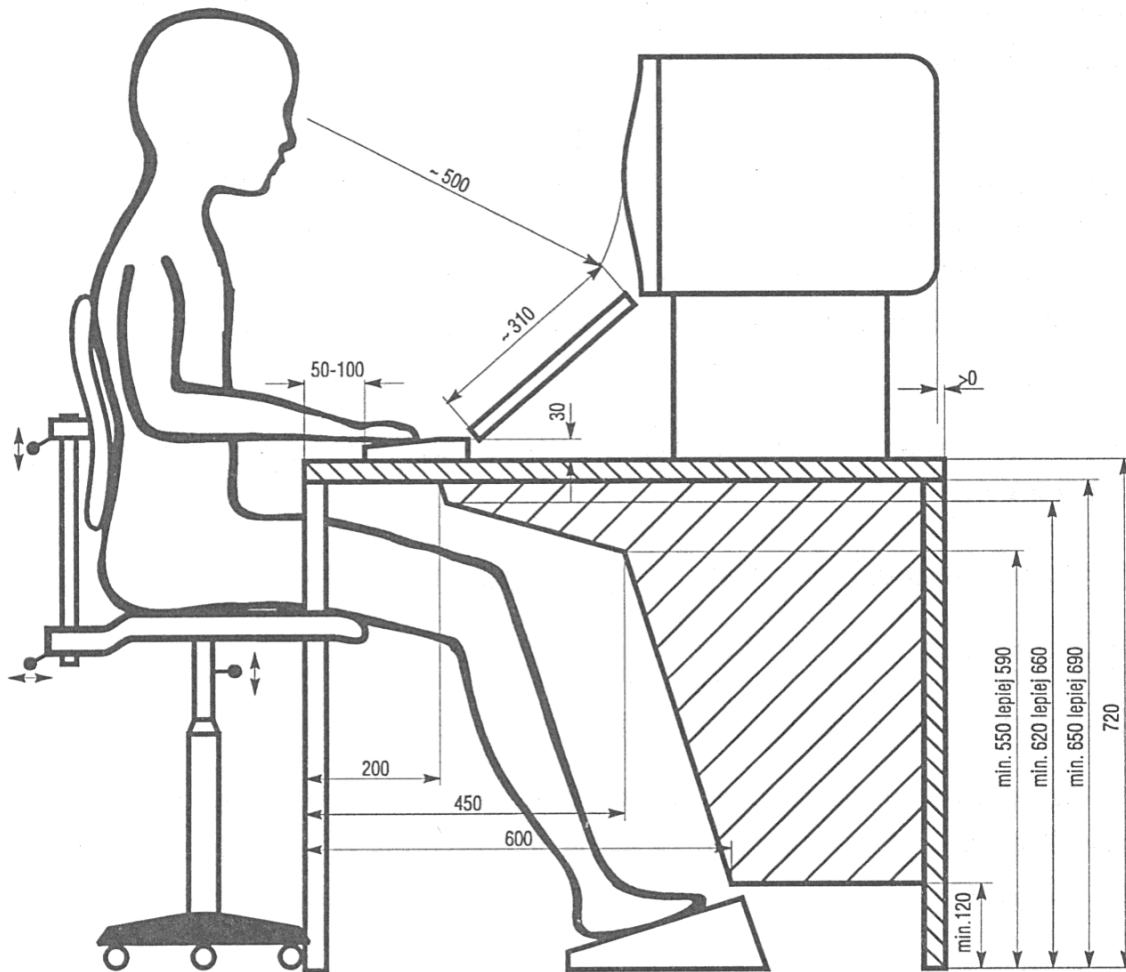
wielkości 12-14 punktów. Dbamy także, by częstotliwość odświeżania obrazu (migotanie obrazu) była optymalna. Zadaniem użytkowników Windows powinno być tylko sprawdzenie, czy system właściwie rozpoznał model monitora i poprawnie ustawił istotne parametry pracy. W systemach Windows 95/98 odbywa się to następująco: start/ustawienia/panel sterowania/ekran/zakładka ustawienia/przycisk zaawansowane/ zakładka karta/ opcja szybkość odświeżania. Zgodnie z najnowszą szwedzką normą TC099 zaleca się ustawienie parametru szybkość odświeżania na 85Hzw rozdzielczości 1024x768. Te niewielkie usprawnienia, bez żadnych nakładów finansowych i zależne tylko od nas, mają duży wpływ na komfort pracy i ochronę naszego wzroku. A jakie są realia? Powszechną praktyką jest ustawianie monitora komputera w miejscu, gdzie mamy trochę wolnej powierzchni. A skutki? Bóle głowy wywołane odbłaskami i refleksami na ekranie, zmęczenie wzroku z powodu złych ustawień parametrów pracy monitora, mroczki przed oczami, łzawienie, swędzenie, szczypanie itp.

**Pomieszczenie.** Pomieszczenie w którym pracują komputery powinno posiadać klimatyzację, lub powinno być często wietrzone. Dobrze jeżeli stoją w nim kwiaty doniczkowe. Na stanowisko powinno przypadać 6 m<sup>2</sup>. Monitory powinny być tak ustawione, by nie odbijało się w nim światło naturalne ani sztuczne. Refleksy świetlne powodują bardzo szybko zmęczenie wzroku, a pochylanie się na boki i do przodu doprowadza do bólów kręgosłupa. Stanowisko pracy powinno być oświetlone światłem rozproszonym nie powodującym olśnienia. Dobrze również jest ograniczyć emitowane przez monitor promieniowanie widzialne.

**Ustawienie komputerów.** Odległości między sąsiednimi monitorami powinny wynosić minimum 0,6 m. Osoba pracująca z tyłu monitora powinna znajdować się w odległości co najmniej 0,8 m. Najlepiej, gdy monitory ustawione są tyłem jeden względem drugiego.

### **1.6. Podsumowanie.**

Jak wykazują badania, większość osób zatrudnionych przy komputerach skarży się na określone dolegliwości: bóle głowy, kręgosłupa, łzawienie oczu, napięcie mięśni, ogólne rozdrażnienie. Oczywiście, każda praca niesie z sobą różnego typu niebezpieczeństwa i schorzenia; nie inaczej jest z komputerami. Można jednak temu przeciwdziałać (patrz rysunek 14).



Rys. 1.2. Prawidłowo zorganizowane stanowisko do pracy z komputerem.

Wymieńmy główne czynniki wpływające na prawidłową i bezpieczną pracę z komputerem (niezależnie od typu posiadanego monitora):

- odpowiednie meble z krzesłem o regulowanej wysokości oraz kącie nachylenia oparcia; stosownie duży blat, pozwalający zachować właściwą odległość monitora od oczu (35-70 centymetrów),
- odpowiednie oświetlenie, dające niezbyt duże różnice jasności ekranu monitora i pomieszczenia oraz eliminujące odbicia,
- odpowiedni mikroklimat: temperatura 21-22 stopnie Celsjusza, wilgotność powietrza 50-65%, lekki ruch powietrza.

Obecnie najwięcej kontrowersji wywołuje ustalenie faktycznych skutków wpływu monitora na zdrowie człowieka, choć nowoczesne monitory emitują mniej promieniowania (typ LR) i mają mniej dokuczliwe migotanie (typ NI). Mimo to ostatnio pojawiły się informacje o konieczności zapewnienia przez pracodawcę okularów pracującym przy ekranie komputera - tak więc problemy nadal istnieją. Pracujący monitor emituje:

- pole elektrostatyczne (wywołujące wiatr elektryczny), które nie powinno przekraczać 60 kV/m,

- pole elektromagnetyczne bardzo małych częstotliwości,
- promieniowanie nadfioletowe powodujące zapalenie spojówek i rumień skóry,
- śladowe promieniowanie rentgenowskie.

Oprócz promieniowania, negatywne skutki mogą wywołać czynniki wynikające z nieprawidłowego ustawienia monitora lub z jego wadliwego działania bądź nadmiernego zużycia:

- odbicie oświetlenia zewnętrznego od ekranu monitora, które zmusza do nadmiernego wyężdżania wzroku,
- migotanie obrazu, rozmycie kolorów i brak ostrości, które występuje w monitorach o niskiej jakości i powoduje szybkie zmęczenie wzroku.

Pozostawiając lekarzom wątpliwości, w jakim stopniu praca z monitorem jest szkodliwa dla zdrowia, postaramy się mimo wszystko poprawić warunki pracy poprzez:

- używanie bezpiecznych, niemigających, o zmniejszonym promieniowaniu monitorów (LR, NI),
- stosowanie filtrów ochronnych zmniejszających najbardziej szkodliwe promieniowanie elektrostatyczne co najmniej 200 razy i odprowadzających ładunki przewodem połączonym z masą komputera,
- zachowanie maksymalnej odległości monitora od oczu oraz skierowanie tyłu monitora do ścian, aby nie narażać innych osób,
- zachowanie podanej wyżej wilgotności powietrza,
- wykonywanie ćwiczeń relaksujących wzrok.

Praca z komputerem to obok wysiłku intelektualnego również zmęczenie statyczne. Dlatego nie zapominajmy o wykonywaniu co jakiś czas ćwiczeń fizycznych poprawiających samopoczucie. Mogą to być: skręty ciała, wyciąganie ramion, zwieszanie głowy, naprężanie i rozluźnianie mięśni nóg, głębokie oddychanie, masaż dłoni, palców, skroni i karku.

Warto też wspomnieć o klawiaturze. Ostatnie badania wykazały, że osoby piszące długo na tradycyjnej klawiaturze cierpią na schorzenia wynikające z nienaturalnego ułożenia dłoni. Powstały więc ostatnio ergonomiczne klawiatury, w których lewa i prawa strona ustawione są pod odpowiednim kątem. Korzystanie z nich wymaga zmiany przyzwyczajeń - czy są rzeczywiście skuteczne, trudno powiedzieć. Po krótkotrwałej modzie ostatnio nie widać ich zbyt wielu.

### **1.7. Praktyka użytkowania sprzętu.**

Na zakończenie rozdziału poruszającego wzajemne relacje komputer-człowiek warto wspomnieć o samym użytkowaniu sprzętu. Chodzi o to, żeby z jednej strony zapewnić trwałość użytkowania, z drugiej bezpieczeństwo danych i użytkownika. Poruszane tu zagadnienia łączą się z zagadnieniami BHP. Przede wszystkim trzeba komputer prawidłowo ustawić i podłączyć. Zaleca się stosowanie listwy z filtrem zabezpieczającym przed zakłóceniami i przepięciami w sieci. Zapewnia to prawidłową pracę komputera oraz zabezpiecza przed ewentualnym uszkodzeniem, które może być spowodowane nagłą zmianą napięcia w sieci. Oczywiście listwa powinna być podłączona do uziemionego gniazda. Znów chodzi tu nie tylko o bezpieczeństwo

użytkownika, lecz także o sprzęt, który może zostać uszkodzony przez ładunki elektrostatyczne powstające przy braku uziemienia. Jeśli komputer "kopie", gdy zbliżymy do niego rękę, świadczy to o nieprawidłowym uziemieniu.

Zasilanie komputera jest bardzo istotne także ze względu na możliwość utraty danych w przypadku nagłej awarii. Przeciążenie sieci, wyłączenie prądu (awaryjne bądź przypadkowe) podczas działania niektórych programów mogą prowadzić do utraty ważnych informacji. Dlatego warto wyposażyć stanowisko pracy w urządzenia do podtrzymywania napięcia, tzw. UPS. W przypadku braku zasilania UPS przejmuje dostawę prądu ze swego układu akumulatorowego. Zależnie od typu urządzenia mamy wówczas od kilku do kilkunastu minut na zakończenie działania programów i zabezpieczenie plików z danymi. Jest to szczególnie istotne w systemach sieciowych oraz wszędzie tam, gdzie przetwarzane są bazy danych (księgowość, magazyn, kadry). Nagłe wyłączenie zasilania powoduje zwykle utratę informacji, a nieprawidłowe zamknięcie plików prowadzi często do ich uszkodzenia.

Jeżeli komputery połączone są w sieć, to kable sieciowe powinny być prowadzone wzdłuż ścian, a nie w miejscach, w których mogą być narażone na nacisk. Ponadto należy chronić je przed wilgocią i dbać o prawidłowe umocowanie wtyczek. Większość problemów sygnalizowanych przy użytkowaniu sieci komputerowych wynika z wad okablowania lub niedokładnego umocowania wtyków.

Komputer powinien być prawidłowo ustawiony. Ważny jest wygodny dostęp do klawiatury i do myszy - dziś wymiennie używamy obu urządzeń i istotna jest możliwość szybkiej zmiany. Mysz powinna być zawsze na podkładce. Przedłuża to jej żywotność i zapobiega zabrudzeniu kulki prowadzącej. Monitor powinien być ustawiony tyłem do ściany i tyłem do światła dziennego. Inne ustawienie może spowodować odbłaski utrudniające pracę. Wpływ odbłasków redukuje ekrany polaryzujące, jeśli na komputer nie pada światło słoneczne. Światło sztuczne także powinno się znajdować za ekranem - monitor musi być właściwie ustawiony względem źródła światła. Ze względu na temperaturę i elektryzowanie się komputer łatwo się kurzy, dlatego należy regularnie czyścić ekran i klawiaturę. Na rynku jest wiele specjalnych płynów i pianek do czyszczenia sprzętu. Zabrudzeniu ulegają także stacje dyskiety oraz czytniki CD-ROM. Co kilka miesięcy warto zawieźć sprzęt do czyszczenia. Ewentualnie można samemu zdjąć obudowę i oczyścić komputer za pomocą odkurzacza. Sygnałem zanieczyszczenia jest nieprawidłowa praca urządzeń.

Do czyszczenia dyskowych głowic odczytu/zapisu służą specjalne dyski czyszczące, które należy zgodnie z instrukcją włożyć do urządzenia raz-dwa razy w miesiącu. To także przedłuża żywotność sprzętu. Trzeba też pamiętać, że układy scalone jednostki centralnej są czułe na temperaturę. Często niespodziewane problemy podczas pracy i sygnalizowanie nietypowych błędów spowodowane jest przegrzaniem pamięci bądź procesora. Jest to zwykle spowodowane awarią wiatraczka znajdującego się w obudowie. Jest to pierwszy element, który należy w takiej sytuacji sprawdzić. Oczywiście komputery trzeba chronić przed wstrząsami i uszkodzenia-

mi mechanicznymi. Nie wolno przenosić sprzętu podczas pracy, szczególnie podczas pracy dysku. Odległość głowic od powierzchni dysku twardego mierzona jest w mikronach i wstrząsy podczas zapisu lub odczytu mogą spowodować uszkodzenia powierzchni. Po wyłączeniu urządzenia głowice są "parkowane" czyli ustawiane w gniazdach z dala od powierzchni nośnika. Wówczas przenoszenie sprzętu nie powoduje żadnych ujemnych skutków.

## **2. ZAPIS MAGNETYCZNY I DyskiETKI.**

Napęd - przy czym na razie nie ma znaczenia, jakiego rodzaju - różni się od innych pamięci (jak np. RAM czy Cache wykonanych przeważnie z półprzewodnikowych obwodów elektrycznych) tym, że składa się z wielu mniej lub bardziej skomplikowanych części mechanicznych. Jedną z konsekwencji tego faktu jest naturalny proces zużywania się napędów co po dłuższej eksploatacji prowadzi do błędów w zapisie i odczycie, a nawet do całkowitej niesprawności. Inaczej jest w przypadku elektronicznych nośników pamięci, które nie zużywają się w normalnej eksploatacji, tylko w wyniku niewłaściwego obchodzenia się z nimi, np. podczas pracy z niewłaściwymi parametrami (tzw. "przetaktowanie"), w efekcie przegrzania lub niewłaściwego napięcia zasilającego. Urządzenie składające się z mechanicznych i obracających się części, jakim jest każdy napęd dyskietek czy twardy dysk wrażliwe jest na działanie takich czynników zewnętrznych, jak upuszczenie czy wstrząsy w czasie pracy i może w ich wyniku zostać poważnie uszkodzone. O tym ważnym rozróżnieniu wśród nośników pamięci powinniśmy zawsze pamiętać. Nie jest niestety rzadkością, że informacje na twardym dysku uważane są za całkowicie bezpieczne, co jest łagodnie mówiąc dużą lekkomyślnością. Więcej o zabezpieczaniu ważnych informacji powiemy w jednym z następnych rozdziałów.

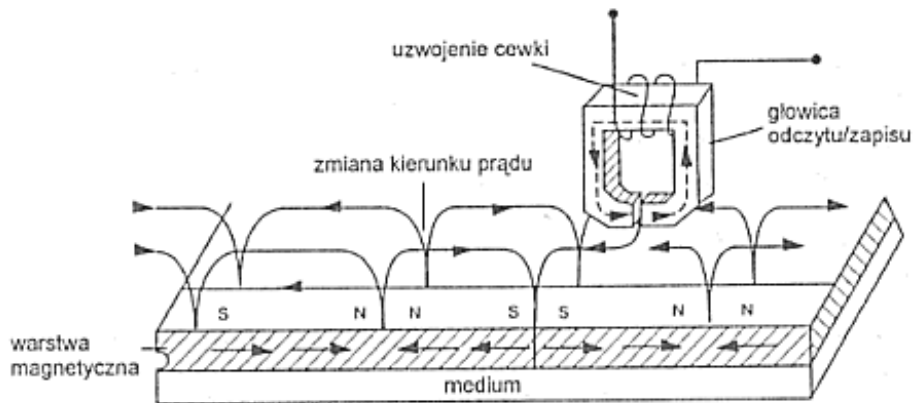
### **2.1. Techniki zapisu magnetycznego.**

Wśród napędów rozróżnia się dwa zasadniczo różne sposoby zapisu, a więc i odczytu informacji: magnetyczny jak w przypadku dyskietek albo twardych dysków i optyczny jak np. CD-R. Spotyka się także techniki stanowiące kombinację obu metod, określane mianem magneto-optyczne (MO). Napędy MO oraz przeróżne warianty zapisu optycznego jak CDR, CD-RW i DVD omówimy w dalszej części skryptu, a teraz skoncentrujemy się wyłącznie na napędach dyskietek wykorzystujących magnetyczny sposób zapisu. W języku angielskim przyjęł się zwyczaj rozróżniania obu nośników pamięci (magnetycznych i optycznych) na podstawie nazwy danego nośnika. I tak nośniki magnetyczne mają w nazwie zawsze literkę „k” np. Floppy Disk, a optyczne „c” np. Compact Disc.

w każdym pececie znajduje się specjalny kontroler napędów, który z reguły steruje zarówno napędem dyskietek jak i dyskiem twardym (jednym lub kilkoma). Cyfrowe (zerojedynkowe) informacje przeznaczone do zapisu trafiają do kontrolera. Następnie zamieniane są w impulsy magnetyczne, a dokładniej, zmienia się kierunek przepływu prądu w cewce powodując zmianę bieguna pola magnetycznego, aby mogły być zapisane na magnetycznej powierzchni dyskietki lub twardego dysku. Nie wystarczy jednak przy tym zapisywać dane bity jeden po drugim, trzeba dodatkowo zaznaczyć gdzie dany bit się zaczyna, a gdzie kończy. Może się to dziać na kilka różnych sposobów:

- FM - Frequence Modulation
- MFM - Modified Frequence Modulation
- RLL - Run Length Limited

- ARLL - Advanced Run Length Limited



Rys. 2.1. Zasada zapisu magnetycznego. Zmiana biegunowości odpowiada (zerojedynkowej) informacji zapisanej na nośniku.

## 2.2. Metoda FM.

W tej najstarszej technice każdej jedyńce odpowiada zmiana biegunowości pola magnetycznego, natomiast przy zerze pozostaje ona niezmieniona. Początek każdego bitu (obojętnie zero czy jedynka) wymaga dodatkowo tzw. sygnału taktującego, czyli dodatkowej zmiany biegunowości. Tak, więc do zapisu „1” potrzeba dwu zmian biegunowości, a do „0” jednej (tylko początek taktu). Metoda ta nie wykorzystuje optymalnie miejsca na dyskietce lub twardym dysku, gdyż wymaga zbyt wielu zmian biegunowości. Im mniej tych zmian, tym więcej danych można zmieścić na tej samej powierzchni. Metoda ta jest już dosyć przestarzała i używana jest właściwie tylko przez stary format dyskietki IBM.

## 2.3. Metoda MFM.

Lepsze wykorzystanie miejsca na dysku umożliwia zmodyfikowana metoda modulacji częstotliwości - MFM. Sygnał taktujący został tutaj niejako "przejęty" przez strumień informacji. Rozwiązanie to zakłada jednak stałą prędkość obrotową dysku, tylko wtedy każdemu bitowi przyporządkowany jest taki sam, co do wielkości obszar. W technice MFM zmiana biegunowości odbywa się dla każdej jedyńki tylko w środku danego obszaru, natomiast każde zero rozpoczyna się na początku takiego obszaru, ale tylko wtedy, gdy wcześniejszy bit nie był jedyńką. (zob. rys. 2.2). Oznacza to 100% wzrost gęstości zapisu w stosunku do metody FM. Przez wiele lat była to najczęściej stosowana technika zapisu na twardych dyskach. Współpracowały one z kontrolerami w standardzie ST506/412, który umożliwiał zapis w 17 sektorach po 512 bajtów w każdym. Standard ten stosowany jest ciągle w spotykanych obecnie napędach dyskietek 3,5" i 5,25".

## 2.4. Metoda RLL.

Dalsze zagęszczenie zapisu, czyli zmniejszenie ilości zmian biegunowości pola magnetycznego na jednostkę informacji, przyniosła metoda kodowania informacji. Pewnym grupom bitów przyporządkowano kod o zmiennej długości. Odkodowanie następowało według następującej



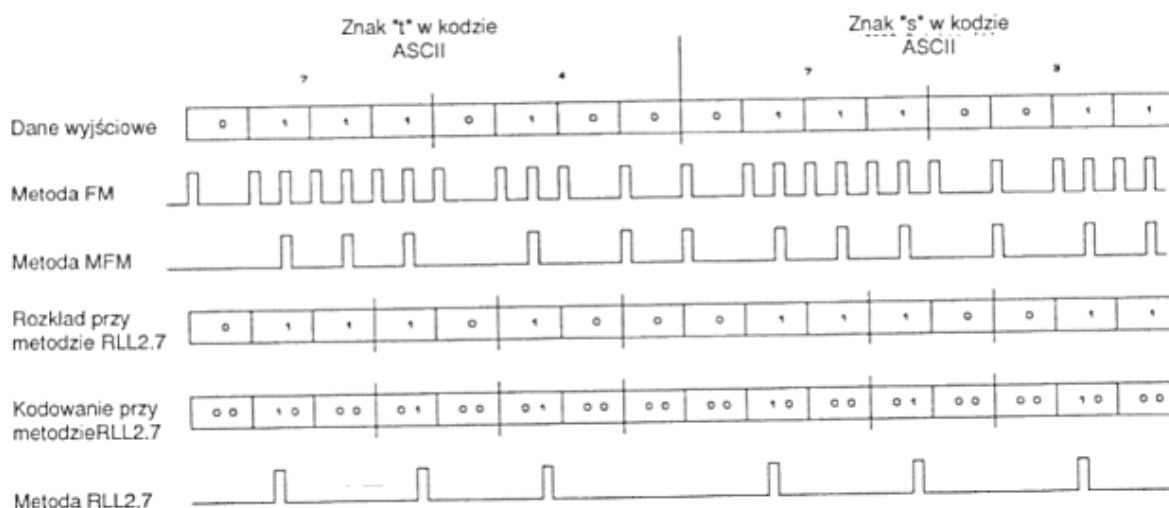
zasady: między dwoma jedynekami znajdować się musi zawsze ściśle określona liczba zer. Procedura ta - a istnieje cały ich szereg - nazywana jest metodą RLL (Run Length Limited). Np. w metodzie RLL2.7 pomiędzy dwie jedynki wstawianych jest od 2 - 7 zer, zaś w RLL3.9, znanej również pod nazwą ARLL (Advanced RLL), od 3 do 9.

Twarde dyski IDE i SCSI wykorzystują zwykle procedurę RLL1.7 lub jakąś jej odmianę. Tabela poniżej pokazuje przyporządkowanie konkretnym grupom bitów (jest ich tylko siedem) kodów w procedurze RLL2.7.

Grupa bitów	Kod RLL2.7
000	000100
10	0100
010	100100
0010	00100100
11	1000
011	001000
0011	00001000

Tabela 1. Kody poszczególnych grup bitów w metodzie RLL.

Procedura RLL2.7 pozwala na 50% zwiększenie pojemności twardego dysku w porównaniu z metodą MFM, na każdej ścieżce zmieści się bowiem 26 a nie 17 sektorów. Procedura RLL3.9 zwiększa tę pojemność już o 100% (34 sektory na każdej ścieżce). Na rysunku 2.2 pokazano różnice w poszczególnych procedurach - najłatwiej tu zrozumieć, na czym polega oszczędność metody RLL.

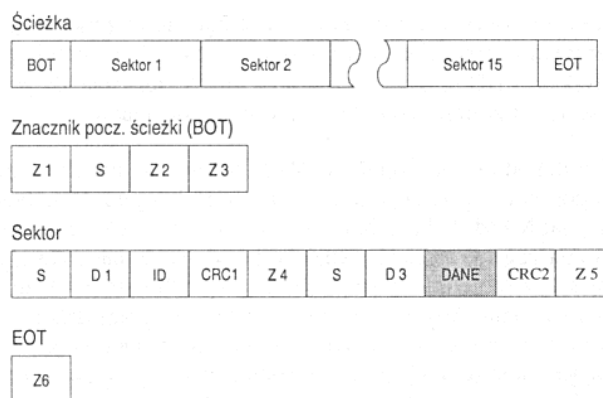


Rys. 2.2. Porównanie technik zapisu magnetycznego. Każdej zmianie biegunowości odpowiada impuls High. Im mniej zmian biegunowości potrzeba do zapisu informacji tym gęściej można ją zapisać.

## 2.5. Fizyczna organizacja danych na dyskiecie.

Informacja nie jest zapisywana na dyskiecie w postaci jednolitego ciągu bajtów, lecz zorganizowana jest w tzw. sektory. Każdy sektor zawiera 512 bajtów, co jest jednocześnie najmniej-

szą porcją informacji, jaka może być z dyskietki odczytana. Z punktu widzenia systemu operacyjnego logiczną strukturę dyskietki tworzą pliki i katalogi, zaś niektóre informacje (zapisane w ściśle określonym miejscu na dysku) mają specjalne znaczenie, np. sektor ładujący (*Boot Sector*), katalog (*Directory*). Każdy sektor należy ponadto do tzw. jednostki alokacji (*Cluster*) o kolejnym numerze logicznym, a jednocześnie, z fizycznego punktu widzenia, wchodzi w skład ścieżki (*Track*). Na szczególną uwagę zasługuje tzw. tablica alokacji FAT (*File Allocation Table*), będąca specjalną „łańcuchową” strukturą informującą system, które jednostki alokacji (i w jakiej kolejności) składają się na dany plik. System operacyjny korzysta bardzo intensywnie ze wsparcia kontrolera napędu dysków elastycznych, któremu powierza się zadanie administrowania fizyczną powierzchnią dyskietki. To kontroler musi wiedzieć, gdzie odszukać żądany sektor oraz jak radzić sobie z sektorami uszkodzonymi. Operacja formatowania dyskietki nanosi na nią określoną fizyczną strukturę ścieżek i ich sektorów. Oprócz znanych nam 512-bajtowych sektorów logicznych dyskietka zawiera wiele dodatkowych pól służących kontrolerowi do administracji sektorami (odnajdywania sektorów i przechowywania informacji o ich stanie) oraz pomagających korygować niedokładności mechanicznego wykonania napędu i dyskietki, jak również wahania prędkości obrotowej dysku. Dodatkowe pola na ścieżce pomagają w uzyskaniu stabilnego ciągu impulsów synchronizacyjnych potrzebnych w obróbce danych (szczególnie MFM). W pierwszym przybliżeniu fizyczna organizacja powierzchni dyskietki składa się z jednokowych, koncentrycznych ścieżek. Każda z nich zawiera znacznik początku ścieżki BOT (*Beginning of Track*), sektory w liczbie zależnej od rodzaju dyskietki oraz znacznik końca ścieżki EOT (*End of Track*). Budowę jednej ścieżki dyskietki obrazuje rysunek 2.3.



Rys. 2.3. Organizacja danych na ścieżce dyskietki.

Kontrolery starszej generacji wyposażone były w mikrokontroler NEC mPD765 lub jego funkcjonalny odpowiednik Intel 8272. Układy te miały zaimplementowaną możliwość przełączenia w tryb pracy FM (format używany na dyskietkach komputerów IBM PC). Organizacja zapisu na dyskietce, tj. układ pól dodatkowych i ich znaczenie, są takie same dla formatów FM i MFM. Inne są tylko długości i zawartości pewnych znaczników. Znacznik synchronizacyjny S (ciąg zer) jest w przypadku formatu FM skrócony do połowy. FM niesie w sobie bezpośrednio impulsy synchronizacji niezależnie od składu strumienia danych i łatwiej jest utrzymywać generator w

stanie zgodności fazowej. Format FM nie ma w obecnych czasach praktycznego znaczenia i nie będziemy się nim dalej zajmować. Poszczególne pola ścieżki formatu MFM mają następującą zawartość:

<b>Pole</b>	<b>Zawartość</b>
Z1	80 bajtów o wartości 4Eh
S	12 bajtów o wartości 00h
Z2	4 bajty o wartościach C2h, C2h, C2h, FCh
Z3	50 bajtów o wartości 4Eh
D1	4 bajty o wartościach A1h, A1h, A1h, FEh
ID	<p>Metryka adresowa sektora - 4 bajty o następującym znaczeniu:</p> <p>bajt 1 - numer ścieżki,</p> <p>bajt 2 - numer głowicy,</p> <p>bajt 3 - numer sektora,</p> <p>bajt 4 - rozmiar sektora kodowany wg klucza:</p> <p>000 - 128 bajtów,</p> <p>001 - 256 bajtów,</p> <p>010 - 512 bajtów,</p> <p>011 - 1 kB,</p> <p>...</p> <p>111-16kB</p>
CRC1	16-bitowy kod CRC <sup>(A)</sup> zabezpieczający pola D1 oraz ID
CRC2	16-bitowy kod CRC <sup>(A)</sup> zabezpieczający pole danych sektora
Z4	22 bajty o wartości 4Eh
D3	4 bajty o wartościach A1h, A1h, A1h, FBh
Z5	80 bajtów o wartości 4Eh
Z6	Bajty o wartości 4Eh - znacznik końca ścieżki (EOT)

<sup>(A)</sup> Budowa kodu CRC omówiona jest na końcu rozdziału.

System operacyjny może mieć wpływ jedynie na zawartość pola ID. Wszystkie funkcje formatujące ścieżkę muszą przekazać kontrolerowi napędu dysków adres w pamięci, pod którym znajduje się odpowiednio przygotowana 4-bajtowa struktura opisująca oddzielnie każdy sektor formatowanej ścieżki. Każda dyskietka ma w pobliżu swego środka mały otwór, który odsłania przy każdym obrocie fotokomórkę zainstalowaną w napędzie. W ten sposób wytwarzany jest przebieg prostokątny, który podawany jest linią IDX przewodu połączeniowego do kontrolera. Stanowi on przybliżoną informację o początku ścieżki. Jeden bajt danych w przypadku dyskietki HD zajmuje około 16  $\mu\text{m}$  na powierzchni nośnika, tak więc dokładne zlokalizowanie początku ścieżki za pomocą otworu o średnicy 2 mm nie jest oczywiście możliwe. Dla kontrolera sygnałem początku ścieżki jest napotkanie znacznika Z1. W następnej kolejności odczytywany jest

blok impulsów synchronizacyjnych S, które dopasowują dokładnie częstotliwość i fazę drgań generatorów potrzebnych do zdekodowania sygnału MFM. Znacznik S zawiera same zera i powoduje generację czystego sygnału synchronizacyjnego. Zmniejszany jest tym samym do minimum szkodliwy wpływ chwilowych zmian prędkości obrotowej dysku. Sygnatura Z2 informuje kontroler, że po niej następować będą sektory ścieżki. Resztę znacznika początku ścieżki wypełnia szczelina Z3. Każdy z następujących teraz sektorów ścieżki ma jednakową budowę i zawiera 10 pól. Pierwsze z nich to pakiet sygnałów synchronizacyjnych S o budowie i znaczeniu takim, jak w znaczniku początku ścieżki. Następnym jest 4-bitowa sygnatura D1 przygotowująca kontroler na przyjęcie metryki adresowej sektora ID. Pole to jest 4-bajtowym (ścieżka - - głowica - sektor - rozmiar) adresem sektora na dysku i stanowi zawartość bloku formatującego, który należy przygotować w pamięci przed wywołaniem funkcji 05h przerwania 13h BIOS-u. Metryka adresowa ID i jej sygnatura D1 zabezpieczone są 16-bitowym kodem CRC umieszczonym w polu CRC-1. Teraz następuje ponowne zsynchronizowanie generatora MFM z aktualną prędkością obrotową dysku (pole S) i pod głowicami pojawia się znacznik początku danych sektora D3. Pole danych sektora nie musi mieć standardowej długości 512 bajtów, choć jest to wartość powszechnie używana. Długość sektora określona jest w czwartym bajcie pola ID. Może ona sięgać teoretycznie nawet 16 kB, jednak w polach danych o długości większej od 2 kB mogą występować częste przekłamania spowodowane rozsynchronizowywaniem się generatorów MFM. Pole danych zabezpieczone jest własną sumą kontrolną umieszczaną w polu CRC-2. Szczelina Z5 daje kontrolerowi czas na obliczenie tej sumy i weryfikację wyniku. Pamiętajmy, że kontroler napędu dysków elastycznych nie ma buforu ścieżki i wszystkie informacje opracowywane są w czasie rzeczywistym, tj. w miarę pojawiania się danych pod głowicami napędu. Szczelina Z5 ma jeszcze dodatkowe znaczenie - jest ona elastycznym buforem między sektorami. Rzeczywiste długości pola danych sektorów mogą się zmieniać, bowiem minimalna zmiana tarcia głowicy o powierzchnię dyskietki (i tym samym zmiana prędkości obrotowej) podczas zapisywania danych nie może być kompensowana zmianą częstotliwości zegara, który jest synchronizowany jedynie polem S. Mogą zatem powstawać lokalne wahania gęstości zapisu, co prowadzi do zmian długości pól danych w sektorach. Pole Z5 zmniejsza zatem szanse na nałożenie się danych sektora na następujące po nich pola. Całkowita długość sektora, po uwzględnieniu wszystkich pól dodatkowych, wynosi więc 654 bajty. Ten fakt, a także istnienie znaczników BOT i EOT, stanowi o różnicy między pojemnością dysku sformatowanego i niesformatowanego. Koniec ścieżki (brak dalszych danych) sygnalizowany jest znacznikiem Z6.

## **2.6. Uwzględnianie mechanicznych własności napędu.**

Głowice zapisujące - odczytujące napędu dysków elastycznych muszą mieć możliwość osiągnięcia każdej ze ścieżek, tj. przemieszczania się nad powierzchnią dyskietki. Głowice umieszczone są na wózku napędzanym silnikiem krokowym lub liniowym. Najczęściej oś tego silnika nawija taśmę przymocowaną do wózka, ale stosowane są również inne mechanizmy, np. przekładnia ślimakowa. Cały ten układ ma pewną bezwładność, częstotliwość rezonansu mecha-

nicznego i inne parametry rzeczywiste. Elektroniczny układ sterowania musi więc uwzględniać charakterystykę dynamiczną całego zespołu. Reguły zachowania kontrolera i układu sterowania zawarte są w mikroprogramie zapisanym w pamięci stałej (Firmware). Możliwość korekcji błędu położenia głowic napędu dyskietek jest bardzo ograniczona. Układ sterowania położeniem wózka z głowicami pracuje w otwartej pętli sprzężenia zwrotnego, tzn. kontroler wydaje polecenie umieszczenia ustawienia głowic nad żądaną ścieżką, ale nie otrzymuje informacji o ich faktycznym położeniu. Jedynym punktem orientacyjnym jest ścieżka zerowa dysku. Algorytm działania układu pozycjonowania jest następujący:

- wydawane jest polecenie ustawienia głowic nad ścieżką zerową. Silnik krokowy otrzymuje impulsy w liczbie równej maksymalnej liczbie ścieżek dla danego typu nośnika. Osiągnięcie tej pozycji przez głowicę potwierdzone jest sygnałem z czujnika położenia (najczęściej fotokomórki). Jeżeli sygnał ten nie nadchodzi, sygnalizowany jest błąd.
- silnik krokowy otrzymuje impulsy w liczbie żądanego numeru ścieżki.
- jeżeli kontroler rozpoznaje inną ścieżkę, mamy do czynienia z błędem pozycjonowania (kalibracji).
- BIOS inicjuje kilkakrotną próbę zapisu i (lub) odczytu. Brak sukcesu oznacza konieczność powtórnej kalibracji, która rozpoczyna się od najazdu na ścieżkę zerową. Próba taka może być ponawiana kilka razy.

Szczególnie podatne na błędy są oczywiście operacje, podczas których głowice przemierzają duże odległości, np. ze ścieżki 0 na 79. Wszystkie niedokładności mechanizmu mogą się wtedy nawarstwiać. Aby zmniejszyć ryzyko błędów pozycjonowania, wprowadzono trzy parametry, które w pewnym stopniu odzwierciedlają rzeczywiste właściwości mechaniczne napędu. Kontroler uwzględnia te dane, wprowadzając do pracy swoich układów wykonawczych trzy opóźnienia:

- T1 Czas opóźnienia opuszczenia głowic - uwzględnia czas wygaśnięcia drgań głowic po zakończeniu procesu pozycjonowania, a przed opuszczeniem ich na powierzchnię dyskietki.
- T2 Czas opóźnienia podnoszenia głowic - jego wprowadzenie ułatwia czytanie całej ścieżki przy realizacji grup rozkazów odwołujących się do kolejnych sektorów.
- T3 Okres impulsów silnika krokowego - uwzględnia bezwładność i czas martwy zespołu silnik - wózek głowic. Zmniejszenie tej wartości powoduje wprawdzie szybszy dostęp do danych, ale niesie w sobie niebezpieczeństwo "połykania" impulsów przez silnik i błędów pozycjonowania.

Powyższe parametry czasowe oraz informacje o wykorzystaniu kanału DMA przekazywane są kontrolerowi rozkazem STOD (Set Type of Disk). Na rozkaz ten składają się trzy bajty. Są one przesyłane przez port danych (3F7h, ewentualnie 377h), naturalnie z zachowaniem każdorazowej kontroli bitu 7 rejestru stanu. Ustawienie tego bitu przez kontroler świadczy o gotowości portu danych.

Bajt 1:

0	0	0	0	0	0	1	1
---	---	---	---	---	---	---	---

*Bity 7-0 liczba 03h, sygnatura rozkazu*

Bajt 2:

bit 7	bit 6	bit 5	bit 4	bit 3	bit 2	bit 1	bit 0
-------	-------	-------	-------	-------	-------	-------	-------

bity 7-4      czas T3 [ms] kodowany w zależności od prędkości transmisji:

	1 Mb/s	500 kb/s	300 kb/s	250 kb/s
00h	8,0	16,0	26,7	32,0
01h	7,5	15,0	25,0	30,0
...	...	...	...	...
0Fh	0,5	1,0	1,7	2

bity 3-0      czas T2 [ms], kodowany w zależności od prędkości transmisji:

	1 Mb/s	500 kb/s	300 kb/s	250 kb/s
00h	128	256	426	512
01h	8	16	26,7	32
02h	16	32	53,5	64
...	...	...	...	...
0Fh	120	240	400	480

Bajt 3:

bit 7	bit 6	bit 5	bit 4	x	x	x	bit 0
-------	-------	-------	-------	---	---	---	-------

bity 7-4      czas T1 [ms] kodowany w zależności od prędkości transmisji:

	1 Mb/s	500 kb/s	300 kb/s	250 kb/s
00h	128	256	426	512
01h	1	2	3,3	4
02h	2	4	6,7	8
...	...	...	...	...
0Eh	126	252	420	504
0Fh	127	254	423	508

bity 3-1      nie używane

bit 0    1 – transmisja kanałem DMA, 0 – transmisja bez udziału kanału DMA

## 2.7. Zabezpieczanie danych – kod CRC.

Nieuniknione niedoskonałości wykonania elementów mechanicznych napędu oraz wady nośnika magnetycznego mogą, mimo prób ich wyeliminowania, powodować przekłamanie utrwalanej informacji. Przyjrzyjmy się więc przez chwilę kontroli poprawności zapisanych danych. Najprostsza, ale i najmniej efektywną metodą kontroli jest tzw. kontrola parzystości. Każdy bajt danych otrzymuje dodatkowy, dziewiąty bit informujący o tym, czy liczba jedynek w bajcie jest parzysta, czy nieparzysta (w zależności od umowy). Łatwo zauważyć, że w najbardziej niesprzyjających warunkach można, bazując jedynie na informacji o parzystości, cały zanegowany bajt (przekłamanie na wszystkich pozycjach) uznać za poprawny. Podobnie ma się sprawa z przekłamaniem każdej parzystej liczby jedynek. W związku z tym opracowano wiele innych, bardziej efektywnych sposobów rozwiązania tego problemu. Magnetyczne systemy cyfrowego zapisu informacji posługują się generalnie tzw. kodami CRC (*Cyclic Redundancy Check*). Dla zobrazowania mechanizmu działania kodu CRC posłużymy się prostym przykładem. Rozpatrzmy na wstępie klasyczne dzielenie dwóch liczb całkowitych. Bez względu na podstawę systemu (dziesiętny, binarny czy inny) obowiązuje zasada:

$$\text{dzielna} / \text{dzielnik} = \text{iloraz} + \text{reszta}$$

Przypomnijmy sobie znany ze szkoły podstawowej algorytm wykonywania takiego dzielenia. Należy ustalić, ile razy mieści się dzielnik w lewostronnej części dzielnej. Wynik ten jest pierwszą cyfrą ilorazu. Cyfrę tę mnożymy przez dzielnik i uzyskaną sumę częściową odejmujemy od wspomnianego lewostronnego fragmentu dzielnej. Do wyniku odejmowania dopisywana jest kolejna cyfra dzielnej i proces się powtarza. Tak samo można postępować z liczbami w systemie binarnym. Nie musimy tu obliczać częściowych ilorazów -wyniki mogą być tylko dwa: dzielna jest większa od dzielnika i częściowy iloraz równa się 1 lub też dzielnik jest większy od dzielnej, a iloraz wynosi zero (dzielenie jest tu zastąpione przez porównywanie).

*Przykład:*

```
324d / 19d = 101000100b / 10011b = 17 d = 10001b, reszta 1
101000100 / 10011 = 10001
10011
000010
00000
00101
00000
01010
00000
10100
10011
00001 <- reszta
```

Do realizacji powtarzających się w tym algorytmie operacji odejmowania używamy zwykłej arytmetyki binarnej, tj. uwzględniamy przeniesienia na sąsiednie pozycje podczas odejmowania liczby większej od mniejszej. Komplikuje to w istotny sposób obliczenia, gdyż wynik na danej pozycji bitowej zależy również od pozycji poprzedniej. Nie jest tak natomiast w tzw. arytmetyce modulo 2. Zastosujmy operację bitowej różnicy symetrycznej (XOR). Oto tablica prawdy dla tej operacji.

X	Y	X XOR y
0	0	0
0	1	1
1	0	1
1	1	0

Tak zdefiniowane dodawanie i odejmowanie daje te same wyniki.

Wróćmy do naszego przykładu, zastępując klasyczne odejmowanie nowo zdefiniowanym działaniem.

$$101000100 / 10011 = 10111$$

10011

001110

000000

11101

10011

011100

10011

011110

10011

01101 <- reszta

Otrzymane wyniki będą się różnić, ale nie należy zapominać, że operujemy inną arytmetyką. Kody kontrolne CRC wykorzystują właśnie reszty z takiego dzielenia. Ilorazy nadają się do tego celu gorzej, gdyż ich długość jest zmienna i w dużym stopniu zależy od długości dzielnej. Reszta musi być krótsza od dzielnika - gdyby była większa, można byłoby przecież podzielić jeszcze raz i zwiększyć iloraz. Pod pojęciem sumy kontrolnej danego ciągu danych w kodzie CRC rozumiemy resztę z dzielenia tego ciągu przez pewien określony dzielnik. Oczywiście jest, że układy wytwarzające sumę kontrolną CRC podczas zapisu posługują się takim dzielnikiem, jaki będzie później użyty przez układy odczytujące do kontroli poprawności danych. Dowolnie długi ciąg danych zabezpieczanych przy użyciu kodu CRC traktuje się jako dzielną. Ustalony z góry dzielnik binarny nazywany jest też generatorem kodu. Reszta z powyższego dzielenia (iloraz jest ignorowany) jest właśnie poszukiwanym kodem CRC, a dokładniej mówiąc -sumą kontrolną w kodzie CRC. Intuicyjnie można przewidzieć, że zdolność do wykrywania błędów przez kod CRC zależy od długości generatora. Dokładna analiza matematyczna, która traktuje powyższe zagadnienia jako element algebry wielomianów, umożliwia wyznaczenie szczególnie opty-



malnych generatorów. Dla celów zabezpieczenia informacji na nośniku magnetycznym stosuje się 17-bitowy generator 10001000000100001, który produkuje 16-bitową resztę. Taki kod CRC (nazywany kodem CCITT) umożliwia rozpoznanie pojedynczego błędnego bitu w ciągach dowolnej długości. Można by w tym momencie stwierdzić, że to samo potrafi bit kontroli parzystości. Kod CRC może jednak dużo więcej. Rozpoznaje on w stu procentach błędy dowolnej pary sąsiadujących bitów. Przekłamanie w szeregu bitów o długości do 16 pozycji wykrywane są w 100%, a przekłamanie w ciągu o dowolnej długości - w 99,9984%. 512-bajtowy blok danych sektora dyskietki zabezpieczany jest przez dopisanie w polu CRC-2 dwóch bajtów zawierających 16-bitowy kod CRC tego bloku. Kontroler, odczytując tak zapisaną informację, korzysta z właściwości arytmetyki modulo 2 -zamiast obliczać sumę kontrolną dla bloku danych i porównywać ją z zawartością pola CRC-2, oblicza sumę kontrolną dla całych 514 bajtów (dane+CRC). Jeżeli nie nastąpiło przekłamanie w polu danych, spodziewany jest wynik równy zero. Tak więc metoda ta jest szybsza i technicznie łatwiejsza.

### **3. DYSKI TWARDE.**

#### **3.1. Historia.**

Historia pamięci masowych sięga połowy dziewiętnastego wieku – już wtedy używano kart perforowanych do wprowadzania danych do mechanicznych maszyn liczących. Pierwsze elektroniczne komputery korzystały z pamięci zbudowanej z lamp elektronowych, potem zaczęły pojawiać się różnej maści pamięci magnetyczne – bąbelkowe, taśmowe, bębnowe.



Pierwszy w historii twardy dysk pojawił się w 1957 roku. Wtedy to IBM zaprezentował urządzenie o nazwie RAMAC 350 – złożony z pięćdziesięciu 24-calowych dysków zespół miał pojemność 5 MB, a koszt jego rocznej dzierżawy wynosił 35 tys. dolarów; jak nietrudno policzyć, oznaczało to 7 tys. dolarów za megabajt... W epoce maszyn mainframe budowano całe „farmy dysków” z zamkniętymi w klimatyzowanych pomieszczeniach zestawami talerzy o średnicach 14 czy 8 cali, wartymi grube dziesiątki tysięcy dolarów. Pojawienie się IBM PC w roku 1981 wcale nie zapowiadało rewolucji w dziedzinie pamięci masowych – system operacyjny „prapoceta” zawierał procedury obsługi pamięci w postaci magnetofonu kasetowego, choć oczywiście istniała także możliwość korzystania ze stacji dyskietek. Lista opcjonalnego wyposażenia IBM PC/XT z roku 1983 obejmuje już twardy dysk o pojemności 5 lub 10 MB – ówczesne napędy o znajomej średnicy 5,25" miały wysokość trzech cali (podobnie zresztą, jak wczesne stacje dyskietek) i stąd właśnie określenie „full height” (współczesny czytnik CD-ROM to „half height”). W roku 1984 Western Digital skonstruował - dzierżący przez kilka lat godność „standardu przemysłowego”, zastosowany w IBM PC/AT interfejs ST506, zaś w 1986 – opracowany do spółki z firmą Compaq dobrze nam znany interfejs IDE (Integrated Drive Electronics). Mniej więcej rok później w komputerach stacjonarnych zaczęto instalować dyski 3,5" (o wysokości 1", czyli „low profile”) – dopiero potem znalazły one zastosowanie w przenośnych laptopach. Postęp technologii powodował ciągły wzrost pojemności i szybkości urządzeń, przy jednoczesnym spadku zapotrzebowania na energię, coraz mniejszej hałaśliwości i większej niezawodności. Wyniki tego wyścigu obserwujemy na co dzień.



W dwudziestoletniej historii PC najbardziej eksponowany był zawsze postęp technologii półprzewodnikowej. Postęp w innych, bezpośrednio z nią związanych dziedzinach technologii był zawsze mniej eksponowany – w ciągu tego samego czasu, gdy stopniowo dokonywano 100-krotnego przyspieszania zegara procesora, pojemność typowego dysku stałego wzrosła 1000-krotnie. Dysk stały, dopóki działa i do momentu, gdy mieszczą się na nim bez kłopotów nasze dane i programy, rzadko bywa przedmiotem szczególnego zainteresowania. Tylko w momencie zakupu staramy się uzyskać możliwy kompromis pomiędzy pojemnościami – dysku i portfela, później dysk schodzi do swojej służebnej roli. Tymczasem od pojemności i szybkości

dysku zależy wydajność komputera i wygoda jego użytkowania, a niezawodność dysku to w wielu przypadkach sprawa nie tylko bardzo ważna, ale wręcz kluczowa.

Przeciętny użytkownik komputera traktuje dysk jako „czarne pudełko” zdolne do zapamiętania pewnej ilości danych – im więcej, tym lepiej. Bardziej dociekliwi zwracają uwagę również na jego parametry wydajnościowe – średni czas dostępu do danych oraz szybkość odczytu i zapisu. Parametry eksploatacyjne każdego urządzenia wynikają z jego konstrukcji, dlatego najdociekliwsi użytkownicy lubią wiedzieć również, co jest wewnątrz „czarnego pudełka” i jak to działa.

### **3.2. Budowa.**

Stosowana w dyskach sztywnych technika zapisu nie odbiega od tej, z jaką mamy do czynienia w magnetofonie czy w kartach telefonicznych, a także w dyskietkach. Wytwarzane przez elektromagnetyczną głowicę pole magnetyczne powoduje uporządkowanie domen magnetycznych w nośniku ferromagnetycznym o szerokiej pętli histerezy, a ruch tak zapisanego nośnika w pobliżu głowicy odczytującej powoduje w niej indukcję sygnału elektrycznego, odpowiadającego zapisanym danym. Współczesna technologia do odczytu danych używa, zamiast głowic indukcyjnych, półprzewodnikowych elementów magnetorezystywnych, umożliwiających zwiększenie zarówno odczytywalnej gęstości zapisu, jak i zwiększenie szybkości odczytu.



Dysk stały naszego PC to wirujący talerz lub zespół talerzy o powierzchni pokrytej nośnikiem magnetycznym, a odpowiednio ustawiane na tych powierzchniach głowice zapisują i odczytują dane. Głowice umieszczone są na przypominającym ramię gramofonu ramieniu pozycjonującym i dociskane do powierzchni dysku sprężynami, ale podczas obrotów dysku nie stykają się z nią – powstająca w wyniku szybkich obrotów talerzy „poduszka powietrzna” utrzymuje głowice nad powierzchnią. Rozwiązanie takie nazywane jest „pływającymi głowicami” i jak na razie jest bezkonkurencyjne i stosowane powszechnie, chociaż są już w toku prace nad innymi sposobami prowadzenia głowic.

Jak już wspomniałem, głowice dysku są zamontowane na konstrukcji obrotowej, budzącej skojarzenie z ramieniem gramofonu. Słuszne optycznie skojarzenie nie jest jednak prawdziwe. Podczas gdy ramię gramofonu było prowadzone przez ścieżkę zapisu na płycie, to z ramieniem głowic dysku jest zupełnie inaczej – musi ono być ustawione tak, by głowice znalazły się nad odczytywaną właśnie ścieżką (czy raczej – na odczytywanym „cylindrze”). W pierwszych konstrukcjach dysków sztywnych pozycjonowanie głowic było realizowane przez mechanizm napędzany silnikiem krokowym (rozwiązanie takie jest do dziś stosowane w napędach dyskietek). W miarę wzrostu wymagań szybkościowych stosowano inne rozwiązania, spośród których optymalnym jak na razie okazało się voice coil, czyli układ magnetodynamiczny, wzorowany na stosowanym w głośnikach (stąd nazwa) – umieszczona w polu silnego magnesu stała cewka porusza się zgodnie z przepływającym przez nią prądem, ustawiając w odpowiedniej pozycji związane z nią mechanicznie ramię głowic dysku. Technika ta pozwoliła na

zmniejszenie czasu pozycjonowania głowic na zadanej ścieżce z kilkudziesięciu do kilku milisekund, a przy przejściach pomiędzy kolejnymi ścieżkami nawet poniżej jednej milisekundy.

Tradycyjnie w komputerze PC AT adresowanie dysku przez przerwanie 13 BIOS-u (INT 13) odbywało się za pomocą trzech parametrów: cylindra, głowicy i sektora (tzw. adresowanie CHS od słów Cylinder, Head, Sector). Konwencjonalne funkcje INT 13 używały 24 bitów do reprezentacji adresów, zatem możliwe było jedynie zaadresowanie obszaru o pojemności 8,4 GB ( $224 \times 512$  bajtów/sektor = 8,4 GB). W celu przekroczenia tej granicznej wartości producenci wprowadzili dwa nowsze sposoby (stosowane właśnie w dzisiejszych dyskach) adresowania.



Pierwszy polegał na rozszerzeniu reprezentacji adresu w konwencji CHS do 32 bitów, drugi – częściej stosowany – używał zupełnie odmiennej metody noszącej nazwę LBA. W metodzie LBA (Logical Block Addressing) stosowane jest adresowanie 28-bitowe, co pozwala na zaadresowanie obszaru do granicznej (znowu!) pojemności wynoszącej:  $228 \times 512$  bajtów/sektor = 137,4 GB. Jest to wartość jak na razie nieosiągalna dla przeciętnego posiadacza komputera (dla producentów – owszem; nie tak dawno Quantum poinformowało o wyprodukowaniu dysku o pojemności ponad 200 GB!). Ten właśnie tryb adresowania jest zalecany i zaimplementowany w BIOS-ach większości dzisiejszych PC-tów.

Zapis na dysku dokonywany jest w formie koncentrycznych ścieżek, podzielonych na sektory. Dość tajemnicze pojęcie „cylinder”, występujące w opisie parametrów dysku i nie znajdujące bezpośredniego odbicia w jego konstrukcji, to grupa ścieżek o tym samym numerze na wszystkich powierzchniach roboczych. Liczba głowic odpowiada liczbie roboczych powierzchni talerzy dysku.

Taki opis parametrów fizycznych dysku oraz wynikający z niego tryb adresowania stanowiły pierwotnie podstawę rozumienia zapisu na dysku przez wszystkie systemy operacyjne. Opis CHS (cylinder/head/sector) sprawdzał się bardzo dobrze w czasach, gdy całością procesu zapisu i odczytu danych zarządzała jednostka centralna przy współdziałaniu dość prymitywnego sterownika. Nietrudno jednak zauważyć, że całkowita długość pierwszej, najbardziej zewnętrznej ścieżki jest znacznie większa od długości ostatniej, najbliższej osi talerza. Liniowa gęstość zapisu jest stała dla wszystkich ścieżek (po prostu – maksymalna), a przy stałej liczbie sektorów na każdej kolejnej ścieżce (licząc od ostatniej do pierwszej) marnowałyby się coraz większa ilość miejsca. Dlatego już od dość dawna stosuje się technikę MZR (Multiple Zone Recording), maksymalnie wykorzystującą dostępną powierzchnię talerzy – liczba sektorów w ostatnim cylindrze dysku, wynikająca z liniowej gęstości zapisu, pozostaje stała w kolejnych cylindrach do chwili, gdy ilość wolnego miejsca pozwoli na dodanie jednego sektora więcej. Na dysku powstają w ten sposób kolejne strefy, charakteryzujące się rosnącą liczbą sektorów w miarę zbliżania się do krawędzi talerza.

W początkowym okresie stosowania MZR praktykowano technikę przeliczania geometrycznej lokalizacji danych na logiczne parametry systemu CHS. Wymagało to dość kłopotliwego, ręcznego wprowadzania parametrów przeliczeniowych konkretnych modeli dysków do pamięci konfiguracji systemu (tzw. Setup). Od problemu indywidualnych parametrów dysków uwolniły nas dopiero: z jednej strony rozwój interfejsu ATA, dzięki któremu system był w stanie samodzielnie odczytać z dysku i przyjąć do wiadomości przeliczeniowe parametry, z drugiej zaś – wprowadzenie BIOS-u funkcji obsługi trybu LBA (Logical Block Addressing), uniezależniającego adresowanie danych na dysku od ich fizycznej lokalizacji na nim.

### **Kodowanie danych.**



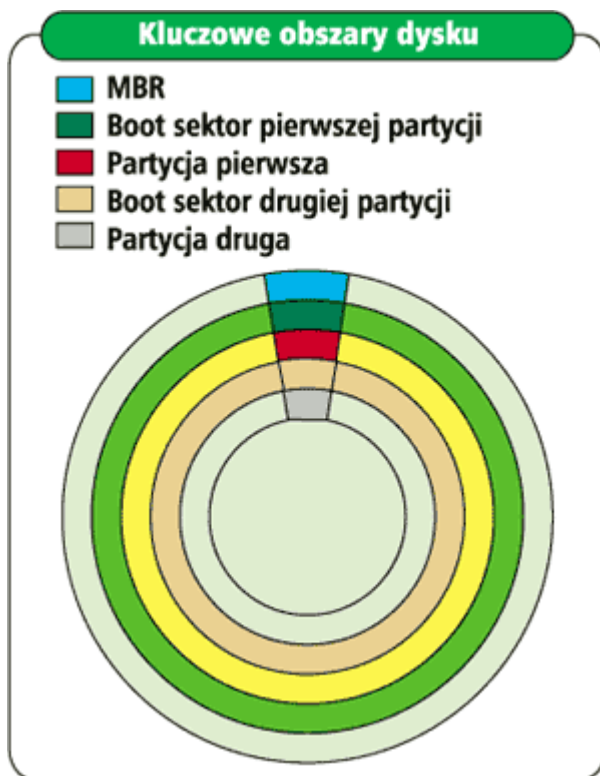
Zapis danych binarnych w formie magnetycznej nie jest dokonywany bezpośrednio „bit w bit” – dane przeznaczone do zapisu są kodowane według pewnych algorytmów, których zadaniem jest usprawnienie odczytu, a także zapewnienie większej jednoznaczności zapisu. Kodowanie danych przeznaczonych do zapisu składa się z dwu faz –

najpierw do zapisywanych danych dodawane są dane nadmiarowe umożliwiające detekcję i korektę ewentualnych błędów odczytu (CRC – Cyclic Redundancy Code – najprostszy, a zarazem jeden z najefektywniejszych algorytmów wprowadzania danych nadmiarowych dla celów korekcji błędów), następnie zaś wynikowe wartości są przekształcane tak, by uniknąć powtarzania dłuższych ciągów powtarzających się zer czy jedynek.

Historycznie pierwszym systemem kodowania danych był MFM, dziś już zupełnie nie stosowany, wyparty następnie przez kodowanie RLL (Run Length Limited) stosowane w dyskach sztywnych do niedawna, a wciąż jeszcze używane przy zapisie na dyskietkach. Obecnie powszechnie stosowaną techniką kodowania danych na dysku jest PRML (Partial Response Maximum Likelihood), która zapewnia największą efektywną gęstość zapisu, a także najniższą stopę błędów odczytu danych. Technika PRML wymaga stosowania w układach sterujących dysku specjalizowanych procesorów o dużej mocy, jednak technologie krzemowe są obecnie na tyle tanie, że uzyskiwane dzięki nim zwiększenie gęstości zapisu z nawiązką wyrównuje nieco wyższy koszt wbudowanej w dysk elektroniki.

### **MZR - Multiple Zone Recording - zapis wielostrefowy**

Nietrudno zauważyć, że w wyniku podziału każdej ścieżki na stałą liczbę sektorów, sektory znajdujące się dalej od osi dysku będą znacznie dłuższe (długość sektorów wewnętrznych jest ograniczona „od dołu” maksymalnym upakowaniem bitów na jednostkę powierzchni). Aby zapobiec ewidentnemu marnotrawstwu, podzielono dysk na kilka stref o określonej liczbie sektorów (od 60 do 120 sektorów na ścieżkę), coraz większej dla stref bliższych obwodowi dysku. Zysk jest ewidentny (o około 25% większa pojemność i wydajność), przy okazji wychodzi na jaw drobne oszustwo: jak to się ma do liczby sektorów na ścieżkę deklarowanej w „Setupie” BIOS? Ano, BIOS mówi swoje, a elektronika dysku po cichu dokonuje przeliczeń...



Rysunek 3.1. Kluczowe obszary dysku.

Mało tego, wewnątrz dysku dzieje się jeszcze coś, o czym ani użytkownik, ani system operacyjny nie mają zielonego pojęcia. Chodzi mianowicie o system obsługi błędów. Oczywiście, dane zapisywane na dysku wyposażone są w dodatkowe informacje umożliwiające funkcjonowanie systemu korekcji „w locie” (ECC on the fly, kodowanie Reed-Solomon itd). Oprócz tego jednak na każdej ścieżce zarezerwowana jest pewna liczba sektorów, które w przypadku pojawienia się fizycznych uszkodzeń nośnika „podstawiane” są przez wewnętrzny mikroprocesor napędu zamiast sektorów wadliwych – dzieje się to całkowicie niezauważalnie dla świata zewnętrznego. Notabene, wewnętrzne układy mikroprocesorowe, w które wyposażone są współczesne napędy, mają moc przetwarzania porównywalną z co najmniej z IBM PC/AT.

### Główki magnetorezystywne.

Twardy dysk ze swoimi maleńkimi elementami wykonanymi z dokładnością przy której, zegarmistrzowska precyzja przypomina raczej kowalską robotę to w istocie arcydzieło technologii. Prawdziwym cudem jest jednak głowica. W nowoczesnych konstrukcjach stosuje się tak zwane głowice magnetorezystywne.

Gwoli ścisłości powinno się raczej używać określenia „hybrydowe” – do zapisu danych służy elektromagnetyczna głowica cienkowarstwowa (jej mikroskopijna ceweczka ma około 10 zwojów), głowica magnetorezystywna służy do odczytu. Wykorzystuje ona efekt zmiany oporności elektrycznej specjalnego materiału (stop żelaza i niklu) przy zmianie pola magnetycznego i jest o wiele czulsza od głowicy elektromagnetycznej. Pozwala to znacznie zmniejszyć powierzchnię zajmowaną przez każdy bit informacji, a więc – zwiększyć gęstość zapisu. Współczesne dyski charakteryzują się gęstością rzędu 1 gigabita na cal kwadratowy, zaś w laboratoriach IBM (to właśnie w nich stworzono pierwsze głowice magnetorezystywne) osiągnięto w grudniu 1996 roku gęstość 5 gigabitów na cal kwadratowy.

Przy tej gęstości na jednym calu długości ścieżki mieści się 240 tysięcy bitów, na jeden cal promienia dysku przypada 21 tysięcy ścieżek, a jeden bit zajmuje powierzchnię 1,2 na 0,1 mikrometra (przekrój ludzkiego włosa zmieściłby około 1000 bitów). Dzięki doskonaleniu technologii GMR (Giant Magnetoresistive Effect) naukowcy przewidują osiągnięcie przed końcem wieku gęstości 10 Gb na cal kwadratowy.

### Pozycjonowanie głowicy.



Kiedyś na potrzeby „nawigacji” zarezerwowana była cała jedna powierzchnia dysku, na której zapisane były znaczniki ścieżek i sektorów dla pozostałych głowic – system taki nazywał się „dedicated servo”. Dzisiejsze napędy wykorzystują technologię „embedded servo” – znaczniki umieszczone są na powierzchniach roboczych i przemieszane z obszarami danych. Wiąże się to co prawda z przydzieleniem elektronice dysku dodatkowych zajęć, pozwala jednak zwiększyć efektywną pojemność urządzenia. W celu uniknięcia błędów odczytu głowica musi znajdować się dokładnie nad środkiem danej ścieżki. Nie jest to wcale łatwe zadanie, gdyż pod wpływem ciepła materiał, z którego wykonane są płyty dysku, może ulec odkształceniom.

W odróżnieniu od tradycyjnej techniki Servo, przy której głowica musiała regularnie korzystać ze ścieżki sterującej, aby zoptymalizować swoją pozycję, mechanizm Embedded Servo wykorzystuje informacje sterujące zapisane na każdej ścieżce. Głowice zapisujące – odczytujące mogą więc korzystać z nich przez cały czas, co umożliwia dokładniejsze pozycjonowanie. Technika Embedded Servo działa na podobnej zasadzie, jak automatyczny pilot, który nieprzerwanie dba o utrzymanie właściwego toru lotu. Stosowana dawniej okresowa kalibracja głowicy dysku powodowała natomiast dodatkowe przerwy w transmisji danych. Inteligentne układy sterujące pozwoliły także zmienić sposób przesuwania głowicy nad szukaną ścieżkę – niegdyś słu-



żyły do tego stosunkowo powolne i zawodne silniczki krokowe (do dziś używane w stacjach dyskietek zdradzają swą obecność charakterystycznym burczeniem), teraz delikatne jak piórko kolibra ramię głowicy wychylane jest na podobieństwo wskazówki miernika elektrycznego za pomocą cewki, przez którą przepływa prąd o odpowiednio dobranym natężeniu (tzw. voice coil) – dzięki temu średni czas

dostępu do danych to dziś 10 lub mniej milisekund. Niektóre firmy stosują technologię „Read on Arrival”, wykorzystującą mechanizm korekcji błędów – pierwsza próba odczytu podejmowana jest jeszcze zanim głowica ustabilizuje się nad żadaną ścieżką; albo próba się powiedzie, albo skutecznie zadziała mechanizm korekcji błędów odczytu, w najgorszym przypadku trzeba będzie ponowić odczyt – nic do stracenia, a można zyskać cenne milisekundy.

### **PRML (Partial Response Maximum Likelihood)**

Większość napędów jeszcze do niedawna podczas odczytu danych używała techniki zwanej peak detection (wykrywanie wartości ekstremalnych – maksimum siły sygnału). W miarę wzrostu gęstości zapisu rozróżnienie sąsiednich wartości szczytowych sygnału od siebie nawzajem i od tzw. tła stawało się coraz trudniejsze. Problem ten rozwiązywano wstawiając pomiędzy sąsiadujące szczyty („jedynki”) rozdzielające chwile ciszy („zera”). Takie postępowanie sprowadzało się do kodowania zerojedynkowych ciągów za pomocą ciągów bardziej przejrzystych, czyli łatwiej identyfikowalnych, lecz z konieczności dłuższych. To oczywiście obniżało efektywną gęstość zapisu danych, a w konsekwencji także wydajność napędu.

Z pomocą przyszła opracowana na potrzeby długodystansowej komunikacji w przestrzeni kosmicznej technologia PRML (Partial Response Maximum Likelihood). Pochodzący z głowicy od-

czytującej analogowy sygnał jest próbkowany w wielu miejscach, a następnie cyfrowo filtrowany przez wbudowany w elektronikę dysku dedykowany procesor sygnałowy DSP. Uzyskaną w ten sposób próbkę analizuje się algorytmem Viterbi. Sprawdza on wszystkie kombinacje danych, które mogły wygenerować zbliżony ciąg i wybiera tę najbardziej prawdopodobną. Umożliwia to dodatkowe zwiększenie czułości kanału odczytu i istotne zmniejszenie prawdopodobieństwa wystąpienia błędów odczytu. Najlepsze efekty daje połączenie technologii PRML z magnetorezystywną głowicą odczytującą ze względu na dobrą jakość generowanego przez nią sygnału analogowego. Głowica magnetorezystywna (MRH) wykorzystuje inne zjawisko fizyczne niż głowice, zbliżone konstrukcją do stosowanych w zwykłych magnetofonach. Element czytający MRH jest wykonany z substancji zmieniającej oporność w polu magnetycznym, więc namagnesowanie nośnika bezpośrednio rzutuje na natężenie płynącego przez głowicę MR prądu. Istotną zaletą technologii MR jest większa czułość, pozwalająca na radykalne zwiększenie gęstości zapisu, a co za tym idzie – wzrost pojemności napędu przy zachowaniu jego rozmiarów. PRML oznacza także inną metodę kodowania danych na dysku: o ile przejście ze starej metody MFM (Multiple Frequency Modulation) na bardziej zaawansowaną RLL (Run Length Limited) oznaczało wzrost upakowania danych o około 50%, PRML daje tu kolejne 20-40% zysku (różne źródła podają różne wartości).

### **Interfejs.**



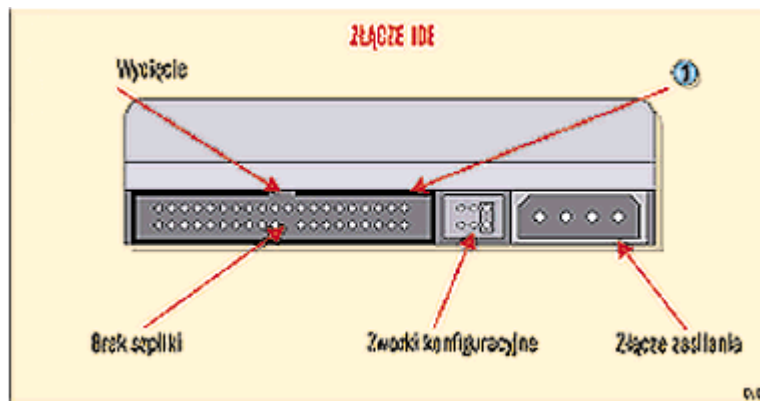
Na pierwszy rzut oka wydawać by się mogło, że w konstrukcji PC używane są obecnie dwa typy interfejsów dyskowych: Ultra ATA i SCSI. Ale po przejrzaniu kilku katalogów sprzętowych opinia ulega zmianie, a w gąszczu nazw takich jak EIDE, ATAPI, Fast Wide SCSI i temu podobnych może się poczuć zagubiony nawet całkiem obyty fachowiec. Tyle

jest standardów – czy nie ma żadnego standardu?

Spróbuję trochę wyprostować kręte ścieżki w dżungli interfejsów dyskowych. Zaczniemy od początku, a na początku, wbrew temu co mówi Pismo, nie było chaosu. W PC był tylko jeden interfejs dyskowy, stworzony przez Seagate i określany jako ST-506/ST-412, od symboli pierwszych dwóch modeli dysków wyprodukowanych przez tę firmę.

Ze względu na niewygodę w posługiwaniu się tak abstrakcyjnym symbolem, żargonowo określano ten interfejs jako MFM lub RLL, od stosowanych w tych dyskach technik kodowania (nb. technika RLL jest stosowana przez niektórych producentów do dziś). Interfejs ten dawno odszedł do lamusa wraz z dyskami o pojemnościach kilkunastu megabajtów, wymagającymi złożonej logiki zewnętrznego sterownika. Jego los podzielił również ESDI (Enhanced Small Device Interface) – pomimo niewątpliwych zalet okazał się zbyt kosztowny w stosunku do swojego konkurenta znanego jako IDE, który stał się pierwszym prawdziwym standardem





Rysunek 3.2. Złącze IDE.

Konstrukcja komputera PC AT stworzyła możliwość zastosowania 16-bitowego interfejsu pomiędzy systemem a dyskiem oraz przeniesienia funkcji sterownika dysku do jego konstrukcji (dyski MFM-RLL były całkowicie „głupie” – całość ich obsługi obciążała wchodzący w skład jednostki centralnej komputera sterownik). Powstał interfejs znany pod nazwami IDE i ATA, który dał początek rodzinie powszechnie stosowanych obecnie interfejsów dyskowych.

Integrated Drive Electronic (IDE) to określenie techniki realizacji nowego interfejsu, w którym całość logiki sterownika dyskowego przeniesiono do konstrukcji dysku, ATA zaś to AT Attachment – relacja nowego interfejsu do konstrukcji AT. Specyfikacja ATA została skodyfikowana przez ANSI jako oficjalny standard, definiując następujący zakres funkcji interfejsu:

- pojedynczy kanał, dzielony przez dwa dyski, skonfigurowane jako master i slave;
- komunikacja w trybach PIO 0, 1 i 2;
- komunikacja przez DMA w trybach 0, 1, 2 dla transmisji pojedynczych słów i w trybie 0 dla transmisji multiword.

Standard ATA sprawdził się bardzo dobrze podbijając rynek, ale rosnące wymagania systemów spowodowały jego rozszerzenie do ATA-2, zaakceptowanego również oficjalnie przez ANSI.

Nowe funkcje interfejsu ATA-2 to przede wszystkim:

- szybsze tryby PIO – ATA-2 wspomaga obsługę w trybach PIO 3 i 4;
- obsługa multiword DMA w trybach 1 i 2;
- rozkazy transmisji blokowych;
- tryb LBA (Logical Block Addressing), umożliwiający, przy odpowiednim wsparciu przez BIOS komputera, przekroczenie systemowych barier pojemności dysku;
- rozszerzenie zakresu identyfikacji dysku przez system.

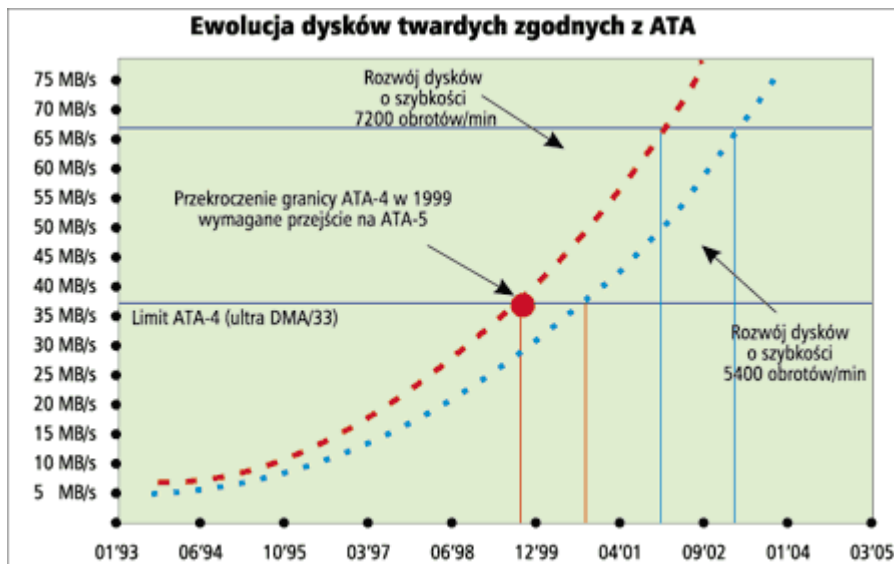
Wprowadzone przez ATA-2 rozszerzenia podniosły sprawność interfejsu, ale w wyniku ciągłego wzrostu mocy PC i wzrostu objętości użytkowanych danych, konieczne stały się kolejne usprawnienia interfejsu tak, by możliwie niewielkim kosztem dało się uzyskać kolejne zwiększenie sprawności. Zanim to jednak nastąpiło, powstało ATA-3, nie wnoszące żadnych nowych trybów transmisji i przyspieszenia obsługi, a jedynie takie, skądinąd bardzo użyteczne zmiany, jak:

- poprawę pewności transmisji danych po kablu, którego możliwości transmisyjne zostały przy uzyskiwanych szybkościach transmisji wyraźnie przekroczone;
- wprowadzenie obsługi SMART (Self-Monitoring Analysis and Reporting Technology) – techniki zapewniającej znaczne zwiększenie poziomu bezpieczeństwa danych dzięki automatycznej diagnostyce dysku.

Następnym znaczącym krokiem w rozwoju interfejsów z rodziny ATA stał się, stosowany obecnie, nieformalny standard Ultra ATA, umożliwiający transmisję danych multiword w trybie 3 DMA, tzw. DMA-33. Oczywiście taki tryb pracy musi być wspierany od strony jednostki centralnej przez BIOS i logikę płyty głównej. Ze strony logiki dysku konieczna jest dodatkowa kontrola poprawności danych i korekcja błędów transmisji z powodu wspomnianego już przekroczenia przepustowości typowego kabla.

Wszystkie interfejsy ATA charakteryzują się pełną kompatybilnością zstępującą, tzn. każda wyższa wersja obsługuje również pełny zestaw funkcji wersji niższej, dzięki czemu możemy np. dołączyć dysk Ultra ATA do systemu przystosowanego do ATA-2, oczywiście tracąc możliwość wykorzystania zwiększających efektywność funkcji wyższego standardu.

Dopełnieniem rodziny interfejsów IDE/ATA jest ATAPI (ATA Packet Interface), protokół komunikacyjny, umożliwiający komunikowanie się przez interfejs ATA z urządzeniami, nie będącymi dyskami stałymi – przede wszystkim z czytnikami CD-ROM. Początkowo protokół ten był obsługiwany przez ładowany do pamięci sterownik, później wbudowany w system operacyjny. Obecnie ATAPI jest wspierany również przez większość implementacji BIOS-u, dzięki czemu możliwe jest np. ładowanie systemu operacyjnego z CD-ROM-u.



Rysunek 3.3. Ewolucja dysków twardech.

**ATA-1** – (IDE) ATA (AT Attachment) jest formalną nazwą tego co często jest nazywane oficjalną specyfikacją IDE. Ale IDE (Integrated Drive Electronics) aktualnie oznacza końcówkę interfejsu twardego dysku. ATA to jedno złącze na płycie głównej umożliwiające podłączenie maksymalnie dwóch urządzeń master i slave o maksymalnej pojemności 528 MB. Wydajność

ATA zależy od trybu transferu danych zwanego PIO (Programed Input/Output). Wszystkie dyski i kontrolery ATA obsługują tryb PIO mode 0 i niektóre obsługują mode 1 i 2. Rodzaj swojego interfejsu można sprawdzić oglądając płytę główną. Płyty z interfejsem ATA mają tylko jedno złącze ATA (nie można podłączyć drugiej taśmy danych oprócz tej do której podpięty jest twarde dyski). Można oczywiście dokupić specjalną kartę ATA-2 mającą dwa kanały ATA i umożliwiającą podłączenie do czterech urządzeń (np. Promise Technology EIDEMax).

**ATA-2** (Enhanced IDE/Fast ATA). Na płycie z tym interfejsem znajdują się dwa gniazda szpilkowe (zwykle do jednego z nich jest podpięta taśma łącząca płytę z twarde dyskiem i napędem CD-ROM). Innym sposobem na sprawdzenie interfejsu bez rozkręcania komputera prowadzi poprzez menu konfiguracyjne BIOS-u (zwykle po włączeniu komputera należy wcisnąć klawisz DEL/ ESC lub CTRL+SHIFT+F1). Wybierz opcję HDD Autodetect. Komputer zacznie rozpoznawać dyski znajdujące się w systemie. Jeżeli na ekranie pojawią się cztery pozycje (nawet puste) to mamy już pewność, że kontroler dysków w komputerze to ATA-2 lub wyższy. ATA-2 akceptuje dyski większe niż 528 MB lecz mniejsze niż 8.1 GB. Standard ten umożliwia podłączenie maksymalnie dwóch urządzeń (2 x master, 2 x slave) do dwóch kanałów kontrolera (primary i secondary). Zawsze należy podłączać szybkie dyski jako master do kanału primary, natomiast wolniejsze (napędy CD-ROM, stare twarde dyski PIO-2 lub PIO-3) jako master i slave do kanału secondary kontrolera. Taki sposób podłączenia jest szczególnie ważny w systemach 486 i wczesnych systemach Pentium, gdyż kanał EIDE primary ma bezpośrednie połączenie z magistralą PCI, natomiast kanał secondary jest połączony z dużo wolniejszą magistralą ISA.

**ATA-3** (Fast ATA). Ta wersja ATA akceptuje dyski pracujące w trybie PIO-4 (znanym także jako „bezprzerwowym”) zapewniającym transfer danych z prędkością 16,7 MB/s.

**ATA-4** (Ultra ATA / Ultra DMA / Ultra DMA-33) Standard podwaja maksymalny transfer trybu PIO-4 do 33 MB/s. Tryb ten zawiera technologię bus mastering używającego kanału DMA w celu zmniejszenia obciążenia procesora.

## **S.C.S.I. (Small Computer Systems Interface)**

Zaawansowany technologicznie typ połączenia komputera z urządzeniami zewnętrznymi takimi jak: streamery, CD-ROM-y, dyski twarde, skanery itp. Standard SCSI umożliwia połączenie w łańcuch do jednego kontrolera 7, a w przypadku wersji rozszerzonej WIDE SCSI nawet do 16 urządzeń (łącznie z kontrolerem). Dla porównania standard EIDE obsługuje tylko 4 urządzenia dzięki czemu SCSI jest szczególnie przydatny w przypadku dysków twardech, gdyż umożliwia jednoczesne połączenie więcej niż 4 takich urządzeń a każdy z nich może mieć pojemność nawet do 50 GB. SCSI oferuje również szybszy transfer danych między urządzeniami, dochodzący do 80 MB/s (EIDE - 66 MB/s).

Standard SCSI jak i wykorzystujące go urządzenia używane są głównie w komputerach Macintosh oraz szybkich serwerach sieciowych i urządzeniach archiwizujących. Rzadziej w domo-

wych pecetach gdyż urządzenia komunikujące się za pomocą tego standardu są zwykle dwukrotnie droższe od takich samych ale wykorzystujących inne standardy jak EIDE czy USB.

Ogólnie SCSI składa się z 4 części: kontrolera SCSI montowanego zazwyczaj jako karta rozszerzająca (ISA lub PCI), lub gotowego elementu wbudowanego bezpośrednio do płyty głównej, kabla połączeniowego SCSI, samego urządzenia (np. dysku twardego) i tzw. terminatorów w postaci zworek lub dodatkowych złącz, które umieszczane są na dwóch końcach łańcucha połączonych urządzeń SCSI. Dzięki nim kontroler otrzymuje informację gdzie są ostatnie urządzenia łańcucha przez co może sprawnie obsługiwać przepływ danych między poszczególnymi urządzeniami a komputerem.

Standard SCSI od czasu swego powstania uległ kilkakrotnie modyfikacjom przez co możemy mieć doczynienia z kilkoma różnymi wersjami tego samego standardu. Jest to szczególnie kłopotliwe w przypadku dobierania i łączenia ze sobą urządzeń obsługujących różne wersje SCSI. Oto zestawienie najważniejszych standardów SCSI:

**SCSI-I** – Leciwy standard wraz z ujednoliconym nieco później Common Command Set (CIS) pracuje na bazie ośmiobitowej magistrali danych i oferuje maksymalną prędkość przesyłania danych około 3 MB/s. Opcjonalny jest synchroniczny tryb pracy i 5 MB/s.

**SCSI-II** – Oferuje jasno zdefiniowany zestaw poleceń i listę parametrów. Dzięki liście urządzeń uniknięto wiele problemów z napędami CD, MO, wymiennymi dyskami, skanerami itp. Magistrala SCSI-II używa 50 żyłowego kabla SCSI-A i zasadniczo nie jest szybsza niż SCSI-I.

**Fast-SCSI** – Jest przyszłością standardu. Pozwala na transfer danych do 10 MB/s, co osiągnięto m.in. podnosząc częstotliwość taktowania magistrali.

**Wide-SCSI** – To 16-bitowa wersja Fast SCSI, Maksymalnie 20 MB /s może być przesyłane synchronicznie za pośrednictwem 68-żyłowego kabla SCSI-B. Dzięki temu, że do adaptera Wide-SCSI na ogół można podłączyć zarówno SCSI-A, jak i SCSI-B, kontroler ten może jednocześnie obsługiwać urządzenia SCSI-II, Fast SCSI oraz Wide - SCSI.

**Ultra-SCSI** – nazywany również Fast-20-SCSI, stosuje większą prędkość taktowania sygnału przesyłanego zwykłym kablem 50-żyłowym i osiąga wydajność 20 MB/s. Dla urządzeń zewnętrznych wymagany jest specjalny kabel połączeniowy. 16-bitowy wariant Ultra-SCSI umożliwia transfer 40 MB/s poprzez kabel SCSI-B nazywany jest Ultra-Wide-SCSI lub Fast-40-SCSI.

Każde z urządzeń SCSI musi mieć przyporządkowany własny adres tzw. ID (od 1 do 15). Dokonuje się tego przełączając odpowiednią zworekę lub pozycję przełącznika (DIP) w urządzeniu SCSI.

### **Zamieszanie w nazwach.**

Gdy oficjalnym standardem został interfejs ATA-2, zaczęły pojawiać się „nazwy handlowe” w rodzaju „Fast ATA” czy „Enhanced IDE”. Miały one znaczenie wyłącznie marketingowe, odnosząc się do sprzętu w jakimś stopniu zgodnego z ATA-2, ale dzięki popularności rynkowej firm używających tych określeń zostały potraktowane i masowo rozumiane jako określenia standar-

dów. Zresztą firmy te, ufne w swoją wielkość i popularność, pozwalały sobie na dość daleko posuniętą elastyczność w stosowaniu się do standardów.

Lansowany w swoim czasie przez Quantum interfejs „Fast ATA-2” był całkowicie zgodny z ATA-2, nie wyłączając szybkości – słowo fast miało charakter wyłącznie marketingowy. Z kolei „Fast ATA” Seagate to nieco ograniczone ATA-2 (bez obsługi trybów PIO 4 i multiword DMA 2). Ale największego zamieszania narobiło określenie EIDE (Enhanced IDE), stosowane przez Western Digital. EIDE to oprócz wszystkich funkcji ATA-2 także ATAPI i obsługa dwóch kanałów, tj. czterech urządzeń EIDE. Ta ostatnia funkcja co prawda nie ma żadnego odzwierciedlenia w konstrukcji interfejsu i w całości jest obsługiwana przez BIOS i logikę płyty głównej, tym niemniej została ujęta w specyfikacji EIDE. W ten sposób każda współczesna płyta główna jest zgodna z tą częścią specyfikacji Western Digital, nawet jeśli konstruktor płyty nigdy nie słyszał ani o tej firmie, ani o Enhanced IDE. Kolejnym źródłem zamieszania stał się, obowiązujący obecnie, standard Ultra ATA. „Ultra DMA-33”, „UDMA” itp. to kolejne marketingowe nazwy, które wprawdzie niekoniecznie odwołują się do standardu, ale dobrze brzmią. Ostatnim marketingowym sloganem jest DMA-66. A jedyną metodą na połapanie się, z jakim dyskiem naprawdę mamy do czynienia, jest odwołanie się do jego specyfikacji technicznej, w której jasno określone są obsługiwane przez jego interfejs tryby PIO i DMA.

Small Computer System Interface (SCSI) to twór bardziej skomplikowany od stosunkowo prostego interfejsu ATA. Podobnie jak ten ostatni, również i urządzenia SCSI mają wbudowane sterowniki, ale podczas gdy ATA jest praktycznie wyłącznie interfejsem pomiędzy dyskiem a jednostką centralną, SCSI stanowi faktyczną szynę systemową, a inteligentne sterowniki urządzeń i bogaty protokół komunikacji sprawiają, że możliwe jest realizowanie przez SCSI wielu funkcji zupełnie niedostępnych dla urządzeń ATA/IDE, jak choćby wymiana danych pomiędzy urządzeniami SCSI bez pośrednictwa jednostki centralnej. W odróżnieniu od ATA, SCSI zostało zaprojektowane od razu z myślą o szerokiej gamie urządzeń, aczkolwiek w początkowej fazie rozwoju również było dość silnie zorientowane na obsługę dysków.

Bałagan nazewnictwa wewnątrz standardu SCSI ma charakter nieco inny niż w przypadku ATA – poszczególne nazwy nie są „hasłami marketingowymi”, lecz raczej opisową charakterystyką poszczególnych wariantów szerokiej rodziny substandardów, np. „Fast Wide SCSI” czy „Ultra SCSI”. Ułatwia to nieco poruszanie się wśród odmian, oczywiście pod warunkiem znajomości ich podstawowych definicji. Na szczęście, mimo nieprecyzyjnego nazewnictwa, poszczególne substandardy są wystarczająco dokładnie zdefiniowane, a ponadto podporządkowane zasadzie kompatybilności zstępującej.

Standard SCSI został zatwierdzony przez ANSI w 1986 roku i bardzo szybko się zestarzał jeśli chodzi o wydajność – szyna 8-bitowej szerokości, taktowana zegarem 5 MHz, mogła stanowić dobry system komunikacji 12 lat temu, ale wymagania systemów komputerowych, nawet tych „small”, wzrastały bardzo szybko. Dlatego już w 1990 roku ANSI zatwierdziło SCSI-2 jako zapis standardu o wariantach dających szeroki zakres możliwości rozwoju.

SCSI w swoich rozlicznych wariantach w stosunkowo niewielkim stopniu znalazł zastosowanie w PC, z dwu głównych przyczyn. Pierwszą z nich była relatywnie wysoka cena urządzeń SCSI i dodatkowy koszt sterownika. Drugą przyczyną ograniczonego stosowania tego interfejsu w komputerach standardu PC był fakt, że zarówno oprogramowanie systemowe (DOS a później Windows), jak i architektura samego komputera nie były przystosowane do wykorzystania możliwości SCSI. Tym niemniej interfejs SCSI bywa stosowany w PC, szczególnie w dwu przypadkach: w maszynach wysokiej klasy, w których koszty nie odgrywają krytycznej roli, natomiast oprogramowanie ma szansę wykorzystania funkcji SCSI, oraz, często w uproszczonej wersji, jako interfejs dla urządzeń takich, jak streamery, skanery, dyski magnetoptyczne i inne urządzenia o dużej objętości danych przesyłanych przez interfejs.

Te najpopularniejsze sterowniki SCSI często mają bardzo ograniczony zakres funkcji – zwykle główne ograniczenie polega na tym, że do interfejsu może być podłączone tylko jedno urządzenie nie zaś, jak wynikałoby ze standardu, aż 7. Sterownik taki, zwykle niewspółmiernie tańszy od "pełnoprawnego", obsługuje podzbiór funkcji SCSI-2 wystarczający do wypełnienia przewidzianej dla niego roli, ale tak naprawdę jest jedynie interfejsem – nie obsługuje zarządzania szyną SCSI.

### **EIDE i S.C.S.I. - podobieństwa i różnice**

Mimo że SCSI powstało nieco wcześniej od swojego konkurenta, jest do chwili obecnej wykorzystywany w dyskach twardych stosowanych w bardziej zaawansowanych systemach komputerowych. Czynnikiem decydującym o wyborze EIDE jest cena. Dyski wykorzystujące ten interfejs są znacznie tańsze, ale czy gorsze? Należy przyjrzeć się bliżej technologii na jakiej bazują obydwie interfejsy.



Rysunek 3.4. Niepożądane efekty zbyt szybkiej pracy dysku.

*Technologia.* Od strony fizycznego zapisu danych na nośniku magnetycznym pomiędzy dyskami SCSI a EIDE nie ma praktycznie żadnej różnicy. Niższa cena dysków EIDE wynika z zastosowanych w nich mechanizmów oferujących nieznacznie mniejsze osiągi przy przesyłaniu danych. Kolejnym, bardzo istotnym czynnikiem kształtującym cenę tych dysków jest masowość ich produkcji. Jedną z podstawowych zasad ekonomii jest to, że przy większej produkcji koszty są niższe. Tak jak EIDE postrzegany jest jako podstawowe rozwiązanie dla biurkowych komputerów PC, tak SCSI uważany jest za standard dla serwerów plików, stacji roboczych i systemów biurkowych klasy wyższej. Historia ewolucji IDE i SCSI pokazuje rosnącą popularność obydwu

standardów. W klasie wyższej rosną możliwości SCSI, natomiast na rynku masowym EIDE zastępuje wcześniejsze generacje SCSI o podobnych pojemnościach i wydajnościach.

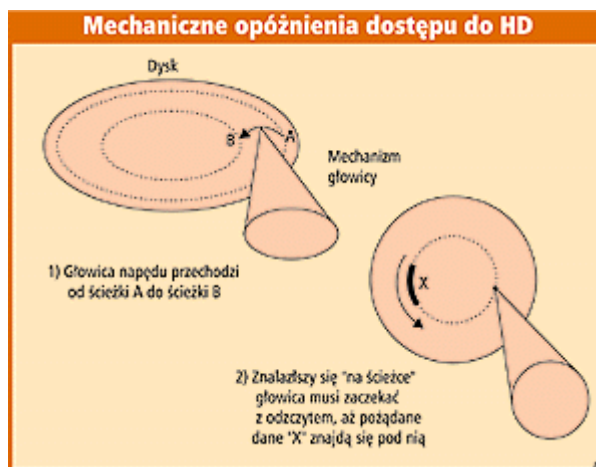
*Dyski coraz szybsze.* Wydajność dysku (zarówno SCSI, jak i EIDE) jest określana głównie przez czynniki mające charakter mechaniczny. Są to: czas pozycjonowania mechanizmu głowicy oraz szybkość obrotowa napędu. Pozycjonowanie określa jak szybko głowica odczytująca dane jest w stanie znaleźć się nad właściwą ścieżką, zaś im większa szybkość obrotowa, tym prędzej dany sektor na tejże ścieżce znajdzie się pod głowicą. Niestety, stałe zwiększanie szybkości obrotowej i zmniejszanie czasu pozycjonowania wymaga zastosowania coraz droższych i wytrzymalszych materiałów.

*Kiedy SCSI jest wydajniejsze.* W systemach komputerowych posiadających jeden dysk twardy, zakup dysku SCSI jest inwestycją nieco chybioną. Podobna sytuacja zachodzi w przypadku starszych systemów operacyjnych (DOS, wczesne wersje Windows), które nie mają wielozadaniowości. Prawdziwy zysk wydajności wynikający z użycia dysków SCSI będzie widoczny w systemach wielozadaniowych zaopatrzonych w wiele nośników pamięci masowej. EIDE to zasadniczo środowisko jednozadaniowe. Nawet przy zastosowaniu zarządzania magistralą (bus mastering) i funkcjach Ultra DMA, system musi poczekać na wykonanie polecenia przez napęd, tak więc przy posiadaniu wielu dysków EIDE podczas komunikacji komputera z jednym z nich inne czekają dopóki zajęty w danej chwili napęd zakończy swoje zadania. SCSI jest interfejsem inteligentnie zarządzającym zasobami. Główną jego zaletą jest ustawianie zadań dla dysków w kolejkę. Dysk SCSI po otrzymaniu od systemu zadania odłącza się od interfejsu wykonując przydzielone mu zlecenia, a w tym czasie pozostałe dyski mogą komunikować się z systemem. Ponadto kolejkowanie zadań pozwala na optymalny odczyt danych z każdego dysku. Polega to na tym, że zadania dla każdego dysku są ustawiane nie w kolejności ich otrzymywania, lecz tak, aby droga głowicy danego dysku odczytująca potrzebne w danym momencie dane była jak najkrótsza, co znacznie skraca czas odczytu i oszczędza mechanikę napędu.

Nie jest to jedyna przewaga SCSI nad EIDE. Interfejs SCSI umożliwia kopiowanie danych z jednego urządzenia SCSI na drugie bez udziału procesora. Interfejs SCSI został zaprojektowany do obsługi urządzeń peryferyjnych w ogóle, a nie tylko i wyłącznie nośników pamięci masowej tak jak to jest w przypadku EIDE. Dlatego też w rozbudowanych systemach komputerowych połączenie za pomocą jednego interfejsu wielu urządzeń np. dysków twardych, skanerów, streamerów potrafi przynieść wymierne korzyści. Wspomniana wyżej np. funkcja SCSI kopiowania (SCSI Copy) z urządzenia na urządzenie pozwala wykorzystywać np. podczas skanowania fotografii (oczywiście skanerem SCSI) moc obliczeniową procesora do innych zadań. Jest to realizowane w znacznym stopniu poprzez jeszcze jedną funkcję SCSI, jaką jest możliwość odłączenia od magistrali aktualnie zajętych urządzeń, a tym samym przyporządkowanie jej wolnym urządzeniom. Należy jednak pamiętać o dość istotnym aspekcie. Użycie tanich kart SCSI powoduje, że dodatkowe rozkazy wchodzące w skład interfejsu SCSI muszą być wykonywane za pośrednictwem procesora i pamięci komputera, co powoduje paradoksalną sytuację –



dobrze skonfigurowana magistrala EIDE może być szybsza od analogicznej SCSI. Dlatego też, gdy najistotniejsza jest wydajność systemu, użycie dysków SCSI o dobrych parametrach nie wystarcza. Konieczne jest uwzględnienie i dopasowanie wszystkich składników magistrali SCSI.



Rysunek 3.5. Mechaniczne opóźnienia dostępu do HD.

### Macierz.

Macierz dyskowa, to zestaw zwykle 3 – 5 dysków widzianych jako jedno urządzenie logiczne, zapewniający redundancję danych wystarczającą do zabezpieczenia ciągłości pracy systemu w wypadku awarii jednego z dysków. Nie jest to może definicja specjalnie ścisła, ale zupełnie wystarczająca w przypadku najczęściej spotykanych macierzy pracujących w systemie RAID3-RAID5 (zbieżność numerów z liczbą dysków jest zupełnie przypadkowa... poziomy RAID – Redundant Array of Inexpensive Disks – odnoszą się do różnych algorytmów zapewniających nadmiarowość danych i związaną z tym odporność na awarie). Gwoli ścisłości, tzw. dyski lustrzane są najprostszą realizacją macierzy w systemie RAID1. Realizacje macierzy w świecie serwerów PC (w komputerach końcowych są stosowane raczej rzadko) można podzielić na trzy rodzaje: sprzętowe, programowe i autonomiczne. W pierwszym przypadku, wszystkie funkcje sterowania macierzą, kontrolę rozmieszczenia poszczególnych fragmentów danych i obliczanie odpowiednich sum kontrolnych realizowane są przez „inteligentny”, dedykowany kontroler dyskowy, który również pamięta szczegóły konfiguracji sprzętowej – np. przyporządkowanie poszczególnych dysków do macierzy (przykładem są np. macierze Mylexa, dostępne również jako OEM u nawet najbardziej markowych producentów sprzętu komputerowego). W przypadku macierzy programowych, cała „inteligencja” podsystemu mieści się w driverze właściwym dla danego systemu operacyjnego, podczas gdy sam sterownik może być niemal dowolny (np. „zwykłe” SCSI-2 Adapteca) – w tym systemie macierze realizuje np. Micropolis. Macierze autonomiczne (zwane niekiedy SCSI-TO-SCSI) są realizowane całkowicie zewnątrz w stosunku do komputera, „inteligencja” jest wbudowana w specjalny moduł sterujący, natomiast całość jest podłączana do zwykłego kontrolera SCSI i widziana jako pojedynczy dysk logiczny.



### **3.3. Parametry.**

Podstawowe parametry dysku sztywnego są takie same, jak dla każdego innego rodzaju pamięci. Jest to pojemność, czas dostępu do danych oraz szybkość ich odczytu i zapisu. Do parametrów podstawowych dochodzi szereg parametrów i informacji szczegółowych, pozwalających na precyzyjne określenie zarówno przeznaczenia dysku, jak i jego zachowania w konkretnych warunkach eksploatacji. Podstawowym „handlowym”, najbardziej eksponowanym parametrem dysku jest, poza pojemnością, czas dostępu do danych. Czas dostępu to czas pozycjonowania głowic plus średnio pół obrotu dysku (bo zawsze jest możliwa sytuacja, gdy poszukiwany sektor znalazł się pod głowicą, zanim była gotowa do odczytu), zatem na czas dostępu wpływa zarówno sprawność mechanizmu pozycjonowania głowic, jak i prędkość obrotowa dysku. Typowy średni czas pozycjonowania głowic to 3-6 ms, a pół obrotu dysku, wykonującego np. 7200 obr./min to około 4 ms – razem, dodając jeszcze czas przełączania głowic, uzyskujemy 8-11 ms. Czas dostępu jest szczególnie ważny w obsłudze baz danych, gdzie występuje często bezpośrednie adresowanie danych plikowych. Ale użytkowo najważniejszym z parametrów dysku jest tzw. transfer wewnętrzny, czyli szybkość bezpośredniego odczytu i zapisu danych. Dlaczego najważniejszym? Rozważmy następujący, bardzo uproszczony, przykład: zadanie polegające na odczytaniu 20 plików po 100 kB każdy w przypadku dwu dysków: jednego o średnim czasie dostępu 10 ms i transferze wewnętrznym 2 MB/s, drugiego o czasie dostępu 7,5 ms (o 25% szybciej!) i transferze wewnętrznym 1,8 MB/s (o zaledwie 10%). Pierwszy z dysków jest w stanie wykonać zadanie w czasie 1200 ms, drugi z nich potrzebuje na to samo zadanie 1261 ms, o przeszło 5% więcej. A ponieważ najbardziej eksponowanym z parametrów jest czas dostępu, ten drugi dysk będzie na pewno sporo droższy, pomimo gorszej wydajności. Transfer wewnętrzny jest stosunkowo rzadko podawany, natomiast dla zwiększenia zamieszania, a także w celu oszołomienia klienta wielkimi liczbami, mocno eksponowanym parametrem dysku jest zwykle szybkość interfejsu, której wpływ na rzeczywistą wydajność systemu jest obecnie marginalny. Odwołując się do naszego przykładu - do przetransmitowania 2 MB danych interfejs SCSI-2 o przepustowości 20 MB/s potrzebuje 100 ms, interfejs Ultra ATA-66 zaledwie 33 ms, ale w obu przypadkach czas transmisji przez interfejs i tak zawarty jest w ogólnym czasie operacji i praktycznie nie ma wpływu na końcowy rezultat.

Skąd wobec tego potrzeba przyspieszania interfejsów, skoro szybkość transmisji wydaje się nie mieć wpływu na ogólną efektywność? Odpowiedź – transmisja przez interfejs zajmuje szynę PCI, przez którą interfejs współpracuje z jednostką centralną, zmniejszając jej dostępność dla innych ewentualnie wykonywanych procesów. W przypadku bardziej zaawansowanych systemów operacyjnych ma to spory wpływ na ogólną wydajność systemu, ale np. w jednozadaniowym środowisku DOS, a także w trybach „wielozadaniowych”, stosowanych w Windows 95/98, nie ma praktycznie prawie żadnego znaczenia. W skrócie:

**Liczba talerzy** – określa liczbę talerzy danego dysku. Uwaga! Liczba talerzy nie oznacza, że dane zapisywane są zawsze po obu stronach talerza dysku. Informację na ten temat otrzymamy porównując liczbę talerzy z liczbą głowic danego dysku.

**Liczba głowic** – określa, ile głowic zajmuje się odczytem/zapisem danych na talerzach. Liczba ta wskazuje także na to, czy wszystkie talerze są wykorzystywane obustronnie. Parzysta liczba głowic wskazuje na to, że dane mogą być przechowywane na każdej stronie każdego talerza dysku, natomiast nieparzysta – że jedna strona któregoś z talerzy dysku nie jest w ogóle wykorzystywana.

**Interfejs** – prawie wszystkie nowe dyski to urządzenia zdolne do pracy w najszybszym obecnie trybie Ultra DMA/66. Jedynie urządzenia ATA-4 nie mają tej funkcji. Mimo to każdy z dysków może pracować także w trybie PIO (przy wyłączonym transferze DMA).

**Średni czas dostępu** – parametr ten określa, w jakim czasie (średnio) od otrzymania przez dysk żądania odczytu/zapisu konkretnego obszaru nastąpi rozpoczęcie operacji. Im krótszy jest ten czas, tym dysk może zapewnić większą płynność odtwarzania, co może mieć znaczenie np. podczas nagrywania płyt CD-R/CD-RW, gdzie wymagany jest ciągły dopływ danych.

**Transfer wewnętrzny** – parametr ten określa w praktyce rzeczywisty transfer danego dysku. Im wartość ta jest wyższa, tym dany dysk jest szybszy. Jednak o tym, czy w danym komputerze będzie osiągał optimum swoich możliwości decyduje konfiguracja komputera (włączenie trybu DMA itp.).

**Transfer zewnętrzny** – właśnie ten parametr często jest używany w marketingowych określeniach i notatkach producentów. Tymczasem nie określa on faktycznej szybkości dysku, lecz przepustowość interfejsu. Oczywiście im ten parametr jest wyższy, tym lepiej – warto jednak pamiętać, że dyski o takim samym transferze zewnętrznym mogą w praktyce pracować z różną szybkością.

**Liczba obrotów na min.** – parametr określający, z jaką szybkością obracają się talerze danego dysku. Im szybkość obrotowa jest wyższa, tym więcej danych może być odczytywanych przez głowice. Pamiętajmy jednak, że ten parametr należy oceniać biorąc pod uwagę także gęstość zapisu. W praktyce jednak przy porównywaniu dysków o podobnej pojemności te z większą szybkością obrotową są zazwyczaj szybsze.

**Cache** – pamięć podręczna dysku twardego. Do tej pamięci buforowane są dane odczytywane i zapisywane na dysku. Im tej pamięci jest więcej, tym sprawniejszy jest proces przesyłu danych.

**MTBF** – akronim od zwrotu Mean Time Between Failure, co można przetłumaczyć jako średni czas międzyuszkodzeniowy. Parametr ten podawany jest w godzinach. Choć wartości, z jakimi spotkamy się w tej tabeli wyglądają na olbrzymie, to należy pamiętać, że czas ten jest wartością średnią ustaloną na podstawie testów dysków danej serii. Warto wiedzieć, że istnieje niezerowe prawdopodobieństwo, że dany dysk ulegnie uszkodzeniu już w pierwszym roku użytkowania.

**Pobór mocy** – zrozumienie tego parametru nie powinno sprawiać kłopotu, jednak jeżeli w naszym komputerze mamy stosunkowo słaby zasilacz, to może się okazać, że zakup dysku wymagającego stosunkowo dużo mocy może spowodować na komputerze, i tak już solidnie obsadzonym różnego typu sprzętem, przeciążenie zasilacza.

#### *Czas dostępu:*

Czas dostępu oznacza ile jednostek (Ms) potrzebuje dysk na pobranie danych. Czas mierzony jest od momentu wydania przez interfejs polecenia odczytu danych. Wartość średniego czasu dostępu jest sumą średniego czasu wyszukiwania (Average Seek Time, czyli czasu potrzebnego na ustawienie głowicy w odpowiednim miejscu nad talerzem twardego dysku) oraz średniego czasu opóźnienia (Average Latency Time – czasu potrzebnego sektorowi do znalezienia się pod odpowiednią głowicą). Przykład: Do komputera do biura zaleca się twardy dysk o średnim czasie dostępu nie dłuższym niż 20 milisekund. Średni czas dostępu idealnego modelu do zastosowań domowych powinien być krótszy niż 16 milisekund, w przypadku komputera multimedialnego – krótszy niż 15 milisekund. Komputer high-end powinien mieć dysk o czasie dostępu 12 milisekund lub krótszym. Producenci w swoich informacjach o produkcie podają zwykle jedynie średni czas wyszukiwania (Average Seek Time) – wartość ta określa wyłącznie czas potrzebny na ustawienie głowicy nad odpowiednią ścieżką dysku. Stąd wartości podawane w opisie dysku są często tak obiecujące. Warto jednak pamiętać, że czas wyszukiwania jest zawsze krótszy od czasu dostępu, z którym nie należy go mylić. Jeśli do czasu wyszukiwania dodamy czas opóźnienia (obliczając w ten sposób czas dostępu), uzyskamy zupełnie inny wynik. Na przykład: czas wyszukiwania dla dysku Diamondmax 90432D2 firmy Maxtor wynosi 9,0 milisekund, a czas opóźnienia, według producenta, 5,5 milisekundy. Po dodaniu okaże się, że czas dostępu wynosi nie - jak mogliśmy mylnie sądzić - 9,0 milisekund, lecz 14,5 milisekund. Wartość Track-to-Track nie mówi niczego o twardym dysku. Miara ta podaje czas potrzebny na przesunięcie głowicy zapisująco – odczytującej nad sąsiednią ścieżkę.

#### *Wewnętrzna szybkość przesyłania danych:*

Informuje o ilości danych, jaka może zostać na twardym dysku zapisana lub odczytana z dysku w ciągu sekundy. Jest to wartość teoretyczna, niezależna od systemu operacyjnego oraz wszelkich właściwości komputera, określająca wydajność wewnętrzną dysku. Mierzona jest ona dla sektorów znajdujących się między krawędzią płyty, a jej środkiem, jako że w zewnętrznej części płyty znajduje się więcej sektorów aniżeli w części bliższej środka. Jeśli wartość ta nie jest podana w danych technicznych o twardym dysku, można ją obliczyć: liczbę obrotów na sekundę należy pomnożyć przez liczbę sektorów znajdujących się na najbardziej zewnętrznej ścieżce i tak uzyskana wartość pomnożyć przez 512, jako że tyle bajtów zawiera jeden sektor. Wartości tej nie da się zmierzyć za pomocą testów, gdyż dokonanie pomiarów możliwe jest tylko przy użyciu odpowiednich urządzeń. Przykład: Twardy dysk do komputera do biura i domu powinien charakteryzować się wewnętrzną szybkością przesyłania danych większą od 140 Mb/s. Komputer multimedialny najlepiej wyposażyć w dysk pozwalający na przesyłanie 160

Mb/s, a komputer high-end - w dysk o szybkości 180 lub więcej Mb/s. Niektórzy producenci podczas obliczania wewnętrznej szybkości przesyłania danych, jako dane liczą również dane administracyjne (takie jak numer sektora, nad który głowica ma być przesunięta i suma kontrolna) uzyskując dzięki temu wartość o 12 do 20 procent wyższą. Stąd pochodzić może ewentualna różnica między wartością wewnętrzną szybkości przesyłania danych obliczona samodzielnie, a podana przez producenta.

#### *Szybkość przesyłania danych:*

Średnia szybkość przesyłania danych oznacza ilość danych przesyłanych w ciągu sekundy przez twarde dysk. Wartość średnia jest ustalana podczas co najmniej dwóch testów, w jednym z nich odczytywane są wszystkie ścieżki twardego dysku po kolei, a w drugim – według przypadkowej kolejności. Szybkość przesyłania danych zależy również od innych charakterystyk dysku, takich jak wewnętrzna szybkość przesyłania danych, szybkość przesyłania danych przez interfejs oraz średni czas dostępu. Przykład: Do komputera do biura zalecamy twarde dysk o średniej szybkości przesyłania danych wynoszącej 5 MB/s. Twarde dyski w komputerach do zadań domowych oraz multimedialnych powinny umożliwiać przesyłanie danych z prędkością 6 MB/s, jako że będą one zapewne przesyłać więcej obrazów oraz sekwencji wideo. Komputer high-end może osiągać najwyższą wydajność tylko wówczas, jeśli wyposażymy go w superszybki twarde dysk o szybkości przekraczającej 7, a nawet 10 MB/s. Dysk powinien przesyłać dane z szybkością przynajmniej 5 MB/s. Producenci, zamiast średniej szybkości przesyłania danych, podają często maksymalną szybkość interfejsu, czyli „Burst-Rate”. Opisy mogą brzmieć: „Interface Transfer Rate” (do dysków firmy Hitachi), „Data Transfer Rate to/from Interface” (do dysków firmy Maxtor), czy też „External Transfer Rate” (do dysków firmy Seagate) mają wartości „up to 33 MB/s”. Aby sugerowana szybkość mogła być przez określony dysk osiągnięta, musi znajdować się również w jego pamięci podręcznej. Tylko wówczas możemy być pewni, że dysk przesyła dane z szybkością określoną przez producenta. Ze uwagi na to, że powyższy warunek jest rzadko spełniany, miara „Burst-Rate” jest wartością, którą nie powinniśmy się sugerować podczas wybierania optymalnego dla nas twardego dysku.

#### *Szybkość obrotowa:*

Szybkość obrotowa dysku twardego to parametr często wykorzystywany przez producentów i dystrybutorów w celach marketingowych. Tymczasem pomyślmy, jakie znaczenie ma on dla użytkownika? Sama szybkość obrotowa dysku jeszcze o niczym nie świadczy. Dopiero w połączeniu z gęstością zapisu możemy mówić o użytkowych walorach tego parametru dysku. Zasada jest stosunkowo prosta. Jeżeli połączymy największą szybkość obrotową z najwyższą gęstością zapisu, otrzymamy najszybszy dysk – a dokładniej: dysk o największym transferze wewnętrznym. O tym, czy dysk taki będzie faktycznie najszybszy w naszym komputerze decydują takie czynniki, jak interfejs pomiędzy dyskiem a płytą oraz przepustowość kontrolera. Warto jednak wiedzieć, że różnice pomiędzy gęstością zapisu dla współczesnych dysków o porównywalnej do siebie pojemności są na tyle małe, że można w pewnym sensie stwierdzić, iż

szybkość obrotowa decyduje o szybkości transferu. Coraz większa pojemność i szybkość obrotowa dysków twardech to jedna strona medalu. Z drugiej mamy przecież do czynienia z konkretnym oprogramowaniem komputera, na którym dany dysk ma pracować. Tutaj trzeba wyraźnie uświadomić sobie możliwe ograniczenia, jakie narzuca używany przez nas system operacyjny czy oprogramowanie BIOS-u komputera.

*MTBF - średni czas międzyuszkodzeniowy:*

Średni czas międzyuszkodzeniowy (Mean Time Between Failures) to podstawowy parametr niezawodności wszystkich urządzeń. Oferowany przez większość współczesnych dysków współczynnik MTBF wynoszący 500 000 godzin lub więcej imponuje wielkością, zwłaszcza gdy po przeliczeniu dowiemy się, że pół miliona godzin to około 60 lat ciągłej pracy. Ale czy to naprawdę wysoka niezawodność? Pamiętajmy, że jest to czas „średni”, określony na podstawie analizy statystycznej wyników testów, a dotyczy całej serii danego modelu dysku. Z punktu widzenia indywidualnego użytkownika ze współczynnika MTBF równego 500 000 godzin wynika aż 1,7% prawdopodobieństwa, że w bieżącym roku dysk ulegnie uszkodzeniu. Przy normalnym użytkowaniu komputer pracuje średnio 6 godzin w ciągu doby, zatem prawdopodobieństwo uszkodzenia będzie odpowiednio mniejsze – rzędu 0,5%. Ujmując to w skali masowej, a nie tylko pojedynczego egzemplarza dysku – co dwusetny dysk ulegnie w tym roku uszkodzeniu! Szansa uszkodzenia dysku jest o wiele rzędów wielkości większa od szansy wygranej w Lotto, a jednak spośród grających regularnie w Lotto użytkowników komputerów tylko nieliczni zwracają sobie głowę czymś tak surrealistycznym jak backup danych. Wiara w szczęście czy fascynacja wielkimi liczbami? Dla przeciętnego domowego użytkownika takie prawdopodobieństwo awarii zwykle jest „do przyjęcia”, zwłaszcza jeśli w kieszeni ma trzyletnią gwarancję, a na dysku nie przechowuje unikalnych wartościowych danych. Ostatnio coraz więcej producentów podaje dla niektórych modeli dysków współczynnik MTBF rzędu miliona godzin i więcej. Liczby astronomiczne, ale zapewniające tylko dwukrotne zwiększenie szans użytkownika. Z możliwością uszkodzenia dysku należy się poważnie liczyć, jeśli używamy komputera nie tylko do zabawy, a na dysku przechowujemy wartościowe dane. A nawet w przypadku komputera „rozrywkowego” awaria dysku to poważny kłopot i wiele godzin pracy nad odtworzeniem konfiguracji aplikacji.

*Pojemność - jaka naprawdę?*

Przeglądając oferty lub informacje dystrybutorów i producentów dysków twardech niejednokrotnie dokonujemy wyboru na podstawie parametrów, jakie przedstawia dana specyfikacja. Tymczasem w przypadku pojemności informacja podawana na ulotce nie do końca musi odpowiadać temu, co zobaczymy po sformatowaniu dysku w naszym komputerze. Po pierwsze, dość często spotykanym „wybiegiem” marketingowym jest podawanie pojemności danego dysku w mega- lub w gigabajtach, z zastrzeżeniem, że 1 MB to 1 000 000 bajtów, a 1 GB to 1 000 000 000 bajtów. Tymczasem stan faktyczny jest inny – 1 kB równy jest 1024 bajtom, a nie 1000 bajtom. Różnica nie jest co prawda wielka, ale przy olbrzymich pojemnościach dzisiejszych

dysków te zaokrąglenia powodują, że różnica pomiędzy informacją producenta a wynikiem formatowania dysku w komputerze może okazać się zaskakująca dla nieświadomego takiej polityki użytkownika. Przykładowo dla dysku o pojemności (przy przeliczniku 1 kB = 1000 B) 18 042 MB otrzymamy, że dysk dysponuje faktyczną pojemnością ok. 17206,20 MB. Jak więc widać różnica sięga ponad 800 MB, co jeszcze nie tak dawno stanowiło całkowitą pojemność dysku twardego! Dlatego też dokonując wyboru musimy pamiętać o tym, w jaki sposób megabajty czy gigabajty są podawane w informacjach producenta.

### **Oznaczenia producenta:**

Zawierają wiele informacji dotyczących twardego dysku. Po dokładnym przemyśleniu wszystkich za i przeciw zdecydowaliśmy się na zakup określonego modelu dysku – może być tak, że modele oferowane przez sprzedawcę nie mają nazw typu Fireball, Cheetah, czy Hornet, lecz nazwy kodowane. Oto przykłady, w jaki sposób należy je odczytywać

#### *Fujitsu - przykład MPB3021*

- Pierwszy znak alfanumeryczny określa rodzaj głowicy odczytująco – zapisującej: „M” – oznacza głowice MR (ang. Magnetoresistive) „GM” – oznacza głowice GMR (ang. giant magnetoresistive).
- Litera druga oznacza rodzaj dysku: „A” – Allegro (czyli dysk SCSI), „H” – Hornet (czyli 2,5 calowy dysk do notebooków), „P” – Picobird (czyli dysk EIDE Ultra-DMA/33).
- Ostatnia z liter określa generacje twardego dysku. Modele nowsze – wyprodukowane po październiku 1998 oznaczone są litera „C”.
- Pierwsza cyfra określa format dysku, gdzie „2” oznacza dysk 2,5 calowy, a „3” – dysk 3,5 calowy.
- Następane trzy cyfry tworzą liczbę, która po podzieleniu przez 10 wskazuje na pojemność dysku mierzona w GB.



#### *Hitachi - przykład DK238A-32*

- Pierwsze dwa znaki („DK”) oznaczają „dysk”. Pierwsza cyfra określa format dysku, gdzie „2” oznacza dysk 2,5-calowy (który Hitachi produkuje wyłącznie z interfejsem EIDE), a „3” – dysk 3,5-calowy (z interfejsem SCSI).
- Druga cyfra oznacza wysokość dysku (dla dysków 2,5 calowych cyfry „1”, „2” i „3” oznaczają odpowiednio: 19,0; 12,7 i 9,5 milimetra; a dla dysków 3,5 calowych „1” oznacza 41mm, a „2” – 25 mm wysokości.)
- Trzecia cyfra określa generacje twardego dysku (obecnie aktualnymi są generacje 7 i 8).
- Litera poprzedzająca myślnik jest oznaczeniem wewnętrznym Hitachi.
- Liczba po myślniku podaje pojemność dysku – w przypadku dysków o pojemności mniejszej niż 10 GB liczbę ta należy podzielić przez 10.



#### IBM - przykład DCAS-34330

- Litera „D” oznacza „dysk”.
- Następne dwie litery, to oznaczenia wewnętrzne IBM.
- Litera poprzedzająca myślnik informuje o rodzaju interfejsu: „A” oznacza AT-Bus (EIDE), a „S” – SCSI.
- Pierwsza cyfra określa format dysku, gdzie „1” oznacza dysk 1-calowy, „2” – dysk 2,5-calowy, a „3” – dysk 3,5-calowy.
- Pozostałe cztery cyfry podają pojemność dysku w MB. Jeśli po myślniku znajduje się nie 5, lecz 6 cyfr, to druga z nich podaje liczbę tysięcy obrotów dysku na minutę. Ostatnie 4 cyfry określają pojemność dysku w MB.



#### Maxtor - przykład 91360D8

- Pierwsza cyfra określa 9 (obecnie aktualną) generację twardego dysku.
- Kolejne cztery cyfry po przemnożeniu przez 10 określają pojemność dysku w MB. Litera określa rodzaj protokołu interfejsu, gdzie „D” oznacza interfejs Ultra DMA/33, a „U” – Ultra DMA/66.
- Ostatnia cyfra określa liczbę głowic odczytujących – zapisujących.
- Informacja dodatkowa: dyski Diamond wykonują 5400 obrotów na minutę, a dyski Diamondmax – 7200 RPM.



#### Quantum - przykład QM36480ST-A

- „QM” to skrót od nazwy firmy (Quantum).
- Pierwsza cyfra określa format dysku, gdzie „3” oznacza dysk 3,5-calowy, a „5” – dysk 5,25-calowy.
- Następne cyfry określają pojemność dysku w MB.
- Litery następujące po cyfrach to wewnętrzne oznaczenia firmy Quantum, gdzie „AL.” oznacza Atlas II, „CY” – Bigfoot CY, „EL” – Fireball EL, „PX” – Viking II, „SE” – Fireball SE, „SG” – Pioneer SG, „ST” – Fireball ST, „TD” – Atlas III, „TX” – Bigfoot TX, a „VK” oznacza Viking.
- Ostatnia litera lub litery określają rodzaj interfejsu: „A” oznacza AT-Bus (EIDE), „LW” oznacza Ultra-II-Wide-SCSI, „S” oznacza SCSI, a „SW” – Wide-SCSI.

Sprzedawcy czasem podają nazwy skrócone, takie jak: Quantum Fireball EL 2500. Tego rodzaju nazwa określa jedynie pojemność twardego dysku i jego nazwę.

#### Seagate - przykład ST39140A

- „ST” określa producenta dysku, czyli firmę Seagate.
- Pierwsza cyfra określa wysokość dysku i jego format: „1” określa dysk o wysokości 41mm (3,5 cala), „3” – 25 mm (3,5 cala), „4” – 82 mm (5,25 cala), „5” – 19 mm (3,5 cala) i „9” na określenie dysków 2,5-calowych o różnych grubościach.
- Kolejne cyfry informują o pojemności dysku mierzoną w MB.



- Jedna lub dwie litery na końcu nazwy określają interfejs: „A” oznacza AT-Bus (EIDE), „N” oznacza SCSI, „W” – Wide SCSI, a „FC” – Fiber Channel.

*Western Digital - przykład WDAC36400L i WDE9100-0007A1*

1.

- Pierwsza cyfra informuje z ilu talerzy składa się dysk.
- Następne określają jego pojemność w MB.
- A litera na końcu oznacza pojemność pamięci podręcznej, gdzie H = 128 KB, a L = 256 KB.



2.

- Pierwsza seria cyfr podaje rozmiar dysku w MB.
- Pierwsze dwie cyfry po myślniku pozostają do dyspozycji klientów OEM, czyli firm składających komputery.
- Kolejne cyfry określają standard SCSI oraz rozmiar pamięci podręcznej: „03” oznacza interfejs SCSI z 512 KB pamięci podręcznej, „07” – Wide-SCSI z 512 KB pamięci podręcznej, „16” – Wide-SCSI z 1 MB pamięci podręcznej, a „30” – Ultra-Wide-SCSI z 4MB pamięci podręcznej.
- Ewentualne kolejne znaki alfanumeryczne w przykładzie „A1” są oznaczeniami OEM.

### **3.4. Montaż.**

Najpopularniejszym interfejsem do przesyłu danych w komputerach PC jest obecnie EIDE. W dzisiejszych komputerach kontroler EIDE jest zintegrowany z płytą główną, w starszych jednostkach korzystanie z dysków wymagało zainstalowania w komputerze dodatkowej karty kontrolera. Jeżeli na płycie głównej komputera znajdują się złącza oznaczone zazwyczaj jako „Primary” i „Secondary”, znaczy to, że komputer ma kontroler EIDE zintegrowany z płytą. Oznaczenia te odzwierciedlają dwa odrębne kanały (pierwszy i drugi lub też nadrzędny i podrzędny) interfejsu EIDE. Pozornie instalacja dysku jest prosta, gdyż w zasadzie polega tylko na połączeniu zamontowanego w obudowie dysku z jednym z dwóch wyżej przedstawionych złączy kontrolera za pomocą specjalnego kabla zwanego popularnie „taśmą”. Teoretycznie, gdy dysponujemy tylko jednym dyskiem, komputer będzie z nim współpracował bez względu na fakt, do jakiego kanału został podłączony. Tym niemniej pamiętać należy o generalnej zasadzie, by najszybszy dysk podłączyć zawsze do pierwszego kanału oraz by (w przypadku, gdy w komputerze zainstalowany jest jeszcze jakiś dysk lub inne urządzenie EIDE) był on urządzeniem nadrzędnym „master”. Nawet jeżeli wymaga to przetasowania dysków (czyli przesunięcia starszego napędu z pozycji „master” na pierwszym kanale na pozycję „slave”), to dzięki temu znacznie lepiej wykorzystamy potencjał technologiczny tkwiący w nowym dysku. Kontroler EIDE łączy się z urządzeniami w kolejności kanałów, uwzględniając priorytety urządzeń, zatem wstawienie najszybszego dysku na uprzywilejowanej pozycji jest rozwiązaniem optymalnym. Wynika to z faktu, że najczęściej na pierwszym dysku zainstalowany jest system operacyjny, a tym samym jest on najczęściej wykorzystywany. Kontroler EIDE jest w stanie obsłużyć maksymal-



nie do czterech urządzeń (dysków, napędów CD-ROM czy DVD-ROM). Na każdy kanał przypadają po dwa urządzenia. Na każdym kanale dane urządzenie musi mieć ustawiony priorytet dostępu do magistrali. Innymi słowy, z dwóch urządzeń podłączonych do jednego kanału EIDE, jedno musi być urządzeniem nadrzędnym, zaś drugie podrzędnym. Każdy dysk twardy EIDE ma zworki pozwalające ustawić go na pożądany priorytet. Oznaczenie dysku jako „master” (władca) określa go jako urządzenie nadrzędne na danym kanale, odpowiednikiem urządzenia podrzędnego jest określenie „slave” (niewolnik). Jeżeli dysk twardy jest jedynym tego typu urządzeniem w systemie, nie ma potrzeby ustawiania dla niego priorytetu, a tym samym nie trzeba manipulować zworkami na dysku.

### **Konfiguracja BIOS-u**

Pierwszym etapem programowej instalacji dysku po jego fizycznym zamontowaniu jest ustawienie odpowiednich parametrów w BIOS-ie. Przeprowadza się to wykorzystując wbudowany w BIOS program Setup uruchamiany najczęściej klawiszem „Del”. Jeżeli komputer, w którym znajduje się dysk poddawany operacji nie jest starszy niż trzyletni, to prawie na pewno w programie Setup znajduje się opcja „HDD Auto Detection” lub podobna. Jej wywołanie spowoduje przeszukanie poszczególnych kanałów EIDE i w razie odnalezienia jakiegokolwiek urządzenia (tu: dysku twardego) – automatyczne wprowadzenie jego parametrów. Gdy dane dysku zostały już poprawnie rozpoznane, nie wolno zapomnieć o zapisaniu zmian podczas opuszczania programu Setup. Tylko w takim przypadku nowo wprowadzone zmiany zostaną zapisane. Ponadto już od jakiegoś czasu BIOS komputerów wyposażane są w funkcje automatycznego konfigurowania i detekcji dysków twardech w komputerze z pominięciem wyżej wspomnianej opcji. Np. w BIOS-ie Awarda wystarczy w tabeli urządzeń EIDE/ATAPI wyświetlanej w sekcji „Standard CMOS Setup” wprowadzić w kolumnach „TYPE” oraz „MODE” wartości „Auto” dla każdego kanału EIDE. Dzięki temu nie tylko nie musimy znać specyficznych parametrów dysku, ale też w razie dołożenia do naszego zestawu komputerowego nowego dysku, BIOS automatycznie go wykryje.

### **Dzielenie dysku na partycje**

Kolejny etap instalacji nowego dysku to podzielenie go na partycje oraz ich sformatowanie. Bez względu na to, czy używany jest system DOS, Windows 3.1, 95 czy 98, program wykonujący to zadanie nazywa się tak samo: fdisk. Mimo że interfejs programu w każdym z wymienionych systemów wygląda praktycznie identycznie, to funkcjonalnie między poszczególnymi wersjami programu są pewne istotne różnice, wśród których najbardziej znamienne jest typ systemów plików, jakie przez daną wersję fdisk są obsługiwane. W przypadku DOS, Windows 3.1 czy Windows 95A (a także OSR1) fdisk obsługuje w zasadzie tylko system FAT (zwany też FAT16). Nowsze (Windows 95 OSR2 oraz Windows 98) wyposażono już w nowszy system FAT32. Bez względu na to, ile partycji chcemy utworzyć na danym dysku, uruchomienie fdiska jest konieczne. Aby na danym dysku mógł zostać zainstalowany system operacyjny, dysk ten musi mieć przynajmniej jedną partycję ustawioną jako aktywną. Ustawianie partycji na dysku za

pomocą programu fdisk nie jest rzeczą trudną. Interfejs programu ogranicza użytkownika w zasadzie tylko do wybierania odpowiednich opcji z wyświetlanego na ekranie menu. Nie trzeba wprowadzać żadnych skomplikowanych komend itp. Jednak trzeba pamiętać o kilku rzeczach, co być może uchroni nas od zmiany układu partycji na nie tym dysku, co trzeba, a co za tym idzie – od straty danych. Jeżeli dysk, który chcemy podzielić na partycje, jest dyskiem podrzędnym, to najprawdopodobniej mamy możliwość uruchomienia systemu operacyjnego z dysku dotychczas używanego. Zatem w tym systemie znajduje się również potrzebny nam program. W przypadku DOS czy Windows 3.1x fdisk znajduje się w katalogu DOS-u (najczęściej C:\DOS\), zaś w przypadku nowszych Windows 95 czy 98 – fdisk znajduje się w (najczęściej) c:\windows\command\. Jednak bez względu na system, wystarczy zazwyczaj z linii poleceń wpisać komendę „fdisk”, by program się uruchomił. Użytkownicy Windows 95/98 mogą to zrobić wywołując z menu Start opcję „Uruchom” i tam wpisując komendę „fdisk”. Lecz w przypadku, gdy nowy dysk twardy jest jedynym dyskiem zainstalowanym w komputerze, pozostaje wykorzystanie dyskietki startowej, dołączanej do każdego z ww. systemów, na której znajduje się fdisk. Wielu użytkowników, a także producentów gotowych zestawów poprzestaje na założeniu tylko jednej partycji bez względu na rozmiary dysku. Mając do dyspozycji kilka GB można oczywiście wszystko umieścić na jednej partycji, ale w przypadku, gdy system odmówi posłuszeństwa lub też jego stan będzie na tyle daleki od doskonałości, że będzie konieczne przeprowadzenie ponownej instalacji systemu albo – co gorsza – formatowania dysku, nie będzie gdzie przechować na czas operacji ważnych dokumentów bądź aplikacji. Dlatego też, w zależności od rodzaju wykonywanej pracy, najlepiej utworzyć na dysku dwie, trzy partycje. Przy czym ich rozmiary zależą od całkowitej pojemności dysku, a także od typu dokumentów, jakie będą na dysku przechowywane. Pamiętać przygotowywanej partycji bądź dysku. Jeżeli zamierzamy przechowywać dużą liczbę małych plików (np. plików tekstowych, dokumentów HTML, VRML, małych plików graficznych, itp.), lepiej skorzystać z systemu FAT32. Z drugiej strony, jeżeli zajmujemy się obróbką dużych plików (np. cyfrowego dźwięku wysokiej jakości czy wideo klipów) użycie FAT32 nie spowoduje wielkich oszczędności przestrzeni dysku. Ponadto jeżeli ciągle jeszcze wykorzystujemy oprogramowanie DOS-owe, również lepiej skorzystać z FAT16, gdyż FAT32 w DOS-ie, a także w trybie awaryjnym Windows 95OSR2 i 98 działa znacznie wolniej. Warto też wiedzieć, że Konwerter dysku nie potrafi dokonać konwersji partycji FAT16 na FAT32, jeżeli ma ona pojemność mniejszą niż graniczne 512 MB. Kolejnym aspektem zakładania partycji na dysku jest ustalenie, z jakiego systemu plików dana partycja ma korzystać. W przypadku DOS, Windows 3.1 oraz pierwszej wersji Windows 95 (nie OSR2) nie ma wielkiego wyboru (tylko FAT16). Gdy używamy nowszych "okienek", program fdisk bezpośrednio po uruchomieniu pyta użytkownika, czy włączyć obsługę dużych dysków. Odpowiedź twierdząca oznacza, że ustawiane partycje będą wykorzystywały nowszy system plików: FAT32. Dotyczy to tylko partycji o objętości większej niż 512 MB, partycje mniejsze będą ustawione na system plików FAT16.

## **Formatowanie**

Samo przydzielenie logicznie odrębnych obszarów na dysku poszczególnym partycjom nie wystarczy, aby dysk taki był „widziany” przez system operacyjny. Konieczne jest jeszcze wykonanie operacji formatowania każdej utworzonej partycji. Formatowanie wykonujemy za pomocą polecenia „format”. Formatowanie wykonywane przez program „format” (bądź przez jego windowsowy odpowiednik) jest formatowaniem logicznym. W odróżnieniu od formatowania fizycznego, na dysk nie są nanoszone żadne ścieżki ani sektory, gdyż te już się tam znajdują po formatowaniu fizycznym. Każdy współczesny dysk EIDE jest formatowany fizycznie jeszcze u producenta i zazwyczaj nie ma potrzeby, a nawet nie wolno ponawiać tego działania. W starszych komputerach program umożliwiający fizyczne sformatowanie dysku twardego znajdował się w BIOS-ie komputera, ale w najnowszych BIOS-ach w programie Setup nie ma już opcji pozwalającej wykonać takie działanie. Jest to zresztą jak najbardziej słuszne. Formatowanie fizyczne dysku EIDE może go bezpowrotnie uszkodzić i w przypadku każdego użytkownika pe-ceta jest to działanie absolutnie niewskazane! Gdy używamy systemu Windows 95 lub 98, używamy opcji „Formatuj” wyświetlanej w menu kontekstowym każdego dysku twardego (a właściwie każdej partycji/dysku logicznego) w oknie „Mój Komputer”. Formatowanie w Windows przebiega znacznie szybciej niż w przypadku zastosowania dosowego polecenia „format”.

### **Ukryte funkcje fdisk i format**

Opis nie udokumentowanych funkcji narzędzi dyskowych: fdisk i format. Programy te, będące zarazem poleceniami linii komend MS-DOS, może nie są najczęściej używane, ale należą do istotnych narzędzi umożliwiających nam skonfigurowanie dysku twardego i przygotowanie go do przechowywania oprogramowania. Narzędzia dyskowe nie należą do „bezpiecznych”, dlatego zanim rozpocznie się pracę z nimi warto zachować na odrębnym nośniku przynajmniej ważne dokumenty i/lub programy. Na początek program ustawiający bądź modyfikujący układ partycji na dysku/dyskach twardych – fdisk.

#### *Fdisk*

Udokumentowanych przełączników polecenia fdisk jest mało, jednak możliwości jest znacznie więcej. Program dzielący dysk na partycje możemy wywołać wpisując po prostu w linii poleceń komendę „fdisk”. Wywoła to najbardziej znany chyba interfejs tego narzędzia. Badanie funkcji, jakie możemy realizować w tym programie, zacznę od analizy przełączników polecenia „fdisk”. Każda komenda w DOS-ie (a więc także w Windows) może mieć jakieś parametry wywołania – inaczej: argumenty zwane też przełącznikami lub opcjami. Aby poznać listę dostępnych dla danego polecenia przełączników, wystarczy w linii poleceń DOS wywołać to polecenie z opcją “/?”. W przypadku programu fdisk otrzymamy tym sposobem informacje o dwóch dodatkowych parametrach: /STATUS oraz /X. O ile pierwszy z parametrów nie wymaga komentarza, to wyjaśnienie towarzyszące drugiemu z nich może się wydać nieco enigmatyczne. Chodzi mianowicie o ignorowanie obsługi partycji LBA. Innymi słowy: fdisk uruchomiony z tą opcją nie

utworzy na dysku partycje LBA. Powyższe dwie opcje to jedynie garstka możliwości, jakimi dysponuje fdisk. A oto znacznie bogatszy zestaw opcji tego programu:

```
/mbr
```

Ta opcja powoduje ponowne zapisanie zawartości obszaru startowego pierwszego dysku twardego o nazwie Master Boot Record (MBR) nową zawartością (więcej informacji o MBR w ramce „Co to jest MBR?”). Opcja ta może być przydatna np. w przypadku, gdy istnieje prawdopodobieństwo, że we wspomnianym obszarze zagnieździł się wirus lub istniejące dane zostały uszkodzone z innych przyczyn. Wywołanie tej opcji nie powoduje usunięcia partycji i nie usuwa żadnych istotnych danych z dysku. Trzeba jednak pamiętać o tym, że w przypadku, gdy na komputerze mamy zainstalowane dwa lub więcej systemów operacyjnych (np. Windows 98 i Linuks), których uruchamianie odbywa się za pomocą odpowiedniego programu, tzw. Boot Managera, to wykonanie instrukcji: fdisk/mbr spowoduje usunięcie Boot Managera i tym samym niemożność uruchomienia innego niż DOS/Windows 95 i 98 systemu operacyjnego. Z drugiej strony jest to najszybszy sposób zlikwidowania możliwości wybiórczego uruchamiania systemu.

```
/cibr
```

Opcja bardzo podobna do powyższej, z tą różnicą, że w tym przypadku istnieje możliwość zdefiniowania dysku, na którym ma być wstawiony lub nadpisany MBR. Wywołanie fdiska z tą funkcją ma następującą składnię:

```
fdisk /cibr <dysk>
```

Tu uwaga. W miejsce sekwencji <dysk> wstawiamy NUMER dysku, a nie jego oznaczenie literowe. Numer oznacza fizyczną kolejność podłączenia dysku w komputerze. Przykładowo - dysk podłączony do pierwszego kanału interfejsu EIDE jako urządzenie master ma numer 1, kolejny dysk na tym samym kanale, ale jako urządzenie slave - ma numer 2 itd.

```
/pri
```

Dzięki tej opcji mamy możliwość utworzenia partycji podstawowej (primary partition) bez potrzeby ładowania interfejsu programu. Partycja zakładana jest bezpośrednio z linii poleceń. Utworzona w ten sposób partycja jest automatycznie ustawiana jako aktywna. Składnia polecenia z tą opcją jest następująca:

```
fdisk /pri:<rozmiar> <dysk>
```

gdzie: <dysk> oznacza numer dysku (wyjaśnienie wyżej), <rozmiar> pozwala na zdefiniowanie, ile miejsca chcemy przeznaczyć na tworzoną partycję. Jeżeli przypadkowo zdarzy się, że podamy rozmiar większy niż ten, jakim dysponuje cały dysk, automatycznie na potrzeby partycji przeznaczona zostanie cała powierzchnia "twardziela". W zależności od tego, jaki podamy rozmiar, partycja zostanie utworzona z wykorzystaniem systemu plików FAT lub FAT32. Ten pierwszy dotyczy partycji mniejszej niż 512 MB.

```
/prio
```

W zasadzie wykonuje to samo co /pri, ale dodano jedną dyskusyjną właściwość. Mianowicie partycja o rozmiarze większym niż 512 MB będzie bazować na systemie plików FAT. Składnia taka sama jak dla opcji /pri.

/ext

Pozwala na utworzenie bezpośrednio z linii poleceń partycji rozszerzonej na dysku. Składnia jest następująca:

```
fdisk /ext:<rozmiar> <dysk>
```

<rozmiar> oznacza oczywiście przestrzeń, jaką przeznaczamy na partycję (jeżeli podamy więcej niż to jest fizycznie możliwe, partycji zostanie przypisane całe wolne miejsce), <dysk> – numer dysku.

Należy pamiętać, że partycja rozszerzona nie może być bezpośrednio wykorzystywana przez system. Aby móc dysponować obszarem na nią przeznaczonym jako kolejnym dyskiem (oznaczonym literą) z poziomu Windows musimy utworzyć na niej dyski logiczne.

/log

Tworzy logiczny dysk na utworzonej poprzednio partycji rozszerzonej. Opcja ta nie może być wykorzystana samodzielnie i musi występować razem z poprzednio przedstawioną opcją /ext. W takim przypadku składnia przedstawia się następująco:

```
fdisk /ext:<rozmiar> <dysk>
```

/log:<rozmiar>

Pamiętać należy o tym, by w tym przypadku rozmiar nie przekraczał maksymalnej, dostępnej, wolnej przestrzeni na dysku, i aby <rozmiar> był identyczny w obu wystąpieniach.

/logo

Działa praktycznie identycznie jak /log. Jediną różnicą jest możliwość utworzenia dysku logicznego FAT o rozmiarze większym niż 512 MB.

/fprmt

Opcja ta wywołuje program fdisk bez wyświetlania ekranu z zapytaniem o obsługę dużych dysków (czyli uwzględniania systemu FAT32 dla partycji/dysków większych niż 512 MB). Zamiast tego użytkownik jest pytany o system plików za każdym razem, kiedy tworzy partycję. Opcja ta ma jedną dodatkową zaletę: umożliwia utworzenie partycji/dysku FAT32 o pojemności mniejszej niż 512 MB. Składnia wygląda tak:

```
fdisk /fprmt
```

/actok

Pozornie wywołanie programu fdisk z tą opcją nie różni się od wywołania bez jakiegokolwiek opcji. Jednak w tym przypadku nie jest sprawdzana integralność struktury dysku (disk integrity) podczas tworzenia partycji/dysku logicznego. Dzięki temu cały proces odbywa się nieco szybciej. Składnia:

```
fdisk /actok
```

## Format

Polecenie format domyślnie zawiera sporą listę opcji, jednak kilka dodatkowych może, przy zachowaniu ostrożności, przyspieszyć pracę

Zazwyczaj kolejną czynnością po założeniu partycji i ewentualnym zdefiniowaniu dysków logicznych jest ich formatowanie. Najczęściej wykorzystywanym programem jest także aplikacja

wywoływana z linii poleceń DOS – format. I tym razem mamy do dyspozycji kilka nie udokumentowanych przełączników.

```
/autotest
```

Opcja wprost idealna do formatowania dyskietek. Jej wywołanie powoduje natychmiastowe sprawdzenie aktualnego formatu nośnika i ponowne jego sformatowanie. Zaletą tej opcji jest to, że program format nie wyświetla żądania włożenia dyskietki, ponadto nie pojawia się żądanie potwierdzenia operacji formatowania, nie ma monitu o podanie etykiety formatowanego dysku. W praktyce użytkownik ogranicza się do wprowadzenia polecenia z opcją, a cała reszta odbywa się automatycznie. Składnia:

```
format <dysk:> /autotest
```

gdzie <dysk:> oznacza tym razem literę danego dysku (np. A: dla dyskietki).

```
/backup
```

Opcja bardzo podobna w działaniu do powyższej, z tą różnicą, że w tym przypadku program format żąda wpisania etykiety sformatowanego nośnika oraz wyświetla podsumowanie działania. Składnia:

```
format <dysk:> /backup
```

```
/u
```

Bardzo niebezpieczna opcja, ale też zalecana, gdy chcemy zamazać wszelkie ślady po naszych działaniach. Standardowo nawet gdy sformatujemy jakikolwiek nośnik, wciąż mamy możliwość odzyskania danych, gdyż polecenie format nie zamazuje fizycznie wszystkich obszarów dysku, a jedynie wprowadza informację niezbędną do tego, by system interpretował dysk jako pusty. Zastosowanie opcji /u powoduje, że operacji formatowania towarzyszy zniszczenie wszelkich danych na formatowanym nośniku przez zapisanie każdego bajtu wartością „F6” (szesnastkowo). Składnia:

```
format <dysk:> /u
```

```
/Z:n
```

Opcja dla dysków używających systemu plików FAT32. Składnia polecenia z tą opcją jest następująca:

```
format <dysk:> /Z:n
```

gdzie <dysk:> oznacza literę formatowanego dysku, natomiast n określa liczbę sektorów (każdy po 512 bajtów) w pojedynczym klastrze (najmniejszej jednostce logicznej dysku w systemach FAT i FAT32). Dla przykładu: n=1 tworzy na formatowanym nośniku klastry 512-bajtowe, n=2 - 1024-bajtowe (1 kB) itd. Należy pamiętać, że choć mniejszy klaster dysku oznacza mniejsze straty pojemności przy przechowywaniu dużej liczby plików, to zmiana domyślnej wielkości klastra dla systemu FAT32 (4 kB) może spowodować nieprawidłowe działanie programów antywirusowych lub innych narzędzi dyskowych.

### **3.5. Uszkodzenia.**

Niezawodność współczesnych dysków twardych wyraża się obecnie średnim czasem między-uszkodzeniowym rzędu miliona godzin. Z pozoru wydaje się to bardzo wiele, ale, gdy się bliżej przyjrzeć, bezpieczeństwo danych na dysku twardym jest co najmniej iluzoryczne. Milion godzin to przeszło 114 lat. Współczynnik MTBF (Mean Time Between Failures) wynoszący milion godzin oznacza nie tylko, że dysk powinien pracować bezawaryjnie przez tyle czasu – w uproszczeniu oznacza to, że spośród 1000 dysków w ciągu bieżącego roku przeszło 8 ma prawo ulec awarii! A jeśli nawet założymy, że dyski pracują zaledwie po 8 godzin dziennie, to i tak wśród tysiąca dysków musimy się liczyć z blisko trzema awariami w ciągu roku. Awarie dysku mają różny charakter. Może to być uszkodzenie układów zapisu i odczytu, w tym złożonej elektroniki dysku, awaria układu napędowego czy układu pozycjonowania głowic, a wreszcie, co jest najczęściej spotykane, mechaniczne uszkodzenie nośnika magnetycznego na powierzchni któregoś z talerzy. Wszystkie dzieła ludzkiego geniuszu mają ograniczoną niezawodność. Zastanawia jednak fakt – w jaki sposób powstają mechaniczne uszkodzenia powierzchni dysku, jeśli głowice nie dotykają bezpośrednio tych powierzchni? Mimo relatywnie dużych rozmiarów, dyski twarde stanowią arcydzieła mechaniki precyzyjnej. Przy typowej gęstości zapisu, szerokość pojedynczej ścieżki zapisu wynosi zaledwie ok. 0,01 mm. Szerokość głowicy odczytującej, wykonanej jako element magnetorezystywny, wynosi ok. 80% szerokości ścieżki – to odpowiada ostrzu nieco tylko stępionej żyłki! Każde dotknięcie powierzchni dysku przez głowicę odpowiada dotknięciu takim właśnie ostrzem. Warstwa nośnika magnetycznego na powierzchniach talerzy dysków pokryta jest bardzo cienką warstwą lakieru ochronnego. W normalnych warunkach eksploatacji twardość powierzchni ochronnej najzupełniej wystarcza – start i lądowanie głowic wiążą się co prawda z przesuwaniem ich po powierzchni talerza, ale występujące naciski są za małe, by zarysować powierzchnię ochronną. Podczas pracy dysku głowice przesuwają się nad powierzchnią talerzy na „poduszce powietrznej” o wysokości kilkunastu mikrometrów, wytwarzanej dzięki ruchowi talerzy, a dopuszczalne nierówności powierzchni nie przekraczają 10% wysokości „lotu” głowicy. W takich warunkach nośnik dysku nie ma prawa ulec uszkodzeniu. Panuje powszechne przekonanie, że dysk nie działający jest odporny na wstrząsy i uderzenia. Tymczasem, co może być zaskakujące, źródłem większości uszkodzeń powierzchni roboczych dysku są właśnie wstrząsy i uderzenia, których napęd doznał w stanie spoczynku. W stanie spoczynku głowice leżą na wydzielonych obszarach powierzchni talerzy, zwanych strefą lądowania (landing zone), przyciśnięte do powierzchni przez odpowiedni układ sprężysty ramienia głowicy. Cóż się stanie, jeśli taki „zaparkowany” dysk dozna silnego, krótkotrwałego wstrząsu? Głowica oderwie się od powierzchni, wyginając sprężyste ramię, a następnie, w wyniku jego drgań, kilkakrotnie uderzy w powierzchnię, za każdym razem odbijając się od niej. Zwraca uwagę fakt, że głowica w takiej sytuacji nie uderza swoją powierzchnią, ale krawędzią! Twarda powierzchnia ochronna jest, niestety, zbyt krucha, by mogła to przy silniejszych wstrząsach wytrzymać – uderzająca głowica odłupuje drobne fragmenty ochronnego lakieru.

Wydawać by się mogło, że nawet ewentualne uszkodzenie powierzchni w strefie lądowania nie powinno spowodować obniżenia sprawności dysku – przecież w tym obszarze nie ma żadnych danych. Rzeczywiście, ale powstałe tam drobne okruchy materiału przemieszczają się, wraz z powietrzem, po całym wnętrzu napędu. Drobne, ale wielokrotnie większe od grubości poduszki powietrznej, unoszącej głowicę. A jeśli któryś z nich dostanie się pomiędzy głowicę a powierzchnię wirującego talerza, następują kolejne drgania głowicy i jej ramienia oraz kolejne uderzenia głowicy – tym razem już w roboczą powierzchnię dysku! Oprócz uszkodzeń powierzchni, na tyle drobnych, że układy korekcji błędów wbudowane w elektronikę dysku, poradzą sobie z powodowanymi przez nie błędami, powstają jeszcze nowe okruchy. Im jest ich więcej, tym częściej zdarza im się wpadnięcie pod głowicę i tym częściej powstają nowe uszkodzenia i nowe drobiny. Proces degradacji wartości użytkowej dysku postępuje lawinowo, tym bardziej, że przy uszkodzonej powierzchni strefy lądowania przy każdym starcie i lądowaniu głowicy mogą powstawać kolejne uszkodzenia. Na jakiego rodzaju wstrząsy narażony jest dysk od momentu opuszczenia taśmy produkcyjnej, do chwili, kiedy trafi do komputera? Co mu grozi po drodze, a czym możemy mu zaszkodzić sami? Odpowiedzi na te pytania może w pewnym stopniu dostarczyć zamieszczony rysunek – wynika z niego, że dysk zamontowany w komputerze jest względnie bezpieczny, nawet upuszczenie komputera na twarde podłoże nie powinno spowodować poważniejszych szkód. Dużym zagrożeniem dla dysku jest również sam proces montażu komputera. W tej fazie łatwo „nabawić się” kłopotów na przyszłość. O uderzenie metalowym narzędziem wcale nietrudno – wystarczy „obsunięcie” ręki, uderzenie dyskiem o konstrukcję obudowy też może się zdarzyć. A jeśli ktoś ma pecha, to i o upadek dysku na twarde podłoże wcale nietrudno. Wszystkie te „gwałtowne zdarzenia” dysk znosi pozornie bez szwanku – po zmontowaniu komputera działa poprawnie i nic nie wskazuje na to, by cokolwiek mu dolegało.

### **Ból głowy producentów**

Uszkodzenia, których źródłem są wstrząsy, jakich doznał dysk w czasie między wyprodukowaniem a zamontowaniem w komputerze, stanowią według danych producentów przyczynę około 40% wszystkich awarii dysków twardych i przeszło 90% uszkodzeń powierzchni dysków. Lekarstwem na to stały się pewne zmiany w konstrukcji dysków, zmierzające do ograniczenia tego typu uszkodzeń. Zmiany te w większości przypadków sprowadzają się do odpowiednich rozwiązań konstrukcyjnych – najważniejsze jest tu wyeliminowanie drgań głowicy i jej wielokrotnego uderzania o powierzchnię po wstrząsie. Tego rodzaju rozwiązaniem jest, stosowany przez firmę Quantum, SPS (Shock Protection System). Również inni producenci od pewnego czasu zwracają uwagę na bezpieczeństwo dysku w czasie między opuszczeniem taśmy produkcyjnej a zainstalowaniem w komputerze, stosując własne rozwiązania, jak np. SeaShield Seagate. Obecnie stosuje się narzędzia które sprawdzają stan dysku. Ważną ich cechą jest zdolność do wykorzystywania w celach diagnostycznych specjalnych procedur, wbudowanych w oprogramowanie napędów. Przy bardzo skutecznych mechanizmach korekcji błędów, jakie są



stosowane w układach odczytu, drobniejsze uszkodzenia pozostawałyby niezauważone – dopiero uszkodzenie uniemożliwiające poprawny odczyt mogłoby zostać zarejestrowane. Należy zwrócić uwagę na fakt, że eliminowanie wadliwych sektorów nie usuwa przyczyn ich uszkodzenia – jeśli wewnątrz obudowy znalazły się luźne okruchy z uszkodzonych powierzchni, to proces niszczenia będzie postępował. Dlatego większość wspomnianych systemów stosuje statystyczną ocenę liczby wykrytych mikrodefektów, umożliwiającą, przy regularnym stosowaniu programu testującego, dość efektywną ocenę aktualnego stanu dysku i jego „perspektyw na przyszłość”.

### **Samonaprawiające się dyski**

Każda usterka sprzętu, którego używamy wywołuje u nas marzenia o urządzeniach, które powiadamiałyby nas, że wystąpiła awaria, a jeszcze lepiej – same się naprawiały. Nawet nie wiemy, że to już nie marzenia, ale rzeczywistość, przynajmniej jeżeli chodzi o dyski twarde. Najnowsze osiągnięcia w dziedzinie technologii dysków twardego sprawiają, że napędy dyskowe uzyskują zdolność nie tylko do monitorowania własnej sprawności, lecz także „samonaprawiania się” w przypadku typowych usterek.

#### *Algorytm ECC*

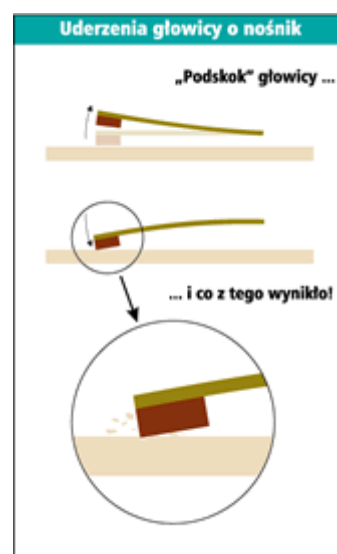
W dysku twardym dane cyfrowe są zapisywane na talerzu magnetycznym, a potem odczytywane, zasadniczo w postaci analogowej. Podobnie jak przy każdym nośniku analogowym, danym zapisanym

na dysku towarzyszą szumy tła, a sam nośnik jest podatny na uszkodzenia fizyczne. Rozpoznanie faktu, że dane zostały uszkodzone oraz podjęcie jakichkolwiek działań naprawczych jest możliwe dzięki temu, że z zasady do zapisywanej informacji dodaje się pewną informację dodatkową, która jest uzależniona od zawartości informacji oryginalnej.

W dyskach twardego stosuje się zaawansowane metody obliczania i kodowania sum kontrolnych, określane jako ECC (Error Correcting Codes – kody korygujące błędy). Chociaż teoria z tym związana jest ogromnie skomplikowana, w praktyce wyznaczenie kodu korekcyjnego dla danych można w miarę prosto zrealizować za pomocą sprzętu lub oprogramowania. Dzięki dobremu algorytmowi ECC możliwe jest nie tylko wykrywanie błędów, lecz także odtworzenie uszkodzonej informacji. Obliczanie kodu korekcyjnego wchodzi w skład procesu odzyskiwania danych, w którym ponadto stosuje się takie techniki, jak wielokrotny odczyt przy kolejnych obrotach talerza z drobnymi zmianami parametrów odczytu, co daje różne „kąty widzenia” uszkodzonych danych. Wszystkie te „sztuczki” pozwalają na odczytanie danych z sektora, który nie nadaje się do dalszego użytku. Co dalej?

#### *Sektory na zapas*

Dyski twarde zawierają pewną liczbę „zapasowych sektorów”, które nie są bezpośrednio dostępne dla użytkownika, lecz służą do zastępowania wadliwych sektorów wykrytych na dysku.



Gdy jeden z zapasowych sektorów zostanie zaalokowany w zastępstwie sektora uszkodzonego, z punktu widzenia użytkownika dysku wygląda to tak, jakby uszkodzenie zostało naprawione. Jeżeli wszystkie uszkodzone sektory są odwzorowywane na dobrych sektorach zapasowych, to dysk z punktu widzenia użytkownika jest całkowicie sprawny. Alokacja zapasowych sektorów może odbywać się z wyprzedzeniem, w miarę zużywania się dysku. Metoda ta polega na tym, że podczas odczytu bloku danych układ elektroniczny, odpowiedzialny za ECC, dokonuje inteligentnej analizy jakości sektora. W niektórych przypadkach dane zostają zapisane nieprawidłowo – na przykład wskutek mechanicznego wstrząsu napędu podczas zapisu – i wówczas całkowita naprawa sprowadza się jedynie do ponownego zapisu tych samych danych. Jeżeli jednak analiza podejrzanego sektora wykazuje, że nie zapewnia on należytej niezawodności, wówczas układ sterowania napędu może podjąć decyzję wykorzystania sektora zapasowego i zapisania w nim odzyskanych danych.

### **SMART - przewidzieć awarię dysku**

SMART oznacza Self Monitoring And Reporting Technology (technologia samoczynnego monitorowania i powiadamiania). Jest to uporządkowana metoda wykonywania przez napęd dyskowy analiz statystycznych własnego funkcjonowania, dokonywania na tej podstawie inteligentnych przewidywań co do zbliżających się awarii oraz powiadamiania o tym użytkownika. SMART wykorzystuje nadmiarową moc obliczeniową procesora napędu dyskowego i prowadzi analizę rozmaitych parametrów operacyjnych, takich jak stopa błędów, liczba powtórzeń, częstość relokacji uszkodzonych sektorów, cykle startu – stopu itd. Informacja ta jest zbierana i poddawana obróbce statystycznej na podstawie znanych charakterystyk operacyjnych sprawnego dysku. W ten sposób uzyskuje się możliwość ostrzeżenia z wyprzedzeniem, że zbliża się awaria dysku. Chociaż obecnie nie ma sposobu, by technologia SMART pozwoliła przewidzieć nagłą awarię dotychczas zupełnie sprawnego dysku, to jednak zapewnia ona skuteczne ostrzeżenie o zbliżającej się awarii w około 30 do 40 procentach przypadków. Aby można było skorzystać z technologii SMART, w systemie musi zostać zainstalowany odpowiedni agent (program obsługi). Odzyskiwanie danych w nowoczesnych napędach dyskowych jest bardzo sprawne – napęd zasygnalizuje błąd odczytu dopiero po wyczerpaniu daleko idących środków zaradczych. Możliwość alokacji zapasowych, dobrych sektorów na miejsce uszkodzonych oznacza, że usterki – które w innym wypadku byłyby klasyfikowane jako awarie dysku - mogą być aktywnie kontrolowane, dzięki czemu wydłuża się użyteczny czas eksploatacji urządzenia. SMART zapewnia prognozowanie możliwych awarii dysku, dzięki czemu dane z dysku o pogarszającej się jakości mogą być zapisane w kopii zapasowej, a dysk wymieniony, zanim dojdzie do katastrofalnej utraty danych. Wszystkie te mechanizmy opierają się jednak na zdolności napędu do właściwego reagowania na usterki przez korekcję błędów, relokację sektorów oraz analizę i rejestrowanie wyników. Działania takie mogą dotyczyć tylko tych części dysku, które są użytkowane, a wskutek tego stan znacznej części powierzchni dysku może przez długi czas być niezna-

ny, a wtedy błędy skądinąd możliwe do naprawienia stopniowo stają się coraz poważniejsze, zaś analizy statystyczne prowadzone przez SMART zostają zafałszowane.

### **Aplikacje testujące i profilaktyczne**

Istnieje wiele łatwo dostępnych programów, służących do testowania powierzchni dysku i wykrywania błędów. Przykładem takiej aplikacji może być Data Lifeguard firmy Western Digital. Działanie programu Data Lifeguard polega na tym, że automatycznie testuje on cały obszar danych na dysku raz na 8 godzin pracy – odpowiada to mniej więcej jednemu dniu pracy typowego napędu dyskowego. Błędy wykryte przez Data Lifeguard podczas testowania mogą zostać naprawione; można też dokonać realokacji uszkodzonych sektorów. Odbywa się to zanim jeszcze system zechce skorzystać z wolnego obszaru, w którym leżą te sektory. Program działa bez interwencji użytkownika, wykorzystując okresowe braki aktywności. Jeżeli podczas działania programu użytkownik wykonuje jakieś operacje, testowanie zostanie po prostu zawieszona, a odczuwana przez niego wydajność całego systemu nie ulegnie pogorszeniu. Podczas testowania powierzchni dysku Data Lifeguard stosuje różnego rodzaju algorytmy odzyskiwania danych i naprawy błędów, dzięki czemu program może przenosić dane z uszkodzonych lub pogarszających się obszarów dysku do obszarów sprawnych. Dodatkową zaletą tej metody działania jest to, że jeżeli równocześnie z Lifeguardem działa SMART, to testowanie dysku zapewnia znacznie lepszą analizę statystyczną i bardziej dokładne prognozy co do stanu dysku.

### **Narzędzia systemowe**

Ważną czynnością jest defragmentacja dysku. Jest to proces mający na celu fizyczne uporządkowanie struktury katalogów i plików na dysku twardym w ten sposób, aby system operacyjny miał do nich jak najszybszy dostęp. Przyspieszy to działanie dysku twardego i, co za tym idzie, całego systemu. Poza tym oszczędza mechanizmy dysku, gdyż ramię nie musi bezustannie przemieszczać się nad talerzami w poszukiwaniu potrzebnych danych. Do defragmentacji dysku wykorzystuje się program Defragmentator. Jego obsługa jest prosta – ogranicza się do wybrania litery dysku, który chcemy defragmentować, i potwierdzenia OK. Przycisk „Zaawansowane...”, ujawniający się podczas przerwania procesu klawiszem „Esc”, zawiera opcje pozwalające wybrać jedną spośród trzech metod defragmentacji. Program ma również opcję sprawdzania dysku na występowanie błędów. Proces defragmentacji jest długotrwały (zależy od rozmiarów dysku twardego oraz od stopnia fragmentacji jego zawartości). Następną formą dbania o dysk twardy jest korzystanie z programów diagnostycznych typu ScanDisk. Programy tego typu sprawdzają poprawność plików folderów, odszukują fragmenty utraconych plików, sprawdzają pliki na występowanie nieprawidłowych nazw, dat i godzin, wykrywają skrzyżowane pliki oraz skanują powierzchnię dysku twardego na wypadek występowania uszkodzeń fizycznych (tzw. bad sectors). Program MS ScanDisk, w który wyposażony jest system Windows 98, znajduje się w tym samym miejscu, co opisywany poprzednio Defragmentator. Po uruchomieniu i wybraniu odpowiedniej litery przypisanej napędowi ScanDisk wykona szereg testów oraz spróbuje dokonać ewentualnych napraw w strukturach katalogów i plików, a uszkodzenia wy-

kryte jako fizyczne zostaną zaznaczone tak, aby żaden program nie starał się ich w przyszłości wykorzystać. Do wyboru są dwa rodzaje testowania: standardowy, w którym sprawdzana jest poprawność plików i folderów, oraz dokładny – który po teście standardowym skanuje powierzchnię dysku w poszukiwaniu błędów. Opcje programu dają do wyboru możliwość skanowania jednocześnie obszarów systemowych i obszarów danych lub każdego z tych obszarów osobno. Można również wyłączyć proces testowania zapisu i naprawiania uszkodzonych sektorów w plikach ukrytych i systemowych. Ustawienia zaawansowane pozwalają modyfikować m.in. takie opcje, jak: wyświetlanie podsumowania, zawierającego informacje o dysku oraz zestawienie wykrytych i poprawionych błędów (wyświetlane po zakończeniu procesu sprawdzania dysku); generowanie pliku dziennika (jest to szczegółowy raport z sesji programu ScanDisk zapisywany w pliku SCANDISK.LOG w katalogu głównym dysku systemowego); postępowanie z plikami wykorzystującymi ten sam obszar dysku (tzw. skrzyżowanymi) oraz z fragmentami utraconych plików. Program ten jest uruchamiany automatycznie za każdym razem, gdy system zostanie nieprawidłowo zamknięty.

### **Boot - manager-y**

Podzielenie twardego dysku na partycje może przynieść szereg doraźnych korzyści, od łatwiejszego zarządzania strukturą dysków i katalogów, do możliwości instalowania innych równorzędnych systemów operacyjnych. Aby jednak wygodnie nimi dysponować potrzebne są programy, które będą trzymać nadzór nad całością i pozwolą efektywnie zarządzać przestrzenią dyskową. Rolę taką pełnią Bootmanagery które zależnie od rodzaju, oferują różne mniej lub bardziej przydatne możliwości, którymi należy kierować się przy ich wyborze, tak aby maksymalnie dopasować możliwości programu do własnych wymagań i oczekiwań. Główne zalety jakie mogą cechować tego typu programy to:

- Umiejętność zarządzania partycjami – czyli możliwość dzielenia dysków na partycje i oznaczania jednej z nich jako aktywną, tzn. tą z której będzie uruchamiany system po włączeniu komputera.
- Funkcjonalność programu – to przede wszystkim umiejętność współpracy z możliwie wieloma systemami różnych plików, takich jak: NTFS, FAT 16, 32 (DOS, Windows), HPFS (OS/2) czy Ext2fs (Linux) oraz umiejętność wzajemnej ich konwertacji z jednego na drugi. Umiejętność uruchomienia programu z dyskietki przydatna w przypadku zarażenia lub uszkodzenia sektora MBR przez wirus a także możliwość utworzenia z poziomu programu dyskietki startowej.
- Interfejs - czyli łatwość i wygoda obsługi programu najlepiej za pomocą graficznych okienek.

Oto niektóre z takich programów: Beret's System Loader, Bootmanager 98/200 Supra, Bootmanager ++, BootIt, Boot Menu, Bootstar, Do it Bootmanager, FODisk, Grub, LILO, MasterBooster, NTLoader, OS/2 Boot, Manager, Paragon Festplatten-Manager, Partition Magic, Ranish Partition Manager, System Commander Deluxe, Vamos, Win

Bootmanager. Przed rozpoczęciem pracy z partycjami jak i w trakcie, należy jednak zachować dużą ostrożność, gdyż łatwo można spowodować utratę wszystkich zgromadzonych na dysku danych. Dlatego też zaleca się, aby zawsze mieć kopię najbardziej istotnych informacji zgromadzonych na modyfikowanych dyskach fizycznych, gdyż nigdy nie wiadomo czy po przepartycjonowaniu dysku, znajdziemy je jeszcze nienaruszone na dawnym miejscu.

## **4. KARTY GRAFICZNE.**

### **4.1. Monitory.**

Podczas pracy z komputerem jednym z najważniejszych czynników wpływających na jej komfort jest monitor. Obecnym standardem są monitory o przekątnej 15 cali, które powoli zostają wypierane przez modele 17 calowe lub większe. Jednak w przypadku tradycyjnych monitorów duży ekran wymusza znaczne wymiary urządzenia. Wad tego rodzaju pozbawione są płaskie wyświetlacze LCD (ang. Liquid Crystal Display, czyli wyświetlacze ciekłokrystaliczne). Początkowo stosowane były one w komputerach przenośnych ostatnio jednak coraz częściej pojawiają się zamiast zwyczajnych monitorów. LCD wyposażone są w ekrany mierzące od 14 do 20 cali, jednak dominują modele o przekątnej 15.1 i 18.1 cala. Wartości te odpowiadają rzeczywistym rozmiarom widocznej powierzchni ekranu, a nie całej lampy jak w przypadku zwykłych monitorów (dla 15 cali rzeczywista wielkość to 13.8 cala, a dla 17cali – 15.7 cala). Najbardziej widocznymi cechami odróżniającymi wyświetlacze LCD od tradycyjnych monitorów są ich wymiary oraz waga. Najgrubsze panele nie przekraczają zwykle 10 cm, a rekordzistą w dziedzinie minimalizacji jest wyświetlacz LCD, który ma zaledwie 14 mm grubości. Najcięższe 20 calowe monitory LCD ważą tyle, ile przeciętny tradycyjny monitor o przekątnej 15 cali.

Technologie wyświetlaczy LCD i tradycyjnych monitorów CRT łączy ta sama metoda tworzenia obrazu z pojedynczych punktów, jednak wszystkie inne zasady działania nie mają ze sobą już nic wspólnego.

Monitor monochromatyczny : najważniejszym, największym i zarazem najdroższym elementem każdego monitora jest tzw. *lampa obrazowa*, czyli *kineskop*. Jest ona zbudowana z opróżnionej z powietrza „bańki” szklanej, w podstawie której znajduje się działło elektronowe z żarzoną katodą (kineskop monochromatyczny). Emituje ono wąski strumień elektronów ogniskowanych przez układ cylindrycznych elektrod działła elektronowego. Elektrony są przyspieszane do ogromnej prędkości przez pole elektryczne wytwarzane przez anodę, do której jest przyłożone bardzo duże napięcie (15 do 30 kV). Anodę tę stanowi wewnętrzna część szyby ekranu która zarazem jest pokryta materiałem fluoryzującym (luminoforem). W przestrzeni między działłem a ekranem elektrony są przyspieszane do takiej prędkości, że energia kinetyczna jaką posiadają wystarcza do pobudzenia luminoforu do świecenia. Poprzez zmianę wartości napięcia na katodzie zmieniamy energię elektronów uderzających w luminofor, a przez to możemy uzyskać płynną regulację jasności wyświetlanych punktów. W przestrzeni tej znajdują się także cewki odchylające, dzięki którym, poprzez padanie na nie odpowiedniego napięcia, cienka wiązka elektronów zostaje skierowana w dowolne miejsce na ekranie i może spowodować zaświecenie na nim pojedynczego punktu. Wiązka elektronów jest odchylana w taki sposób, że najpierw przebiega ona po kolejnych punktach górnej linii, od lewej do prawej strony ekranu. Po dotarciu do końca linii, wiązka elektronów zostaje na krótko wygaszona (przez dodanie niskiego napięcia na anodę) i zwrócona na lewą stronę ekranu do następnej linii. Po na-

rysowaniu ostatniej linii obrazu (wiązka znajduje się w dolnym prawym rogu ekranu) zostaje znowu wygaszona i następuje powrót do lewego górnego rogu ekranu i rysowanie następnego obrazu. Za odchylenie wiązki elektronów w poziomie i pionie odpowiadają układy odchylenia monitora.

Monitor kolorowy : kineskopy zastosowane w monitorach kolorowych mają nieco odmienną konstrukcję. Mają one trzy oddzielne działa elektronowe, z których każde emituje wiązkę elektronów, odpowiedzialną za wyświetlenie jednego z trzech podstawowych kolorów. Sposób ogniskowania i odchylenia trzech wiązek elektronów jest również bardziej złożony niż to miało miejsce w przypadku kineskopów monochromatycznych. Są one odpowiednio ogniskowane przez zespół soczewek i pryzmatów elektronowych. Stanowią je cewki siodłowe lub toroidalne, umieszczone tuż za działem elektronowym (są przyklejone do tzw. szyjki kineskopu). Dzięki temu otrzymuje się odpowiednią zbieżność kolorów, czystość oraz geometrię obrazu. W takim przypadku wewnętrzna część ekranu nie jest pokryta jednolitym luminoforem, tak jak to było w przypadku kineskopów monochromatycznych, ale trzema warstwami i to w taki sposób, iż ekran pokryty jest pojedynczymi triadami RGB, składającymi się z leżących obok siebie trzech mikroskopijnej wielkości plamek R, G i B, (po jednej z każdej warstwy). Wiązka z pojedynczego działu elektronowego (np. Red) pada na odpowiadającą jej plamkę (Red), itp. Dzięki trzem niezależnym strumieniom elektronów każda z tych cząstek może być naświetlana z inną intensywnością. Nawet jeśli zostaną wzbudzone wszystkie trzy plamki i spojrzymy na nie z pewnej odległości, będą one zlewały się w całość, tworząc jednolitą barwę pochodną. Możliwe jest to dzięki wykorzystaniu ograniczonej rozdzielczości oka ludzkiego. Swobodna zmiana natężenia poszczególnych strumieni pozwala na uzyskanie pełnej palety barw. Aby wiązki elektronowe z odpowiednich dział RGB trafiały we właściwe plamki RGB i nie powodowały świecenia plamek sąsiednich, stosuje się specjalnej budowy maskownicę. Ponadto kineskopy kolorowe wyposażone są w tzw. pętlę rozmagnesowującą (opłata ona bańkę kineskopu), wytwarzającą niewielkie stałe pole magnetyczne, zabezpieczające kineskop przed rozproszonymi polami magnetycznymi. Proces rozmagnesowywania ma miejsce najczęściej po włączeniu zasilania, lub może być inicjowany specjalnym przyciskiem. Kineskopy maskowe możemy podzielić na dwa podstawowe typy: *kineskopy typu Delta* - w kineskopie tego typu zastosowano maskownicę (maską) perforowaną. Jest nią cienka, czarna folia posiadająca określoną liczbę okrągłych otworów. Nazwa „Delta” odzwierciedla sposób położenia poszczególnych pikseli: jeden kolorowy punkt na ekranie tworzą trzy leżące obok siebie jednobarwne punkty, tworzące trójkąt równoboczny. Tak samo względem siebie umiejscowione są trzy działa elektronowe. Pierwsze monitory z tym kineskopem wyróżniały się wypukłą powierzchnią (kineskop jest wycinkiem kuli), dużymi zniekształceniami geometrycznymi obrazu. Przy porównaniu z kineskopami typu Trinitron lampy kineskopowe typu Delta odznaczają się mniejszą ceną, są lżejsze, mają mniejsze gabaryty oraz są bardziej odporne na wstrząsy mechaniczne i dostarczają zwykle bardziej jednorodnego obrazu. Mogą one uzyskać większą rozdzielczość; *kineskop Trinitron* - został

skonstruowany dużo później przez firmę Sony. Podstawową różnicą między nim a „Delta” jest inna konstrukcja maskownicy. Tworzą ją cienkie, czarne pionowo rozpięte, metalowe druciki grubości 0,1 mm. Dzięki takiemu rozwiązaniu wyświetlane na ekranie punkty mają kształt prostokątny, co zapewnia większy kontrast i ostrość oraz lepszą geometrię obrazu. Dodatkową zaletą tego kineskopu jest fakt, iż jest on wycinkiem walca (a nie kuli, jak w przypadku kineskopów Delta), co w efekcie sprawia, że ekran jest bardziej płaski, przez co zniekształcenia geometryczne obrazu są mniejsze, a także posiadają lepsze właściwości przeciwodblaskowe.

#### Podstawowe parametry:

Jednym z najważniejszych parametrów monitora, określającym jego rzeczywistą wartość, jest *rozdzielczość* z jaką może on wyświetlać obraz. Jest ona ściśle związana z maksymalną częstotliwością *odchylania poziomego* i *pionowego* monitora, dlatego więc te parametry odgrywają największą rolę. *Częstotliwość odchylania poziomego* określa prędkość, z jaką strumień elektronów wyświetla jedną linię poziomą na ekranie (stanowi ona odwrotność czasu jaki upływa na narysowanie jednego punktu). Natomiast *częstotliwość odchylania pionowego* (odświeżania obrazu) określa liczbę kompletnych ekranów, które monitor jest w stanie wyświetlić w czasie 1 s. Im obie powyższe częstotliwości są większe, tym rozdzielczość monitora może ulec zwiększeniu. Należy jednak pamiętać aby obraz (o danej rozdzielczości) był wyświetlany z odpowiednią częstotliwością odświeżania. Luminofor, zastosowany w kineskopach świeci tylko krótką chwilę - gdy wiązka elektronów przestaje padać na dany punkt, ulega on wygaszeniu. Jeśli częstotliwość odświeżania jest zbyt niska, możemy zauważyć zjawisko migotania obrazu, które jest bardzo męczące i szkodliwe dla oczu. W celu zapewnienia odpowiedniej stabilności obrazu strumień musi w odpowiednio krótkich odstępach czasu przebiegać przez całą powierzchnię ekranu. W nowoczesnych monitorach częstotliwość odświeżania nie powinna być mniejsza niż 75 Hz. W monitorach, w których częstotliwość odświeżania i pasmo przepustowe były za małe, w celu uzyskania obrazu o wysokiej rozdzielczości zastosowano dosyć chytrą „sztuczkę”. Metoda wyświetlania obrazu polega na tym, że najpierw są rysowane linie o numerach parzystych (rysowana jest co druga linia obrazu), a później są dorysowywane ważne linie o numerach nieparzystych. Można zatem powiedzieć, iż obraz podzielony jest na dwie części (jedna zawiera linie parzyste, druga - nieparzyste) i są one wyświetlane na przemian. Wtedy rzeczywista częstotliwości odchylania pionowego jest dwukrotnie mniejsza niż podawana wartość nominalna. Taki rodzaj wyświetlania nosi nazwę „*wybierania międzyliniowego*” lub „*wyświetlania z przeplo-*tem” (ang. *interlace*) i pozwala stosować monitory o dwukrotnie niższej częstotliwości odchylania pionowego. W normalnym trybie wyświetlania obrazu, tzn. przy zastosowaniu wybierania kolejno liniowego (ang. *non-interlaced*), monitor tworzy obraz rysując po kolei linię po linii.

Zupełnie innym rodzajem monitorów są wyświetlacze LCD. Wyświetlacz składa się z filtrów polaryzujących (które polaryzują wiązkę światła w określonym kierunku) dwóch kawałków szkła z nadrukowanymi od wewnątrz elektrodami. Pomiędzy szklanymi powłokami znajduje się warstwa orientująca, pod którą znajduje się ciekły kryształ i kolorowe filtry RGB umożli-



wiające uzyskanie kolorów. Stałą odległość pomiędzy dwiema warstwami szkła utrzymuje izolator. Ciekły kryształ jest substancją organiczną o ciekłej formie i krystalicznej strukturze molekularnej. Jego cząsteczki w kształcie podłużnych walców mogą zmieniać swoją pozycję, ale dzięki siłom wzajemnego oddziaływania są one ułożone równolegle względem siebie. Do sterowania nimi używane jest pole elektryczne. W większości dzisiejszych wyświetlaczy LCD wykorzystuje się typ ciekłych kryształów zwany *skręconym ciekłym kryształem nematycznym*. W celu skręcenia molekuł w spiralny lub skręcony wzór, na wewnętrzne powierzchnie szkła nanosi się warstwę zestrzajającą. Źródłem światła jest lampa fluorescencyjna umieszczona za wyświetlaczem. Gdy promień świetlny pada na wyświetlacz, w pierwszej kolejności przechodzi przez filtr polaryzujący a następnie w postaci spolaryzowanej wpada do warstwy ciekłego kryształu, gdzie przechodzi przez molekuły skręcające, na całej szerokości warstwy ciekłego kryształu. W wyniku tego zjawiska polaryzacja światła ulega zmianie o kąt równy kątowi skręcenia molekuł ciekłego kryształu, który wynosi 90 stopni i przechodzi przez kolejny filtr polaryzujący obrócony również o 90 stopni w stosunku do poprzedniego. Jeżeli do elektrod, umieszczonych na powierzchni warstwy ciekłego kryształu, przyłożone zostanie napięcie, walcowate molekuły ustawiają się w kierunku pola (prostopadle do warstwy) i zapobiegają polaryzacji światła o kąt 90 stopni, a co za tym idzie również przejściu przez drugi filtr polaryzujący. W celu uzyskania kolorów stosuje się serię kolorowych filtrów RGB przez które przechodzą wiązki światła po przejściu przez drugą warstwę polaryzującą. Każdy piksel obrazu tworzony jest podobnie jak w monitorach CRT, za pomocą trzech punktów odpowiadających trzem barwom podstawowym RGB. W pasywnej matrycy LCD każdy piksel obrazu musi być nieustannie pobudzany sygnałem, ponieważ molekuły ciekłego kryształu szybko powracają do stanu skręcenia. W matrycach aktywnych za pobudzenie odpowiada dodatkowy tranzystor pełniący funkcję bufora zapamiętującego wartość sygnału sterującego. Poprawił się dzięki temu kontrast, na lepsze zmieniły się też szybkość reakcji na impuls (czas odpowiedzi matrycy) i jakość odwzorowania kolorów. Pochodzący od nazwy cienkiej warstwy z tranzystorami - *Thin Film Transistor* - skrót TFT został powszechnie zaakceptowany jako nazwa wyświetlacza LCD z matrycą aktywną. W tradycyjnych wyświetlaczach LCD - nawet TFT - kontrast obrazu maleje wraz ze wzrostem kąta, pod którym obserwator patrzy na ekran. Zmieniają się także kolory obrazu, ponieważ oko ludzkie z różną intensywnością reaguje na zmiany kontrastu każdej z barw podstawowych RGB. Aby zminimalizować tego rodzaju wady wyświetlaczy, opracowano nowy typ aktywnej matrycy TFT nazwanej IPS-TFT (*In-Plane-Switching TFT*). W matrycy tej molekuły ciekłego kryształu ulegają skręceniu jedynie w płaszczyźnie poziomej, równoległej do powierzchni wyświetlacza, co znacznie zwiększa kąt patrzenia na obraz. Jedną z nielicznych cech przemawiających przeciw monitorom LCD jest ich *wysoka cena*. Przekracza ona średnio pięciokrotnie cenę monitora CRT o tych samych parametrach obrazu. Wynika to z bardzo skomplikowanego procesu wytwarzania matrycy i bardzo wysokiego współczynnika wadliwych egzemplarzy. Każdy punkt obrazu wyświetlany jest za pomocą trzech elementów wyposażonych w miniaturowe tranzystory. Dla rozdzielczości

1024x768 potrzeba ich aż 2 359 296, więc wyprodukowanie w 100% sprawnej matrycy jest zadaniem bardzo trudnym. W związku z tym producenci przyjęli, że liczba wadliwych tranzystorów nie może przekraczać sześciu. W praktyce jednak najczęściej spotykamy modele z 3 i mniej uszkodzonymi tranzystorami. Na obniżenie kosztów wpłynąć może bardzo obiecująca technologia sklejanego dużych ekranów LCD z tanich, małych paneli TFT o przekątnej zaledwie kilku cali. Gdy granice między poszczególnymi „klockami” nie będą widoczne, możliwe będzie wyprodukowanie monitora LCD, którego cena będzie porównywalna z monitorem CRT o takiej samej przekątnej. Kolejną wadą paneli LCD jest *niedokładne skalowanie obrazu*. Na ogół najwyższą jakość obrazu uzyskujemy tylko dla rozdzielczości będącej maksymalną rozdzielczością danego panelu. Jeżeli chcemy przeskalować obraz z rozdzielczości 1600x1200 na 800x600 nie jest to jeszcze problem, wystarczy jedynie podwoić wysokość i szerokość każdego piksela. Problemy zaczynają się przy przeskalowaniu do rozdzielczości np. 1024x768 lub 1280x1024. Tu stosunek liczby pikseli w pionie i poziomie wynosi 1,56 i 1,25 więc za podejmowanie decyzji, które piksele mają być przeskalowane odpowiada układ elektroniczny monitora. Aby skalowany tekst był czytelny układ stosuje metodę cieniowania, polegającą na wyświetlaniu każdej litery za pomocą dwóch lub więcej odcieni koloru. Wzorce opracowywane są oddzielnie dla każdej litery. Dla grafiki stosuje się algorytm krzywych s wygładzających krawędzie wyświetlanych obiektów. Innym sposobem obejścia tego ograniczenia jest tzw. centrowanie. Centrowanie polega na tym, iż obraz o mniejszej rozdzielczości umieszczany jest pośrodku ekranu, a pozostałą przestrzeń zajmuje czarna ramka. Dzięki temu nie dochodzi do zniekształcenia wyświetlanego obrazu. Z drugiej strony, stosując tę technikę nie wykorzystujemy całego ekranu i możemy mieć trudności z odczytaniem tekstu (gdy np. wyświetlamy obraz 320x200 na ekranie LCD o rozdzielczości 1024x768 punktów). Bardzo ważnym parametrem charakteryzującym panele LCD jest także *czas odpowiedzi matrycy*. Jest on zdefiniowany jako maksymalne opóźnienie pomiędzy sygnałem pobudzającym daną komórkę matrycy a momentem ustalenia się jej stanu (czas ten jest w granicach kilkudziesięciu milisekund). W modelach charakteryzujących się długim czasem odpowiedzi można zaobserwować efekt rozmazywania się obrazu podczas wyświetlania szybko przesuwających się scen. Zakłócenia widać np. podczas dynamicznych scen filmów oraz gier komputerowych. Może to doprowadzić do szybkiego męczenia się oczu i znacznego spadku komfortu pracy. Efekt rozmazywania obrazu zaobserwować można dla wyświetlaczy o czasie reakcji powyżej 50 milisekund. Wiele starszych monitorów LCD ma problemy z prawidłowym odwzorowaniem kolorów. Powodem jest niska rozdzielczość wbudowanych w monitory przetworników analogowo-cyfrowych. Tanie modele mają 6-bitowe przetworniki dla każdego z podstawowych kolorów, co umożliwia uzyskanie jedynie 262 144 kolorów ( $2^{18}$ ). Choć dla wielu użytkowników komputerów różnica jest niewidoczna, to tak duża redukcja przestrzeni barw całkowicie uniemożliwia pracę grafikom. Na szczęście wszystkie monitory z cyfrowymi interfejsami, a także najnowsze monitory z interfejsami analogowymi wyświetlają na ekranie pełną, 24-bitową paletę barw (16 777 216 odcieni).

Kąt widzenia. W zależności od położenia molekuł względem płaszczyzny ekranu zmienia się poziom kontrastu obrazu. Ponieważ oko ludzkie z różną intensywnością postrzega zmiany kontrastu podstawowych składowych barw RGB, wraz ze wzrostem kąta widzenia zmianie ulegają kolory. Przy dużym kącie obraz staje się nieczytelny a w skrajnych przypadkach – jeżeli kontrast spadnie poniżej progu czułości oka ludzkiego, niewidoczny. Radą na to jest zwiększenie podświetlenia ekranu jednak ma to znaczący wpływ na żywotność elementów wyświetlacza. W produkowanych obecnie panelach zakres dobrej widoczności waha się od 100 do 170 stopni w poziomie i 50 do 170 stopni w pionie. Dodatkowo opracowano kilka technologii mających na celu polepszenie kąta widzenia. Są to np. IPS (*In-Plane Switching*), MVAN, TN, FSTN. Wszystkie w/w technologie polegają na odpowiednim sterowaniu ułożenia molekuł ciekłokrystalicznych za pomocą warstwy sterującej. W najpopularniejszej obecnie technologii TN (*Twisted Nematic* - skręcone ciekłe kryształy nematyczne) cząstki ciekłego kryształu stopniowo skręcają się względem siebie o 90 stopni pomiędzy płaszczyznami ekranu dzięki specjalnej obróbce wewnętrznych powierzchni ekranu i leżą równoległe do tych płaszczyzn. Przyłożenie napięcia powoduje obrót cząstek ciekłego kryształu do pozycji prostopadłej do płaszczyzn elektrod umieszczonych na przeciwległych ściankach ekranu i zablokowanie światła. W wyświetlaczu takim światło skręcające swoją płaszczyznę polaryzacji podczas przejścia przez warstwę ciekłego kryształu widoczne jest pod małym kątem. Technologia TN ma także odmianę nazywaną STN, która charakteryzuje się kątem skręcenia od 200 do 270 stopni. Takie wyświetlacze są łatwiejsze do sterowania, ponieważ uzyskuje się zaciemnienie od 10 do 90 procent przy różnicy napięć tylko 0,2 V (normalnie ok 1,5 V). Wadą jest dodatkowy efekt chromatyczny związany z podwójnym załamaniem światła, co wymusza stosowanie dodatkowych filtrów - po uzupełnieniu wyświetlacza o specjalną warstwę uzyskujemy ekran FSTN (*Film Super Twisted Nematic*). Ekran z matrycą aktywną TFT-IPS (*In Plane Switching*) ma elektrody umieszczone tylko na tylnej ściance, a cząstki LC nie są skręcone względem siebie w stanie neutralnym. Przyłożenie napięcia powoduje, że drobiny ciekłego kryształu skręcają się w jednej tylko płaszczyźnie i pozostają równoległe do przodu i tyłu ekranu. Można obrazowo powiedzieć, że „przewodzenie” światła odbywa się wzdłuż krótszych brzegów molekuł i dzięki temu obraz widoczny jest pod większym kątem. W najnowszej technologii Super-IPS wprowadzono elektrody łamane, dzięki którym istotnie zmniejszyły się przebarwienia dla dużych kątów obserwacji oraz zwiększył się kontrast obrazu. Technika VAN (*Vertically Aligned Nematic*) pozwala na zwiększanie kąta widzenia i przyspieszenie reakcji kryształu dzięki skośnemu ustawieniu molekuł w stosunku do płaszczyzny ekranu, ale w zamian powoduje duże zmiany obrazu przy zmianach kąta obserwacji ekranu. Skośne ustawienie uzyskuje się dzięki zastosowaniu specjalnych roztworów poliamidowych, w których cząstki ciekłego kryształu mogą być ustawiane praktycznie w dowolnych kierunkach. W technologii MVAN (*Multidomain VAN*) grupy molekuł są odpowiednio skręcone względem siebie, tak aby uzyskać identyczny obraz z każdego kąta obserwacji.

Cyfrowy czy analogowy? Poza niedoskonałościami skalowania, źródłem utraty jakości obrazu w monitorach LCD może być konwersja analogowo-cyfrowa. Chociaż wyświetlacze są urządzeniami cyfrowymi, większość z nich wyposażono w standardowe analogowe złącze D-Sub. Główną przyczyną wykorzystywania analogowego toru do przesyłania sygnału z karty graficznej do monitora jest brak jednolitego standardu i znikoma liczba kart ze złączem cyfrowym. Dlatego też producenci monitorów LCD z wyjściem cyfrowym dostarczają zwykle zestawy zawierające odpowiednią kartę graficzną. Cechą charakterystyczną monitorów ze złączem cyfrowym jest brak systemu OSD, a co za tym idzie, brak możliwości regulacji parametrów obrazu. Nie jest to jednak potrzebne, gdyż przesyłany cyfrowo obraz wolny jest od zakłóceń synchronizacji czy przekłamań kolorów. Ewentualnej korekty kontrastu lub nasycenia barw można dokonać za pomocą opcji dostępnych z poziomu sterowników.

#### **4.2. Zasada działania karty graficznej.**

Monitor wymaga zawsze karty grafiki przetwarzającej informacje z płyty głównej na analogowe, bądź cyfrowe (w zależności od typu karty) sygnały, które służą z kolei monitorowi do budowania obrazu. W niektórych, zwłaszcza tańszych, kompaktowych modelach PC, karta grafiki zostaje od razu zintegrowana z płytą główną. Z wielu względów nie poleca się tego rozwiązania. Po pierwsze, zdarza się często, że kiedy wzrosną nasze wymagania co do grafiki, to okazuje się, że wyłączenie takiej, zintegrowanej z płytą główną, karty oraz instalacja nowej jest po prostu niemożliwe (albo bardzo skomplikowane). Po drugie, takie „oszczędnościowe” pecety nie posiadają często odpowiednich slotów (ISA, PCI, AGP), co zupełnie uniemożliwia instalację nie tylko karty grafiki ale i jakiegokolwiek innej. Po trzecie, rzadko kiedy daje się w tych modelach poszerzyć pamięć video i wreszcie po czwarte, o wiele trudniej zdobyć uaktualnione sterowniki.

Pamięć video, zwana również pamięcią grafiki, wymagana jest przez każdą obecnie produkowaną kartę grafiki. Odpowiedzialna jest ona, nieco upraszczając, za to, że obraz może być w ogóle odbierany jako nieruchomy. Nawet jeśli tego nie dostrzegamy, obraz na monitorze „budowany” bywa na nowo przynajmniej 50 razy w ciągu sekundy. Spokojny obraz uzyskiwany bywa z częstością 75 obrazów na sekundę (75 Hz), co umożliwiły dopiero karty i monitory VGA.

Pamięć video adresowana bywa podobnie jak inne rodzaje pamięci peceta. Zawiera ona cyfrowy zapis wyświetlanego aktualnie obrazu. Na karcie grafiki znajduje się specjalny chip - kontroler grafiki - komunikujący się z resztą systemu poprzez adresy I/O, wykorzystując do sterowania obrazem specjalny rejestr. Począwszy od kart VGA wzwyż, do każdego kontrolera grafiki dołączony jest tzw. DAC (Digital Analog Converter), który przetwarza cyfrowe informacje na sygnały analogowe sterujące obrazem monitora (VGA).

W czasie krótkiej historii PC zastosowanie znalazł cały szereg różnego rodzaju kart do odtwarzania obrazu. Nie wszystkie były zdolne do zastosowań graficznych, tym niemniej, dla potrzeb poniższych rozważań, nazywać je będziemy zawsze kartami grafiki, by w ten sposób

podkreślić, że sterowały one pracą monitora. Te dwa komponenty: monitor i karta grafiki tworzą zawsze pewną całość (system graficzny) - ich parametry muszą zostać wzajemnie dopasowane. W przeciwnym wypadku dojść może do uszkodzenia jednego z nich.

Z dzisiejszego punktu widzenia, można by się skoncentrować wyłącznie na kartach graficznych VGA z minimum 1MB pamięci przy minimalnej rozdzielczości 640 x 480 pikseli i standardowym interfejsem PCI. Jednakże karty te miały swoje poprzedniczki, na które się od czasu do czasu natykamy i które mogą jeszcze z powodzeniem zostać sensownie wykorzystane. Już choćby z tego powodu warto co nieco o nich wiedzieć. Jeszcze bardziej atrakcyjne finansowo od starszych kart grafiki są starsze monitory. Czy jednak taki monitor będzie działał wraz z posiadaną przeze mnie kartą grafiki? Oto pytanie, którym m.in. zajmiemy się w następnych punktach.

### **4.3. Przegląd kart graficznych.**

Pierwszym sterownikiem instalowanym w oryginalnych komputerach IBM PC był sterownik MDA (Monochrome Display Adapter), który pracował wyłącznie w trybie tekstowym, umożliwiając wyświetlenie 80 znaków w 25 wierszach w dwóch kolorach, czarnym i białym (właściwie w czarnym i jasnozielonym, gdyż taki luminofor miały ówczesne monitory). Później pojawił się -opracowany przez firmę IBM - sterownik CGA (Colour Graphics Adapter), umożliwiający wyświetlanie obrazu o rozdzielczości 320 na 200 punktów w czterech kolorach lub obrazu o rozdzielczości 640 na 200 punktów w dwóch kolorach. Mimo „oszałamiającej” grafiki sterownik został nisko oceniony, gdyż w trybie tekstowym znaki miały rozmiar tylko 8 na 8 punktów. Wkrótce na rynku pojawił się sterownik o lepszych parametrach graficznych - Hercules. W trybie tekstowym miał te same parametry, co sterownik MDA, ale umożliwiał także wyświetlanie dwukolorowego obrazu graficznego o rozdzielczości 720 na 350 punktów. Dzięki niskiej cenie i wysokich jak na owe czasy parametrach zdobył on ogromną popularność i przez długi czas był stosowany wszędzie tam, gdzie kolorowy obraz nie był konieczny.

W odpowiedzi na sterownik Hercules, firma IBM (która nigdy de facto nie uznała tego standardu) opracowała sterownik EGA (Enhanced Graphics Adapter), o rozdzielczości 640 na 350 punktów i możliwości jednoczesnego wyświetlenia 16 kolorów spośród 64 możliwych. Potomkiem karty EGA w prostej linii jest sterownik VGA (Video Graphic Array) i jako taki jest zgodny ze swoimi poprzednikami. Jest to jego ogromną zaletą, gdyż większość programów działających ze sterownikami MDA i CGA będzie działała prawie ze sterownikiem VGA. Sterowniki VGA niektórych producentów umożliwiają także emulację karty Hercules ale nie jest to opcja występująca w standardzie VGA. Oprócz wymienionych sterowników opracowano kilka innych kart graficznych, między innymi przeznaczony dla systemu OS/2 sterownik MCGA (Multi-Colour Graphics Array) oraz sterownik PGA (Professional Graphics Adapter), stanowiący ogniwo pośrednie pomiędzy kartą EGA i VGA.

Standardowy sterownik VGA umożliwia wyświetlenie na ekranie 25 wierszy znaków w 80 kolumnach. W tym trybie znaki mają rozmiar 9 na 16 punktów. Znak może być wyświetlony

w jednym z 16 kolorów, niezależnie od koloru tła, które dla każdego znaku może być dobrane oddzielnie. Przy wykorzystaniu wszystkich kolorów znaków i tła i ustawieniu różnych atrybutów (rozjaśnienie, migotanie) można jednocześnie wyświetlić 256 kombinacji znaków i tła. W pewnych przypadkach możliwe jest podwojenie liczby wierszy tekstu do 50, ale dzieje się to kosztem zmniejszenia rozmiarów znaku do 8 na 8 punktów. Można także zwiększyć liczbę wyświetlanych kolumn. W trybie graficznym standardowy sterownik VGA umożliwia wyświetlenie 640 punktów w 480 liniach przy 16 dostępnych kolorach (w jednym ze standardowych trybów możliwe jest wyświetlenie jednocześnie 256 kolorów, ale przy rozdzielczości 320 na 200). Jeśli sterownik ma odpowiednio dużo pamięci, przez właściwe ustawienie rejestrów można uzyskać rozdzielczość nawet 1 024 punktów w 768 liniach (przy odpowiedniej ilości pamięci nawet w 256 kolorach jednocześnie). Niezależnie od trybu graficznego wszystkie wyświetlane kolory mogą być wybrane z palety 262 144 barw. Jednym z elementów różniących sterownik VGA od poprzedników jest sposób generowania sygnału wyjściowego. Sterownik VGA generuje sygnał analogowy, dzięki czemu do standardowej karty VGA można podłączyć zarówno monitor kolorowy, jak i monochromatyczny, który zamiast kolorów będzie wyświetlał 64 poziomy szarości. Istotną cechą sterownika VGA jest możliwość zamontowania go w jednym komputerze razem ze sterownikiem monochromatycznym, na przykład kartą Hercules. W ten sposób stworzony został pierwowzór tak oczywistych w dzisiejszych czasach systemów Dual-Monitor. Ułatwiały one znacznie pracę programistom posługującym się obiektami graficznymi. Możliwość pracy z dwoma monitorami była stopniowo akceptowana przez większość oprogramowania narzędziowego i graficznego (Borland, Autocad), a obecnie również przez same systemy operacyjne. W miarę rozwoju technologii parametry standardowego VGA stały się niewystarczające i wielu producentów zaczęło swoje produkty wyposażać w dodatkowe opcje. Rozbudowa obejmowała głównie zwiększenie pamięci karty graficznej (standardowy sterownik VGA miał 256 kB pamięci) oraz dodanie nowych trybów graficznych (1 024x768, 800x600, 1 280x1 024). Oprócz rozdzielczości zwiększeniu uległa także liczba jednocześnie wyświetlanych kolorów. Powstały karty umożliwiające jednoczesne wyświetlenie ponad 60 tysięcy kolorów lub wyboru koloru z palety 16 milionów odcieni. Nigdy nie powstał jednolity standard kart SVGA (Super VGA; tak nazywane są karty VGA mające rozbudowane możliwości). Główni producenci procesorów graficznych - kart i chipów graficznych - prześcigali się w dodawaniu nowych możliwości do swoich produktów. Niestandardowe tryby pracy, w zależności od użytego procesora graficznego, miały różne numery, pamięć zorganizowana była w różny sposób i różne były adresy niestandardowych rejestrów.

Próby zapanowania nad tym chaosem doprowadziły do powstania stowarzyszenia producentów urządzeń graficznych o nazwie VESA (Video Electronics Standards Association). Postawiło sobie ono za cel opracowywanie norm umożliwiających standaryzację:

- parametrów monitorów współpracujących z kartami SVGA,
- numeracji rozszerzonych trybów graficznych,

- interfejsu programowego (dodatkowych funkcji BIOS-u).

Sterowniki współczesnych kart SVGA obsługują sprzętowo założenia standardu VESA - dla starszych typów konieczne było stosowanie programowych emulatorów, opracowanych specjalnie dla danej karty.

#### 4.4. Karta VGA.

Nowy standard firmy IBM - Personal System 2-Computer (PS/2) z nową architekturą (Micro-channel) wyposażono w nową kartę grafiki VGA. Karta ta może emulować wszystkie inne standardy, przez co nadaje się zwłaszcza do rozbudowy starszych PC. VGA to skrót od *Video Graphics Array* i oznacza specjalny, wysoko zintegrowany układ (*GateArray*) zastępujący i jednoczący funkcje wielu pojedynczych, starszych komponentów, które do tej pory odpowiadały za grafikę w pececie. Karty VGA zaczęto wkrótce produkować również w architekturze ISA, co stało się standardem na wiele lat. Na karcie VGA znajduje się, podobnie jak na karcie EGA, specjalna kość pamięci z własnym BIOS-em (zachowującym kompatybilność z rejestrem EGA). Obszar adresowania BIOS-a rozpoczyna się również od C0000h ale sięga aż do C7FFFh. Adresowanie pamięci graficznej rozpoczyna się, podobnie jak w karcie EGA, od adresu A0000h. Karta VGA wymaga minimum 256 kB pamięci, zajmującej w wersji minimalnej ten sam obszar co karta EGA. Poza tym spotyka się karty z 512 kB lub 1MB pamięci video. Ponieważ konstruktorzy IBM-PC przewidzieli dla karty grafiki obszar adresowania od A0000h do BFFFFh, a więc tylko 128 kB, pamięć video musi zostać stronicowana. Odbywa się to na podobnej zasadzie jak dostęp pamięci Expanded do obszaru adresowania peceta. Zawartość pamięci przenoszona jest kawałkami do 128 kB-owego obszaru adresowania. Rezultatem tej procedury, po raz pierwszy zastosowanej w karcie EGA, jest niestety znaczne spowolnienie tempa powstawania obrazu.

Matryca znaku	9x16 pikseli w trybie VGA
Pamięć	Zwykle 256 kB - 1 MB
Obszar adresowania	A0000h - BFFFFh
BIOS grafiki	C0000h - C1FFFh

*Tabela 2. Parametry techniczne karty VGA.*

Cechą charakterystyczną karty VGA jest zastosowane po raz pierwszy analogowe sterowanie monitorem. Sygnał analogowy jest znacznie bogatszy w szczegółowe informacje, co poprawia zdecydowanie jakość obrazu. Monitory z cyfrowym wejściem, jak np. w standardzie EGA, nie potrafią oczywiście współpracować z analogowym sygnałem karty VGA (odnosi się to również do odwrotnej konfiguracji: monitor VGA, karta cyfrowa np. EGA.). Ewentualne nieprawidłowe połączenie może doprowadzić do uszkodzenia monitora lub karty grafiki. Karta EGA potrafi jednocześnie wyświetlić 16 kolorów. Wiązki elektronów odpowiedzialne za barwy podstawowe (czerwony, zielony, niebieski) sterowane są cyfrowo (tzn. możliwy jest jeden z dwóch stanów: kolor włączony lub wyłączony). Umożliwia to jednoczesne przedstawienie 8 kolorów, a biorąc pod uwagę dodatkowy sygnał intensywności - 16 kolorów.

Rozdzielczość graficzna	Rozdzielczość tekstowa	Ilość kolorów	Tryb
320 x 200	40 x 25	2	CGA
320 x 200	40 x 25	16	CGA
640 x 200	80 x 25	2	CGA
640 x 200	80 x 25	16	CGA
720 x 348	80 x 25	2	HGC, MDA

Tabela 3. Tryb tekstowy.

Informacje o kolorach zostają przetworzone przez znajdujący się na karcie VGA cyfrowo-analogowy konwerter (ADC - *Analog Digital Converter*). Każdy kolor to mieszanka trzech barw podstawowych o 64 stopniach intensywności, co daje teoretycznie:  $64^3 = 262144$  kolory. Oczywiście monitor musi być przeznaczony do pracy z sygnałem analogowym. Jeśli zamiast monitora kolorowego VGA zastosujemy monitor czarno-biały (ale również VGA), to informacje o kolorach zostaną tak przetworzone (za pomocą specjalnej formuły w BIOS-ie), że w efekcie otrzymamy 64 stopnie szarości. Musimy jedynie uważać na monitor, nie każdy monochromatyczny pracuje w tym trybie.

Rozdzielczość graficzna	Rozdzielczość tekstowa	Ilość kolorów	Tryb
320 x 200	40 x 25	2	CGA
320 x 200	40 x 25	4	CGA
640 x 200	80 x 25	2	CGA
320 x 200	40 x 25	16	EGA
640 x 200	80 x 25	16	EGA
640 x 350	80 x 25	2	EGA
640 x 350	80 x 25	16 z 64	EGA
320 x 200	80 x 25	256	VGA
640 x 480	80 x 25	2 z 16	VGA
640 x 480	80 x 25	16 z 256	VGA
640 x 480	80 x 25	256 z 64k	VGA

Tabela 4. Tryb graficzny.

W trybie graficznym jak widać (zob. tab. 4) możliwe jest jednoczesne przedstawienie 256 kolorów w rozdzielczości 640 x 480 pikseli. W takim jednak przypadku pamięć video musi wynosić minimum 512 kB. Dysponując minimalną dla standardu VGA pamięcią (256 kB), możemy w rozdzielczości 640 x 480 przedstawić tylko 16 kolorów. Matryca znaku w trybie tekstowym składa się z 9 x 16 pikseli. Większość kart VGA potrafi jednak jeszcze więcej. Wyższe rozdzielczości (jak 800 x 600 lub 1024 x 768) określa się często mianem Super-VGA, VGA-Deluxe lub Extended-VGA. Ważne jest przy tym by podłączony monitor również potrafił wyświetlić obraz w



takich rozdzielczościach oraz aby karta wyposażona była w odpowiedni sterownik. W DOS-ie, bez specjalnego sterownika, możliwy jest tylko tryb standardowy VGA (640 x 480). Standardowa karta VGA potrafi wszystko to, co już potrafiła karta EGA, a więc emulację wszystkich



Rys. 4.1. Standardowe 15-biegunowe gniazdko DSUB karty VGA.

innych rozdzielczości oraz odpowiednie przełączanie matrycy znaków. W dokumentacji technicznej niektórych kart grafiki znaleźć można często oznaczenia typu APA lub A/N. APA oznacza tryb graficzny (A/I Points Addressable), a A/N tryb alfanumeryczny (tekstowy). Karta VGA doczekała się również nowego łącza - 15-biegunowego gniazdka. Spotykane się je w dwóch różnych wariantach, jak to np. widoczne jest na rys. 4.2. W pecetach rozpowszechniło się gniazdko Mini-DSUB, podczas gdy w komputerach Macintosh

firmy Apple zastosowanie znalazł wariant nieco większy. Przyporządkowanie sygnałów jest w każdej wersji inne, ale nie ma to żadnego znaczenia, o ile do monitora dołączono odpowiedni kabel połączeniowy (co powinno być regułą). W czasie przejściowym pomiędzy kartą VGA i EGA w użyciu znajdowało się również 9-biegunowe łącze VGA, co było jednak dosyć problematyczne ze względu na możliwość pomylenia z cyfrowym łączem EGA. Niektóre karty posiadają nawet oba gniazda, przy czym 9-biegunowe (cyfrowe) przeznaczone jest do pracy w trybie EGA, a 15-biegunowe (analogowe) do pracy w trybie VGA. Na kartach takich znajduje się specjalny przełącznik (DIP), który zmienia tryb pracy karty i aktywizuje odpowiednie łącze.

#### **4.5. Inne rodzaje kart.**

##### **Karta AGA.**

Firma Commodore, przed pojawieniem się kart VGA, stosowała w swoich pecetach własne rozwiązanie: tzw. kartę AGA (*Advanced Graphics Adapter*). Karta AGA to kombinacja trybu CGA i Hercules. W trybie CGA zachowuje się ona jak karta CGA, z tą jednak różnicą, że potrafi wyświetlić w obu rozdzielczościach (tzn. 320 x 200 oraz 640 x 200) 16 kolorów na raz. Jednocześnie potrafi pracować w monochromatycznym trybie karty Hercules (720 x 348 -grafika, i 80 znaków x 25 wierszy w trybie tekstowym) oferując dodatkowo dwa dalsze tryby (tekstowe): 132 znaki x 25 wierszy oraz 132 znaki x 44 wiersze.

##### **Karta PGA.**

W roku 1984, a więc w rok po pojawieniu się karty EGA, firma IBM wprowadziła na rynek kartę PGA (*Professional Graphics Adapter*). Zastosowano w niej po raz pierwszy analogową transmisję sygnałów z karty grafiki do monitora. Umożliwiło to jednoczesne wyświetlanie 256 kolorów z palety 4096 barw, przy rozdzielczości 640 x 480 pikseli. Karta PGA wymagała już specjalnego procesora graficznego, przez co była stosunkowo droga i nie zdobyła szerszej popularności. Karty PGA, podobnie jak AGA nie są już produkowane. Zastąpił je standard VGA.

##### **Karta MCGA.**

Komputery firmy IBM Personal System 2-Computer (PS/2 modele 25 i 30) wyposażono w pracującą analogowo kartę o nazwie *Multi Color Graphics Adapter* (MCGA). Kartę tę pod względem wydajności porównać można z kartami CGA i EGA, gdyż podobnie jak tamte pracuje ona w rozdzielczości 320 x 200 (CGA) z 256 kolorami z palety 262144 barw. W trybie tekstowym oferuje rozdzielczość 80 znaków w 25 wierszach z matrycą znaku 8 x 16 pikseli. Żaden inny producent nie podjął się jej wytwarzania, z czasem została całkowicie zastąpiona kartami VGA. Karty MCGA nie należy mylić z kartą MGA, gdyż skrót ten - *Matrox Graphics Accelerator* (lub *Adapter*) oznacza kartę firmy Matrox np. Millenium albo Mystique.

### **Karta 8514/A**

Firma IBM pracowała nadal nad rozwojem standardu VGA. Jednym z tego efektów jest, przeznaczona pierwotnie dla modeli PS/2 zbudowanych w architekturze Microchannel, karta 8514/A. Wymaga ona specjalnego monitora, np. 8514, w którym możliwe jest osiągnięcie rozdzielczości 1024 x 768 pikseli, przy jednoczesnym odwzorowaniu 256 kolorów. Znajdujący się na tej karcie procesor graficzny odciążał znacznie procesor główny, co zwłaszcza w pracy z wymagającymi programami typu CAD miało niebagatelne znaczenie. Nie zadbano jednak o kompatybilność ze standardem VGA. Karty 8514/A dostępne są również w wersji ISA, nie zwiększyło to jednak ich popularności. IBM zastąpił z czasem ten standard adapterem XGA, co praktycznie oznaczało koniec jego stosowania. Niektóre dzisiejsze „akceleratory” graficzne, jak np. Graphics Ultra czy Graphics Pro Turbo firmy ATI, potrafią emulować za pomocą specjalnego sterownika tryb karty 8514, przez co niektóre starsze programy wymagające tego trybu pracować mogą również na nowszym sprzęcie.

### **Karta XGA.**

Karta XGA (*Extended Graphics Array*) cechuje się kompatybilnością zarówno ze standardem VGA jak i 8514/A. Szybkość przetwarzania danych w konfiguracji XGA jest szczególnie wysoka, a osiągnięta zostaje dzięki specjalnemu chipowi (Array). Komunikacja pomiędzy kartą grafiki i płytą główną odbywa się zazwyczaj poprzez porty I/O (np. 3C0h 3CFh w przypadku kart EGA i VGA). Inaczej rzecz się ma z kartą XGA. Porty I/O zostają odwzorowane w ściśle określonym obszarze pamięci operacyjnej komputera, a karcie grafiki przyporządkowany specjalny obszar tej pamięci. Przyspiesza to ogromnie dostęp do karty, a przez to prędkość powstawania obrazu. Dodatkowy wzrost wydajności i szybkości kart XGA związany jest z „inteligencją” chipa, który potrafi np. samodzielnie kreślić linie albo prostokąty na podstawie tylko kilku parametrów. Odwzorowane mogą zostać 1024 x 768 punkty w 256 kolorach albo 640 x 480 punkty w 65536 kolorach. Oprócz tego karta ta umożliwia wyświetlanie kursora składającego się z 64x64 punktów.

### **Standard TIGA.**

Firma Texas Instruments stworzyła własny, bazujący na specjalnym procesorze graficznym (TMS340XX), standard TIGA (Texas Instruments Graphics Architecture). TIGA to w zasadzie programowalny interfejs pomiędzy CPU, daną aplikacją i procesorem graficznym. TIGA dyspo-

nuje całą gamą instrukcji, które na życzenie mogą jeszcze zostać uzupełnione. Nowe instrukcje zostają, w zależności od potrzeb danego programu, załadowane przez procesor komputera do procesora graficznego. Karta grafiki TIGA dla swoich aplikacji nie potrzebuje więc wcale procesora głównego. Umożliwia to zbudowanie obrazu w niezwykle krótkim czasie, a więc dużą częstotliwość odświeżania przy możliwości jednoczesnego wyświetlania kilku obrazów (jeden na drugim). Wkrótce karty TIGA zaczęły być produkowane także przez inne firmy. Oprócz Texas Instruments np. Eizo, Hercules i Opta. Oryginalne karty TIGA firmy Texas Instruments noszą nazwy: TIGA 10, TIGA CARD, TIGA STAR i TIGA DIAMOND, a ich obszar zastosowania waha się od normalnych zastosowań w Windows, po profesjonalne programy tj. trójwymiarowe modelowanie, animacje komputerowe czy zastosowania multimedialne. Karty te kompatybilne są z trybem VGA, choć nie wszystkie w ten sam sposób. I tak na przykład karta TIGA 10 musi dopiero zostać połączona (poprzez tzw. Feature-Connector zwany też „VESA FeatureConnector”) z tradycyjną kartą VGA, podczas gdy inne karty mają już wbudowany specjalny chip kontrolera VGA. Windows, aż do wersji 3.1 współpracował ze standardem TIGA, jednakże nie zyskał on takiej popularności, jak się przez pewien czas wydawało. Winić za to należy niewątpliwie stosunkowo wysoką cenę tych kart w porównaniu do „akceleratorów” graficznych dla Windows, które w efekcie (niestety) wyparły karty TIGA prawie całkowicie. Niestety, bowiem karty TIGA wskazywały właściwą drogę, zwalniając procesor z funkcji graficznych - dobra jakościowo grafika wcale nie wymaga procesora Pentium, ma on „ważniejsze sprawy na głowie”. Dopiero wraz z AGP (Accelerated Graphics Port) sytuacja ta gruntownie uległa zmianie. Architektura ta pozwala bowiem karcie grafiki na bezpośredni dostęp do pamięci operacyjnej bez obciążania procesora transferem danych.

### **Tryby VESA**

Na rynku pojawiło się w międzyczasie całe mnóstwo kart grafiki wielu różnych producentów oferujących wyższe rozdzielczości i więcej możliwych kolorów niż w standardzie VGA. Nazywa się je często SuperVGA (SVGA, 800 x 600 pikseli) lub VGA-Deluxe itp., o czym już w poprzednich rozdziałach wspominaliśmy. Brak określonych standardów doprowadził w krótkim czasie do sporego zamieszania (zwłaszcza na rynku sterowników), zakończonego dopiero przez konsorcjum VESA (*Video Electronics Standard Association*). To zrzeszenie firm produkujących hardware i software dla grafiki i video, zdefiniowało jednolite wytyczne dotyczące m. in. interfejsu BIOS-u, adresowania pamięci video oraz znormalizowało używane częstotliwości synchronizacji poziomej i pionowej. Należy zwracać szczególną uwagę, aby, przy rozdzielczościach większych niż VGA, karta grafiki i monitor pracowały w możliwie wielu jednakowych trybach VESA. Często jeden z tych trybów to jedyna kombinacja, która zostaje automatycznie rozpoznana przez oba komponenty systemu graficznego (monitor i kartę grafiki), a więc jedyna, która może być używana. W zależności od rodzaju karty graficznej, BIOS grafiki samodzielnie rozpoznaje i udostępnia wymagany tryb VESA (zob. poniższa tabela). W przeciwnym przypad-

ku, przed wystartowaniem danego programu, załadowany musi zostać tzw. „rezydentny” sterownik odpowiedniego trybu VESA.

Tryb VESA	Kolory	Znaki x Wiersze	Piksele	Matryca znaku	Syn. pozioma	Syn. pionowa
10A (54h)	16/256	132 x 43	1056 x 350	8x8	31.5 kHz	70Hz
109 (55h)	16/256	132 x 25	1056 x 350	8x 14	31.5 kHz	70Hz
102 (58h)	16/256	100 x 37	800 x 600	8x 16	35.2 kHz	56Hz
102 (58h)	16/256	100 x 37	800 x 600	8x 16	37.8 kHz	60Hz
102 (58h)	16/256	100 x 37	800 x 600	8x 16	48.1 kHz	72Hz
103 (5Ch)	256/256	100 x 37	800 x 600	8x 16	35.2 kHz	56Hz
103 (5Ch)	256/256	100 x 37	800 x 600	8x 16	37.9 kHz	60Hz
103 (5Ch)	256/256	100 x 37	800 x 600	x 16	48.1 kHz	72Hz
104 (5Dh)	16/256	128 x 48	1024 x 768	8x 16	35.5 kHz	87Hz (i)
104 (5Dh)	16/256	128 x 48	1024 x 768	8x 16	48.3 kHz	60Hz
104 (5Dh)	16/256	128 x 48	1024 x 768	8x 16	56 kHz	70Hz
104 (5Dh)	16/256	128 x 48	1024 x 768	8x 16	58 kHz	72Hz
101 (5Fh)	56/256	80 x 30	640 x 480	8x 16	31.5 kHz	60Hz
101 (5Fh)	256/256	80 x 30	640 x 480	8x 16	37.9 kHz	72Hz
105 (60h)	56/256	128 x 48	1024 x 768	8x 16	35.5 kHz	87 Hz (i)
105 (60h)	256/256	128 x 48	1024 x 768	8x 16	48.3 kHz	60Hz
105 (60h)	256/256	128 x 48	1024 x 768	8x 16	56 kHz	70Hz
105 (60h)	256/256	128 x 48	1024 x 768	8x 16	58 kHz	72Hz
111 (64h)	64k/64k		640 x 480		31.5 kHz	60Hz
111 (64h)	64k/64k		640 x 480		37.9 kHz	72Hz
114 (65h)	64k/64k		800 x 600		35.2 kHz	56Hz
114 (65h)	64k/64k		800 x 600		37.8 kHz	60Hz
110 (66h)	32k/32k		640 x 480		31.5 kHz	60Hz
110 (66h)	32k/32k		640 x 480		37.9 kHz	72Hz
113 (67h)	32k/32k		800 x 600		31.5 kHz	56Hz
102 (6Ah)	16/256	100 x 37	800 x 600	8x 16	35.2 kHz	56Hz
102 (6Ah)	16/256	100 x 37	800 x 600	8x 16	37.8 kHz	60Hz
102 (6Ah)	16/256	100 x 37	800 x 600	8x 16	48.1 kHz	72Hz
106 (6Ch)	16/256		1280 x 1024		48 kHz	87Hz (i)
107 (6Dh)	256/256		1280 x 1024		48 kHz	87 Hz (i)
10E (6Fh)	64k/64k		320 x 200		31.5 kHz	70Hz
10F (70h)	16M/16m.		320 x 200		31.5 kHz	70Hz
112 (71hJ)	16M/16m.		640 x 480		31.5 kHz	60Hz
117 (74h)	64k/64k		1024 x 768		35.5 kHz	87 Hz (i)

Tabela 5. Niektóre z najbardziej rozpowszechnionych trybów VESA (standard oraz extended; i oznacza tryb interlaced – z przeplotem)

#### 4.6. Akceleratory graficzne.

Dla większości programów działających w środowisku DOS - z wyjątkiem niektórych gier - karty VGA oferują wystarczającą graficzną wydajność za przystępną cenę. Ich rozdzielczość i szybkość budowy obrazu są z reguły wystarczające. System operacyjny Windows ma już o

wiele większe wymagania jeśli chodzi o graficzne zdolności peceta. Przy każdym otwarciu okna, lub rozwinięciu tzw. Pull-Down-Menu, informacja graficzna musi zostać w całości na nowo przetworzona przez procesor. Przykładowo, zbudowanie wycinka obrazu o wielkości 100 x 100 pikseli - a nie jest to wcale dużo - wymaga przetworzenia informacji o 10.000 pojedynczych pikselach, a w wypadku jego przesunięcia nawet 20.000, tzn. wygaszenia punktów na starej pozycji i wyświetlenia ich na nowej. Powinno to odbywać się w miarę szybko, aby nie cierpiała na tym płynność pracy przy komputerze. Jak już wcześniej wspominaliśmy, jedno z możliwych rozwiązań, jak odciążyc procesor od funkcji graficznych, pokazały karty TIGA. Zastosowano tam specjalny procesor graficzny, przez co jednak karty te stały się stosunkowo drogie. Od roku 1992 na rynku obecne są tzw. akceleratory graficzne dla Windows (czyli karty przyspieszające budowę obrazu pod Windows), oparte na podobnej idei przewodniej tzn. odciążeniu procesora przez pewien rodzaj „własnej inteligencji”. Nie jest to jednak sposób tak konsekwentny, jak w przypadku kart TIGA. Nieznana do tej pory, założona w 1989 r, firma S3 Incorporated poszła nieco inną drogą, która w międzyczasie stała się standardem i praktykowana jest obecnie przez wszystkie bez wyjątku karty graficzne dla PC. Zamiast osobnego procesora graficznego z własnym katalogiem poleceń, zastosowano tzw. chip akceleratora, ze zintegrowanymi w nim funkcjami graficznymi, które Windows najczęściej wykorzystuje. Typowe funkcje akceleratora:

#### **BitBlit: Bit-Blit**

Przesunięcie zawartości okna (bloku pikseli). Funkcja BitBlit przesuwa prostokątny wycinek obrazu i wypełnia go pierwotną zawartością.

#### **Hardware-Cursor**

Ruchy kursora lub strzałki myszy sterowane są chipem akceleratora. CPU przetwarza tylko informacje o współrzędnych myszy.

#### **Line Drawing**

Rysowanie linii. CPU dostarcza wyłącznie informacje o współrzędnych początkowych i końcowych.

#### **Circle Drawing**

Rysowanie okręgów. Tylko środek okręgu i jego promień dostarczane są przez procesor.

#### **Polygon Fill**

Wypełnianie wieloboków. Chip akceleratora wypełnia wielokąty poszczególnymi pikselami.

Karty przyspieszające są kompatybilne ze standardem VGA. Oznacza to, że po włączeniu PC zachowują się one jak „normalne” karty VGA. Niektóre (starsze) karty potrzebują jednak, również dla środowiska DOS, specjalnego sterownika, który aktywizuje funkcje przyspieszające, co w nowszych kartach odbywa się automatycznie podczas inicjalizacji peceta w fazie POST (Power On Self Test). Akcelerator taki pokazuje swe właściwości przyspieszające dopiero podczas pracy z aplikacjami realizującymi złożone zadania graficzne. Wszystkie dostępne obecnie na rynku karty grafiki stosują zasadę graficznego „wspomagania”, przy czym w coraz większym

stopniu zastosowanie znajdują funkcje przestrzenne, co omówimy szerzej w następnym punkcie.

#### **4.7. Karty grafiki 3D.**

Trójwymiarowe odwzorowanie obrazu (3D) stawia przed kartą grafiki szczególnie wysokie wymagania. Z tego powodu wykorzystuje się do tego celu specjalne akceleratory 3D, które pracują na podobnej zasadzie jak chipy 2D, z tą jednak różnicą, że zamiast funkcji 2D wyposaża się je w mniej lub więcej, w zależności od typu, funkcji 3D. Obliczenia geometryczne pozostają w dalszym ciągu w gestii procesora głównego, a akcelerator 3D zajmuje się przeważnie operacją tzw. „Renderingu” i obliczaniem tekstur<sup>1</sup>. Nowsze karty grafiki posiadają oba chipy 2D i 3D, często zintegrowane w jednej obudowie. Obok tego istnieją karty „czysto” 3D, jak np. Monster 3D firmy Diamond, wykorzystywane równoległe z tradycyjną kartą 2D. Głównym kryterium oceny karty 3D jest oczywiście jakość obrazu. Zwłaszcza tzw. efekt „klocków” (okrągłe obiekty robią się wielokątne) nie powinien, o ile to możliwe, występować. Funkcja, która zapobiega jego powstawaniu zwie się w żargonie technicznym „bilinear filtering”. Powoduje ona, że tekstury zostają przedstawione bardzo miękko i płynnie. Niektóre karty - np. Matrox Mystique - nie dysponują jednakże tą funkcją. Dlatego wybierając kartę 3D powinniśmy zwrócić uwagę, czy i jakie funkcje zintegrowane zostały w chipie akceleratora. Zdarza się również, że niektóre funkcje nie są zintegrowane w chipie tylko emulowane przez specjalny program (sterownik). Odbywa się to oczywiście kosztem szybkości i jakości odwzorowania. Typowe funkcje akceleratorów 3D:

##### **Bi- i Trilinear Filtering**

Funkcje filtrujące wytwarzające miękkie przejścia i zapobiegające powstaniu tzw. efektu klocków.

##### **Flat Shading**

Barwienie wieloboków, przez co np. krawędzie wyglądają ostrzej i wyraźniej.

##### **Fogging**

Funkcja pozwalająca na realistyczne przedstawienie oddalających się obiektów, które niejako znikają we mgle.

##### **Gouard Shading**

Płynne przejście barw pomiędzy zadaną wartością początkową i końcową.

##### **MIP Mapping**

Jednej powierzchni przypisać można kilka tekstur o różnym stopniu szczegółowości. Obiekty znajdujące się bliżej wykazują większe bogactwo detali od obiektów znajdujących się dalej.

##### **Shading**

---

<sup>1</sup> Tekstura to struktura powierzchni trójwymiarowego obiektu, a Rendering to jej wypełnianie i cieniowanie.

Realistyczny wygląd zakrzywionych powierzchni możliwy jest przez ich odpowiednie cieniowanie. Dana powierzchnia zostaje w tym celu podzielona na malutkie trójkąty.

### **Texture Mapping**

Powierzchnie wypełniane są wzorami (np. ceglany mur), co prowadzi do bardziej realistycznego wrażenia niż w przypadku Gouraud Shading.

Innym kryterium świadczącym o jakości danej karty 3D jest płynność w budowie obrazu, co ma, zwłaszcza w grach komputerowych, decydujące znaczenie. Odpowiednia wielkość nazywana zwykle z angielska „Framerate” określa ile obrazów na sekundę (*frames per second - fps*) zostaje zbudowanych. Płynność osiągnięta zostaje dopiero po przekroczeniu 25 fps. Niższe wartości odbierane są przez ludzkie oko jako nieprzyjemne „skakanie” czy „przerywanie”.

Trójwymiarowe odwzorowanie wymaga specjalnych, standardowych interfejsów programowych. W przeciwnym wypadku każdy producent kart 3D realizowałby własne koncepcje, co z jednej strony wymaga zwiększonego wysiłku programistów, a z drugiej nie gwarantuje kompatybilności i pogłębia chaos na rynku kart i programów 3D.

### **Direct-3D**

Część składowa Direct X, stosowana przeważnie w grach dla Windows 95/98. Direct X zainstalowany musi zostać dodatkowo po instalacji Windows 95/98. W praktyce Direct X jest sprawcą wielu problemów, gdyż umożliwia on każdej grze dokonywanie zmian w sterowniku karty grafiki.

### **Open-GL**

Produkt firmy Silicon Graphics przeznaczony jest do pracy w profesjonalnej dziedzinie trójwymiarowego konstruowania, animacji komputerowych (np. w filmie "Dzień niepodległości") itp. Standard ten jest częścią składową Windows NT oraz Windows 95b (od wersji OSR2 z FAT32).

### **HEIDI**

Standard 3D programu konstrukcyjnego AUTOCAD oraz 3D Studio Max firmy AUTODESK, przeznaczony przede wszystkim do profesjonalnych zastosowań. To programowe łącze musi zostać dodatkowo zainstalowane.

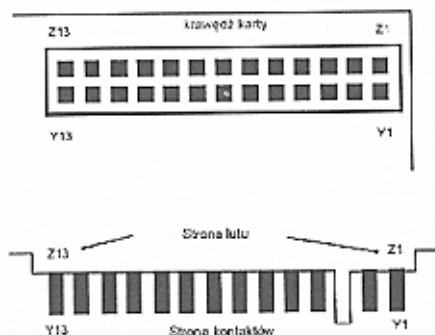
## **4.8. Accelerated Graphics Port.**

Komputery PC, nawet z najszybszymi kartami 3D, mimo znacznego wzrostu wydajności w odtwarzaniu grafiki i video, nie są w stanie konkurować z profesjonalnymi stacjami roboczymi firmy Hewlett Packard czy Silicon Graphics (SGI), które mogą obrabiać kompletne, pełnometrażowe filmy. Za „wąskie gardło” uważana jest powszechnie magistrala PCI, która tego typu zastosowaniom nie zapewnia odpowiedniej szybkości transferu informacji. Z inicjatywy firmy Intel, konsorcjum producentów kart graficznych zdefiniowało nowy standard, a mianowicie Accelerated Graphics Port (AGP). AGP ani nie zastępuje ani nie rozszerza magistrali PCI. Łącze to znajduje się na płycie głównej jako dodatkowe tzw. gniazdo Point-to-point (a więc nie magistrala), przewidziane na przyjęcie odpowiedniej karty graficznej AGP. Slot AGP nie jest ani elektrycznie ani mechanicznie kompatybilny z gniazdem PCI, od którego jest zresztą o wiele

mniejszy. AGP zintegrowany jest jako tzw. mostek (Bridge) z chipsetem płyty głównej i pracuje niezależnie od procesora, co po raz pierwszy umożliwia równoległą aktywność CPU i akceleratora graficznego, mogącego pracować jako tzw. *busmaster*. Wymiana informacji odbywa się bezpośrednio pomiędzy kartą AGP i pamięcią główną, co stawia szczególnie wysokie wymagania przed modułami RAM. Według firmy Intel, pamięć operacyjna do pracy z kartą AGP powinna się składać z kości SDRAM taktowanych z częstotliwością 66 lub 100 MHz. AGP, podobnie jak magistrala PCI, umożliwia transfer informacji o szerokości 32 bitów (32 bity na jeden takt), pracuje jednak z częstotliwością 66 MHz (jak PCI wersja 2.1), co daje teoretycznie maksymalną przepustowość - jeśli wykorzystane są obie flanki sygnału taktującego - rzędu 533 Mbps. Także niektóre sygnały PCI (sterujące) wykorzystane są przez AGP, a inicjalizacja takiej karty odbywa się kompletnie łączem PCI. Dopiero później do akcji wkracza karta AGP.

#### 4.9. VESA Feature connector

Liczne karty VGA oraz niektóre EGA wyposażone zostały w tzw. *Feature Connector*. Ta dodatkowa listwa kontaktowa potrzebna jest np. do rozgałęzienia sygnału VGA do innej, niekompatybilnej ze standardem VGA, karty grafiki lub video. Jako przykład mogą posłużyć niektóre karty TIGA firmy Texas Instruments lub adapter 8514 firmy IBM. Poza tym potrzebny jest on (Feature Connector) niektórym kartom video (np. VideoBlaster) lub dekodernom MPEG, które bezpośrednio wykorzystują kartę grafiki jako kartę odtwarzającą. Sygnały z tego gniazda karty grafiki przesyłane są za pomocą specjalnego kabla (płaski, przypominający taśmę kabel danych)



Rys. 4.2. Oba warianty Feature Connector spotykane na kartach graficznych.

do karty video lub grafiki, która musi być oczywiście wyposażona w odpowiednie łącze. Monitor podłączany bywa wtedy często zamiast do karty VGA, do tej drugiej karty (niekompatybilnej ze standardem VGA, EGA), jak to ma miejsce w przypadku prostszych kart video, które po prostu „wpuszczają” swój sygnał do informacji graficznej płynącej do monitora. Łącze Feature Connector spotyka się w dwóch różnych wariantach:

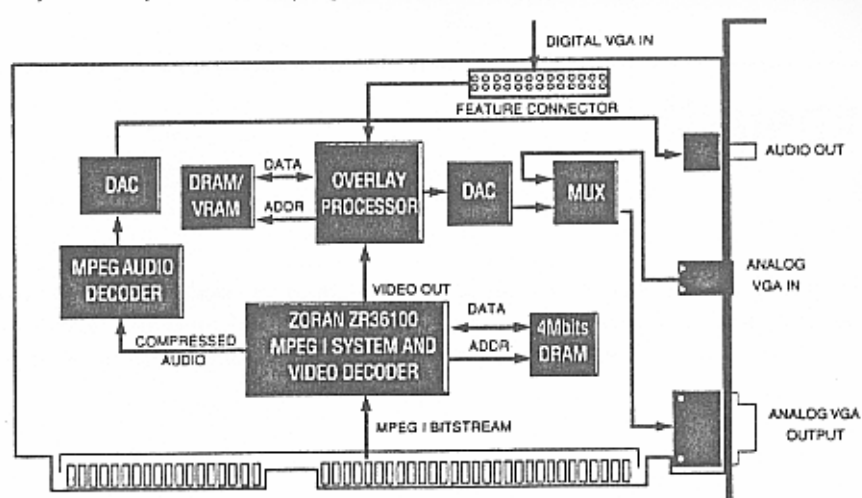
jako 26-biegunowa (dwurzędowa) listwa kontaktowa lub jako wtyczka „grzebieniowa”, która w przedłużonej formie (34-biegunowa) spotykana bywa w napędach dyskietek 5,25". Nie wszyscy producenci stosują podane poniżej przyporządkowanie sygnałów (np. karta SPEA VEGA). Kolumny Z i Y są wtedy pozamieniane. W takim przypadku (błędne połączenia) karta nie będzie działać jednak nie powinno dojść do żadnych elektrycznych uszkodzeń. Poza tym może się zdarzyć, że korzystanie z łącza Feature Connector pogarsza jakość odwzorowywanej grafiki, tak że trzeba przejść do pracy w niższej rozdzielczości lub głębi barw. Często dopiero wtedy jakiegokolwiek odwzorowanie staje się w ogóle możliwe.

Obok Feature Connector na rynku pojawiły się w międzyczasie inne (lepsze) możliwości połączenia karty grafiki z kartą video, dekodernem MPEG, tzw. kartą capture czy z kartą tunera TV,



przy czym funkcjonują one zwykle tylko w obrębie produktów tego samego producenta, jak np. Multimedia Channel firmy ATI czy Analog-Loop-Through firmy Diamond.

Konsorcjum VESA już od kilku lat próbuje wprowadzić jakiś standard, który umożliwiłby łączenie tego typu kart między sobą. Powstała nawet odpowiednia specyfikacja tzw. VESA Media Channel (VMC), która jednak została faktycznie zrealizowana tylko przez niewielu producentów. W pewnym sensie jako wyjątek posłużyć może angielska firma VideoLogic, która wyposaża niektóre swoje karty w tę magistralę (np. VideoLogic968 z chipem S3). VMC nie zdobyła jednak większej popularności. Czy uda się to innym, jak np. firmie Siemens-Nixdorf z tzw. SNI-Channel, który jest już bardziej rozpowszechniony niż VMC, pozostaje na razie kwestią otwartą. Jakby nie było, Feature Connector to ciągle najbardziej rozpowszechniony standard jeśli chodzi o połączenie różnych kart związanych mniej lub bardziej z grafiką pceta.



Rys. 4.3. Także dekodery MPEG do połączenia z kartą grafiki stosują Feature Connector.