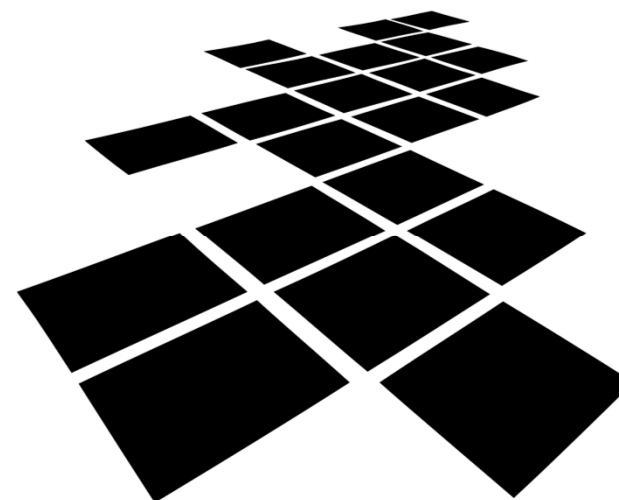


Urządzenia wejścia-wyjścia

**Wykład prowadzą:
Jerzy Brzeziński
Dariusz Wawrzyniak**



**UCZELNIA
ONLINE**



Plan wykładu

- Klasyfikacja urządzeń wejścia-wyjścia
- Struktura mechanizmu wejścia-wyjścia (sprzętu i oprogramowania)
- Interakcja jednostki centralnej z urządzeniami wejścia-wyjścia
 - odpytywanie
 - sterowanie przerwami
 - bezpośredni dostęp do pamięci
- Buforowanie, przechowywanie podręczne i spooling
- Wirtualne wejście-wyjście



Rodzaje urządzeń wejścia-wyjścia

- Urządzenia składowania danych (dyski, dyskietki, taśmy, CD ROM, DVD itp.)
- Urządzenia transmisji danych na odległość (karty sieciowe, modemy)
- Urządzenia do komunikacji z człowiekiem (monitory, projektory, klawiatury, myszy, drukarki, skanery itp.)
- Urządzenia specjalizowane
 - układy sterowania (np. elektrownią, samolotem, systemem obrony antyrakietowej itd.)
 - kasy i drukarki fiskalne itp.
 - urządzenia medyczne



Właściwości urządzeń wejścia-wyjścia ⁽¹⁾

- Tryb transmisji danych:
 - znakowy — przekazywanie danych odbywa się bajt po bajcie, przykład: port szeregowy
 - blokowy — przekazywanie danych odbywa się w blokach (np. po 512 bajtów), przykład: dysk
- Sposób dostępu do danych:
 - sekwencyjny — dane przekazywane są w określonym porządku, narzuconym przez urządzenie, przykład: karta sieciowa
 - bezpośredni (swobodny) — możliwe jest określenie lokalizacji danych na urządzeniu, przykład: dysk



Właściwości urządzeń wejścia-wyjścia (2)

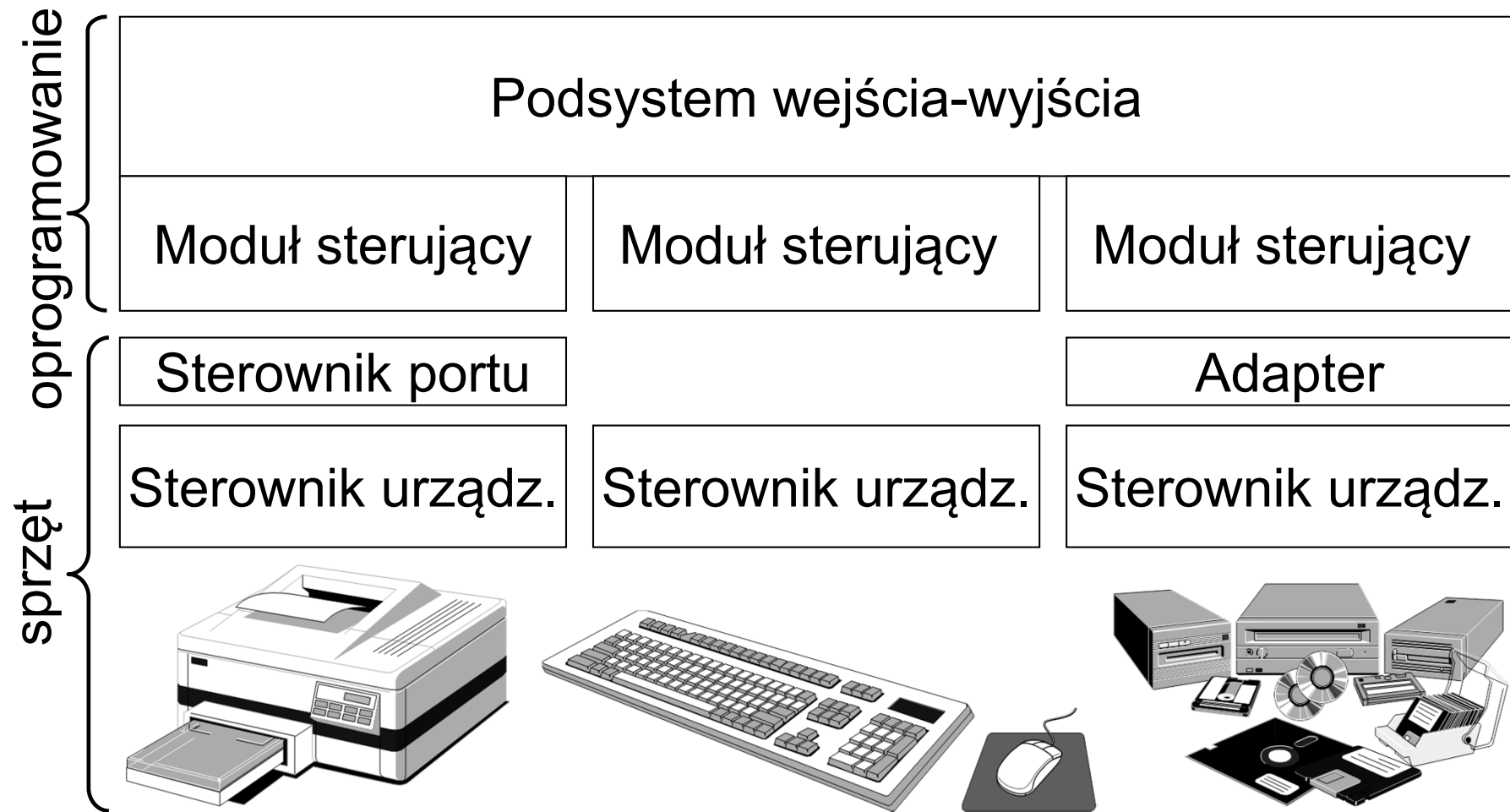
- Tryb pracy urządzenia:
 - synchroniczny — dane zostaną przekazane w znanym z góry (przewidywalnym) czasie, przykład: dysk, karta graficzna
 - asynchroniczny — dane mogą zostać przesłane w dowolnym, trudnym do przewidzenia, momencie, przykład: klawiatura, karta sieciowa
- Tryb użytkowania:
 - współdzielony — dopuszczalne jest współbieżne używanie urządzenia przez wiele procesów, np.: dysk
 - wyłączny — niemożliwe jest współbieżne używanie urządzenia przez wiele procesów, przykład: drukarka



Właściwości urządzeń wejścia-wyjścia ⁽³⁾

- Szybkość działania (transmisji)
 - od bardzo wolnych, przykład: drukarka
 - do stosunkowo szybkich, przykład: dysk
- Kierunek przekazywania danych
 - urządzenia wejścia i wyjścia — możliwość zarówno zapisu jak i odczytu, przykład dysk, karta sieciowa
 - urządzenia wejścia — tylko możliwość odczytu z urządzenia, przykład: klawiatura
 - urządzenia wyjścia — tylko możliwość zapisu, przykład: drukarka

Struktura mechanizmu wejścia-wyjścia





Oprogramowanie obsługi wejścia-wyjścia

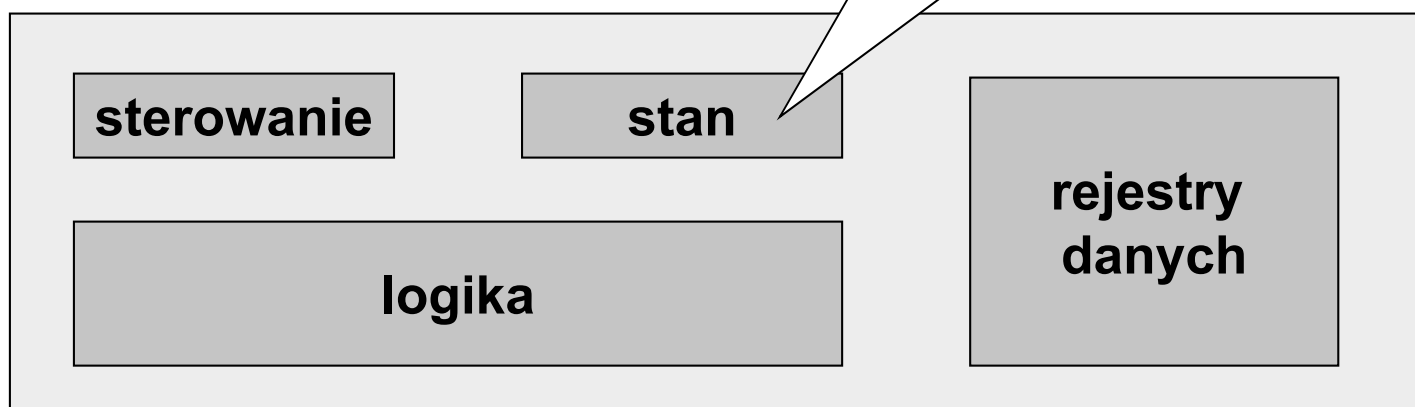
- Podsystem wejścia-wyjścia:
 - interfejs wejścia-wyjścia — specyfikacja operacji (API), umożliwiających jednolity sposób dostępu do urządzeń wejścia-wyjścia na poziomie aplikacji
 - buforowanie
- Moduł sterujący — ukrywanie sprzętowych szczegółów realizacji danego urządzenia przed interfejsem wejścia-wyjścia:
 - dostarczanie implementacji dla operacji z interfejsu wejścia-wyjścia w odniesieniu do danego urządzenia
 - kontrola realizacji operacji na urządzeniu (sprawdzanie stanu, poprawności, transfer danych)



Sterownik urządzenia

	zajętość	gotowość
bezczynność	0	0
zakończenie	0	1
praca	1	0
(stan przejść.)	1	1

... zajętość gotowość kod błędu ...





Miejsce urządzeń wejścia-wyjścia w architekturze systemu komputerowego

- Odwzorowanie w przestrzeni adresowej wejścia-wyjścia (tzw. izolowane wejście-wyjście) — rejestry sterownika widoczne są w przestrzeni adresowej wejścia-wyjścia systemu komputerowego i dostępne są przez specjalne rozkazy (np. `in` i `out` w procesorach firmy Intel).
- Odwzorowanie w przestrzeni adresowej pamięci — rejestry sterownika widoczne są w przestrzeni adresowej pamięci fizycznej i dostępne są pod odpowiednimi adresami tak samo, jak inne komórki pamięci.

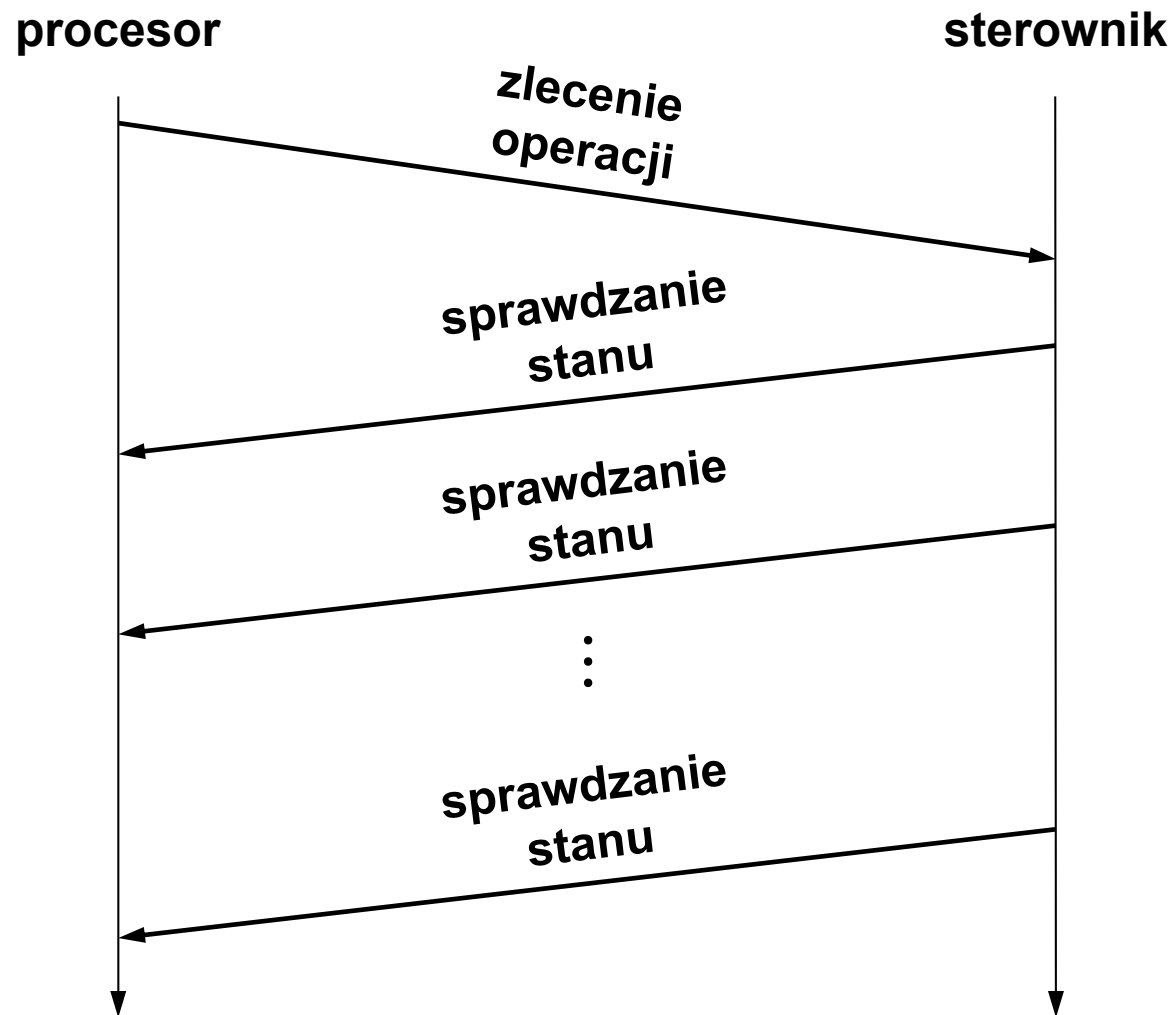


Interakcja jednostki centralnej ze sterownikiem urządzenia wejścia-wyjścia

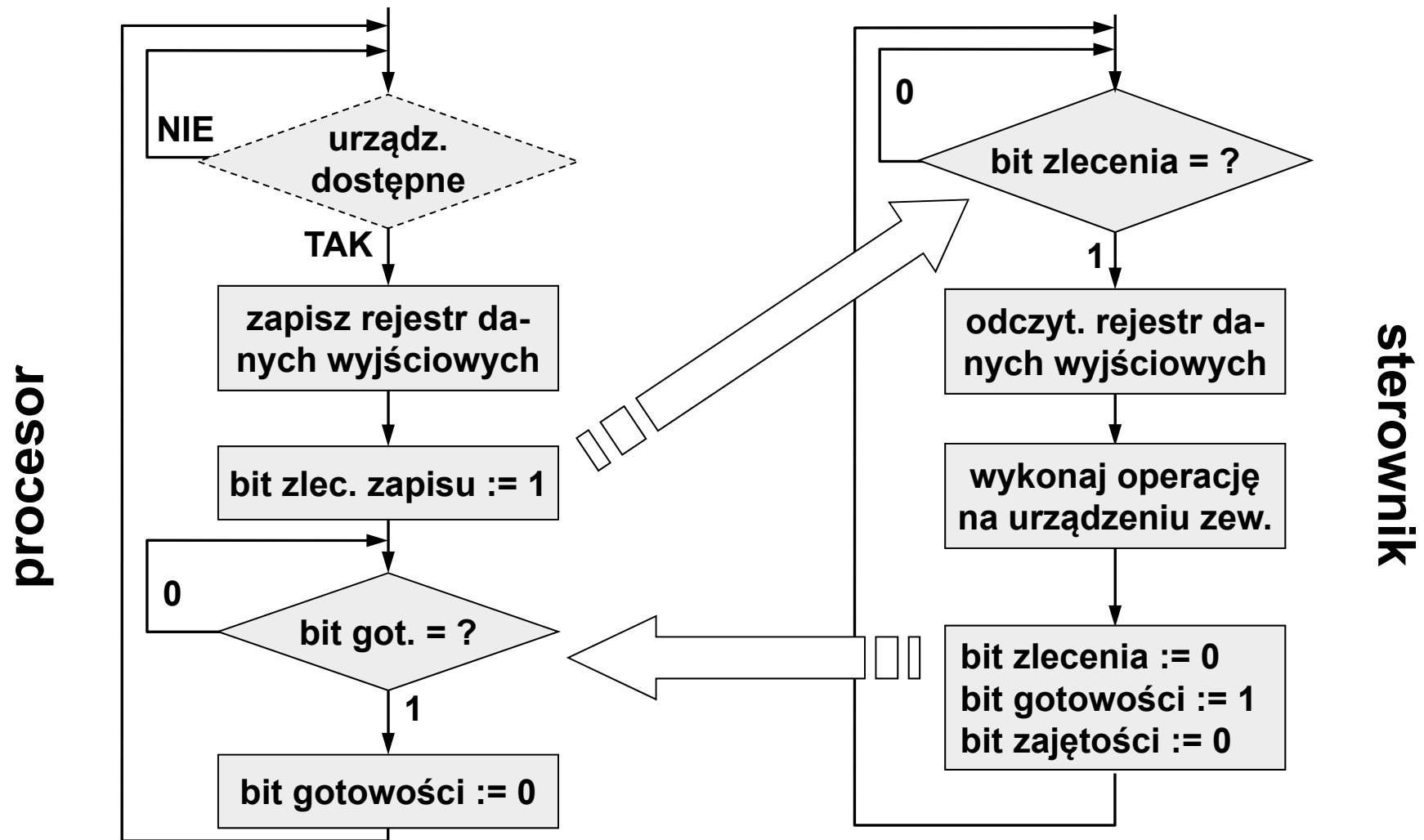
- Odpytywanie (ang. polling) — ciągłe lub okresowe sprawdzanie stanu sterownika
- Sterowanie przerwami (ang. interrupt-driven I/O) — inicjalizacja pracy sterownika przez procesor i obsługa urządzenia po zakończeniu działania w ramach reakcji na przerwanie
- Bezpośredni dostęp do pamięci (ang. direct memory access) — inicjalizacja pracy sterownika przez procesor i uruchomienie układu bezpośredniego dostępu do pamięci w celu realizacji transferu danych pomiędzy sterownikiem a pamięcią



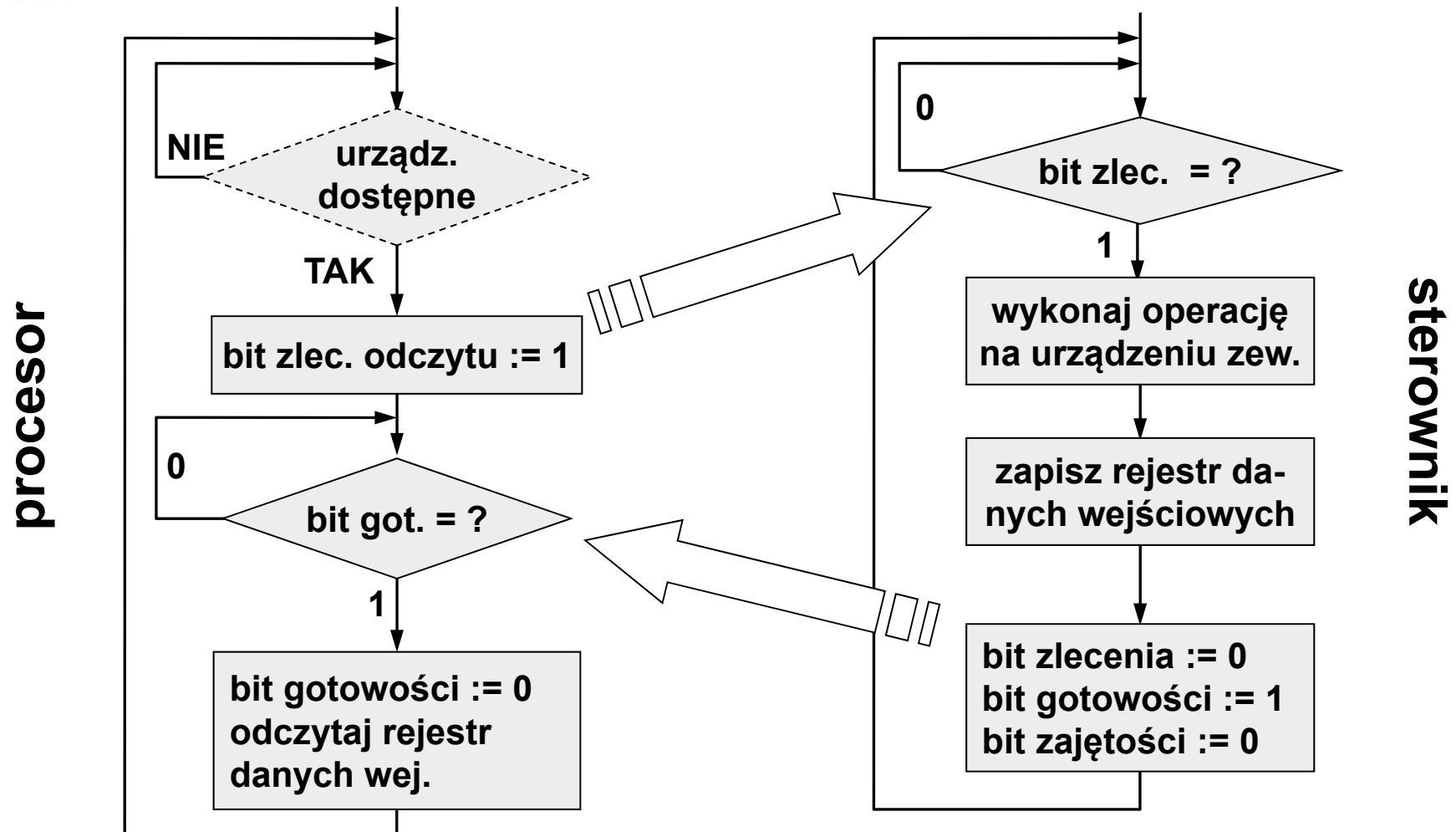
Przebieg czasowy odpytywania



Interakcja procesor — sterownik w operacji wyjścia w trybie odpytywania

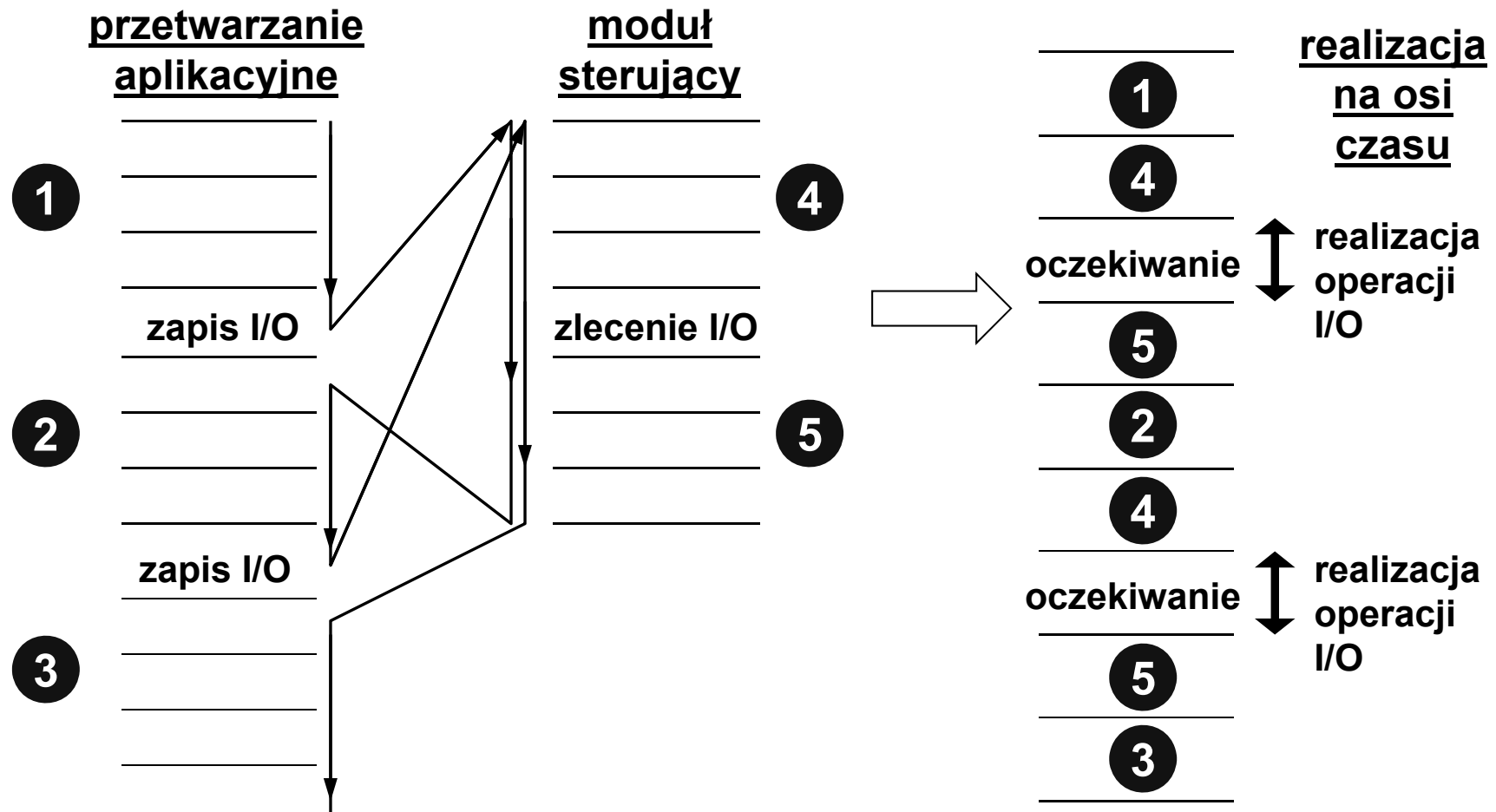


Interakcja procesor — sterownik w operacji wejścia w trybie odpytywania



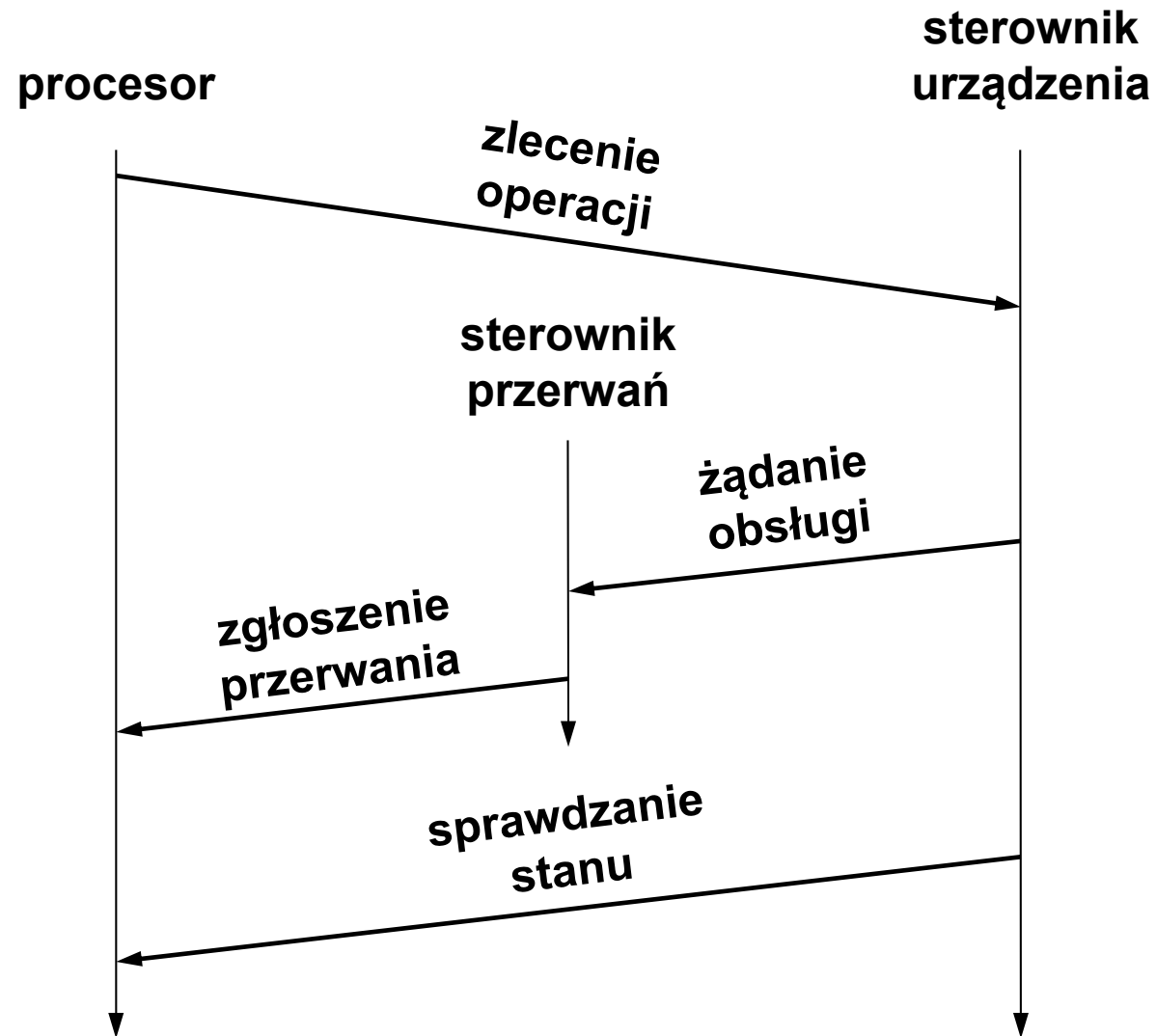


Odpytywanie — przepływ sterowania



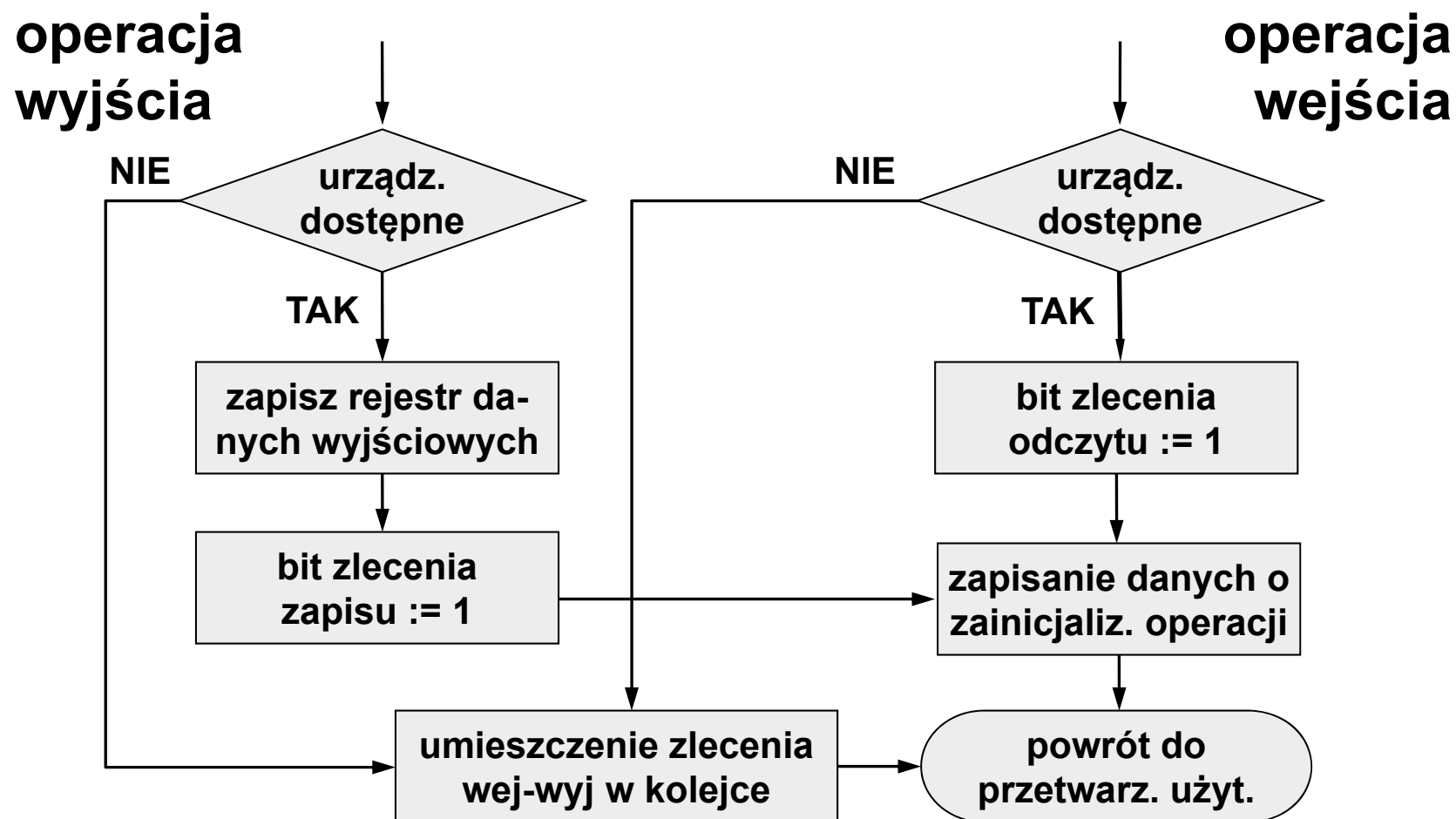


Przebieg czasowy obsługi urządzenia sterowanej przerwami





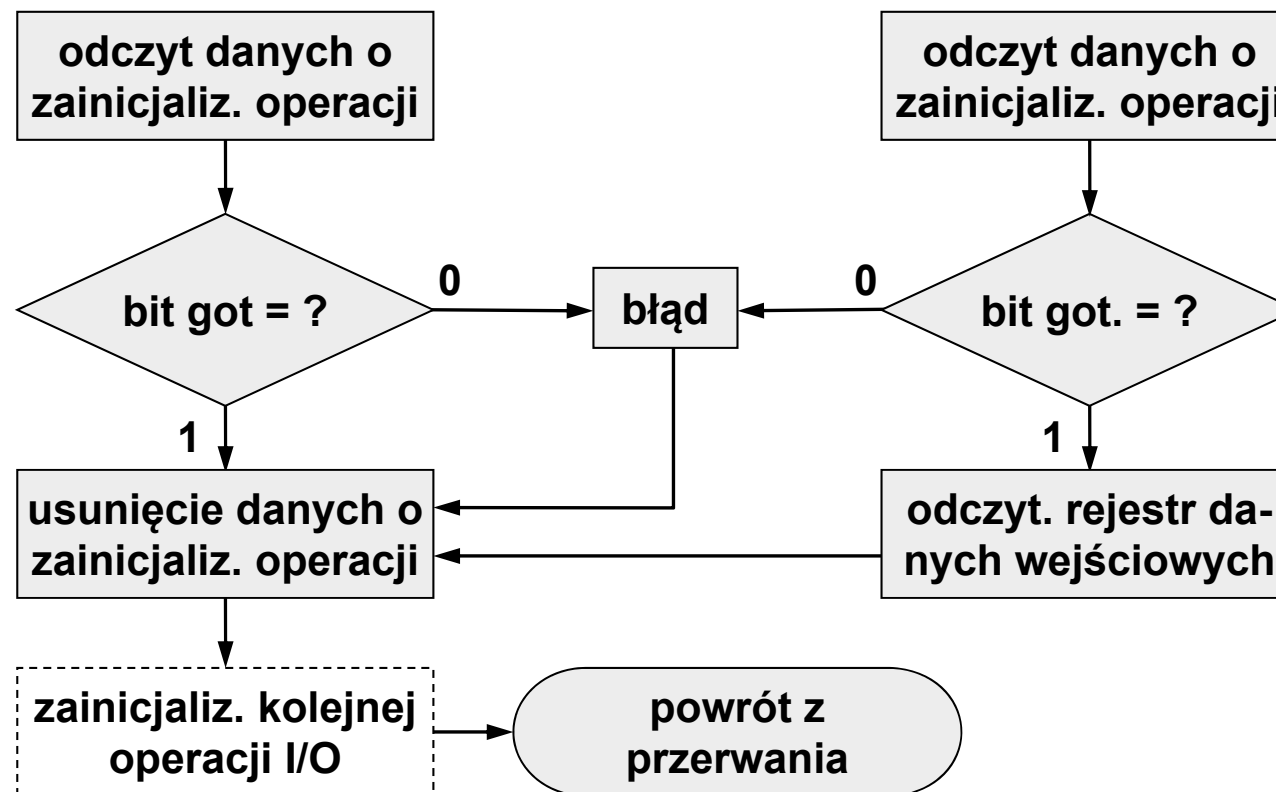
Obsługa sterowana przerwaniem — zlecenie operacji



Obsługa sterowana przerwaniem — reakcja na przerwanie

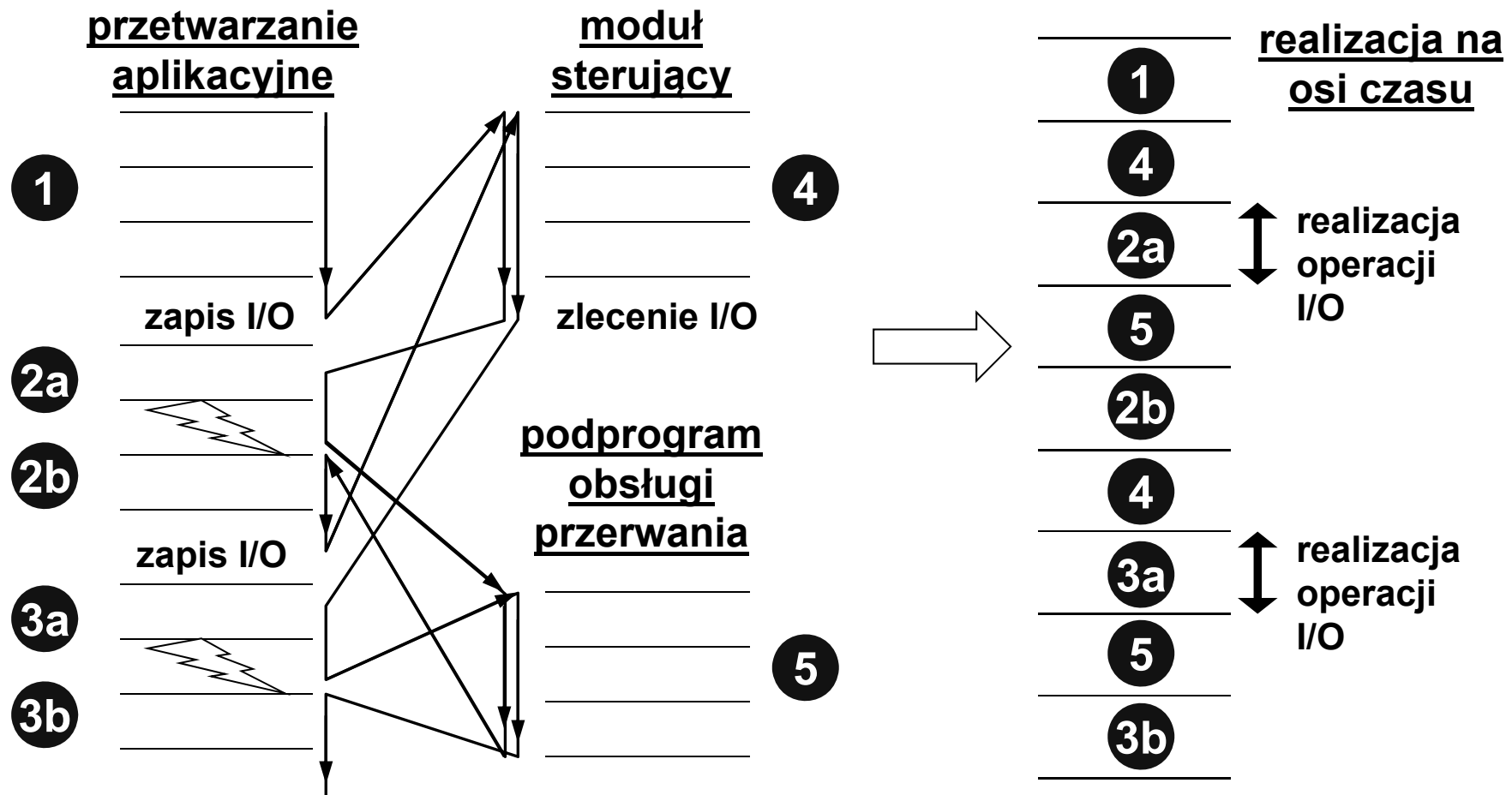
operacja wyjścia

operacja wejścia





Sterowanie przerwaniem — przepływ sterowania



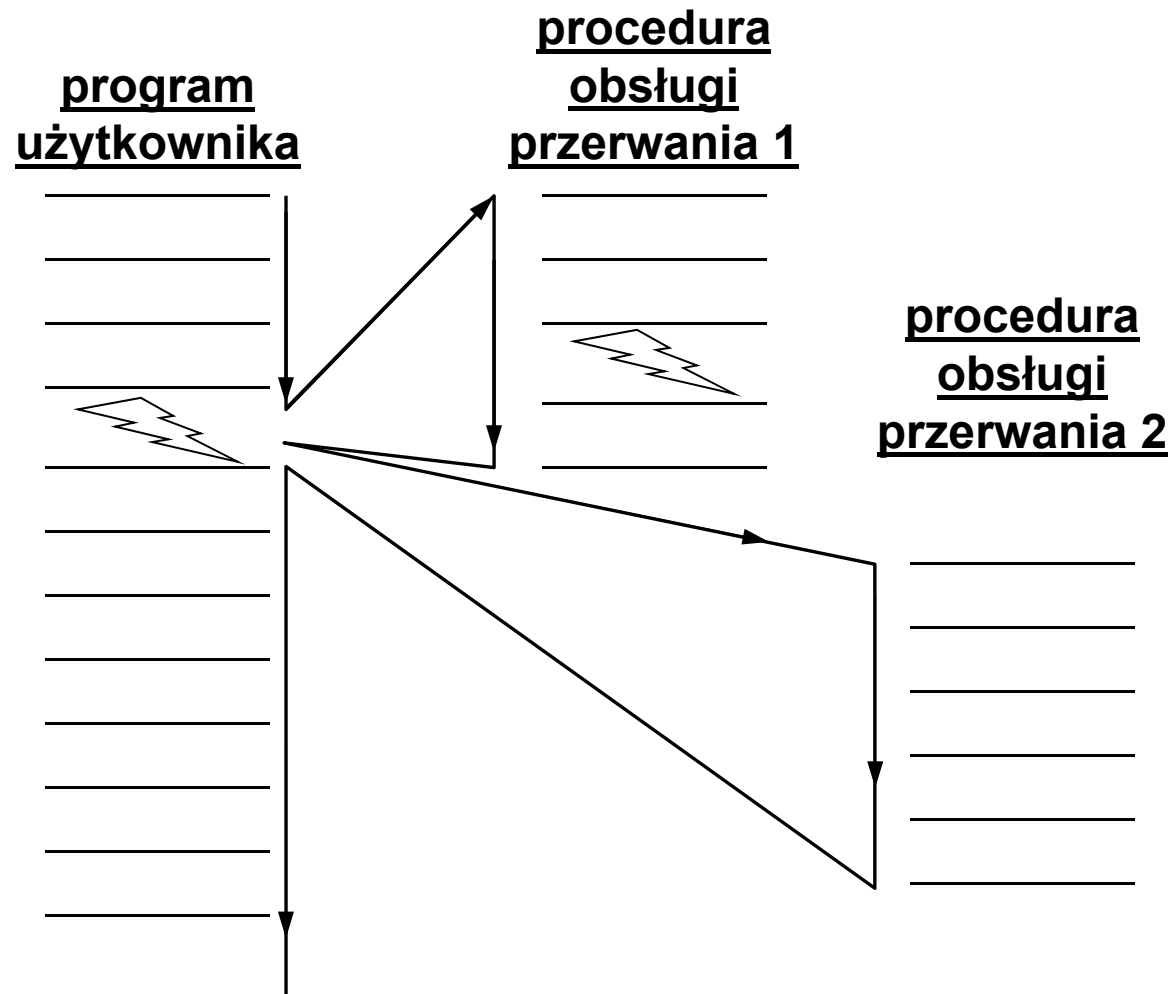


Obsługa przerwań wielokrotnych

- Przerwania wielokrotne oznaczają zgłoszenie kolejnego przerwania w czasie obsługi innego przerwania.
- Podejścia do obsługi przerwań wielokrotnych:
 - obsługa sekwencyjna,
 - obsługa zagnieżdżona,
 - obsługa priorytetowa

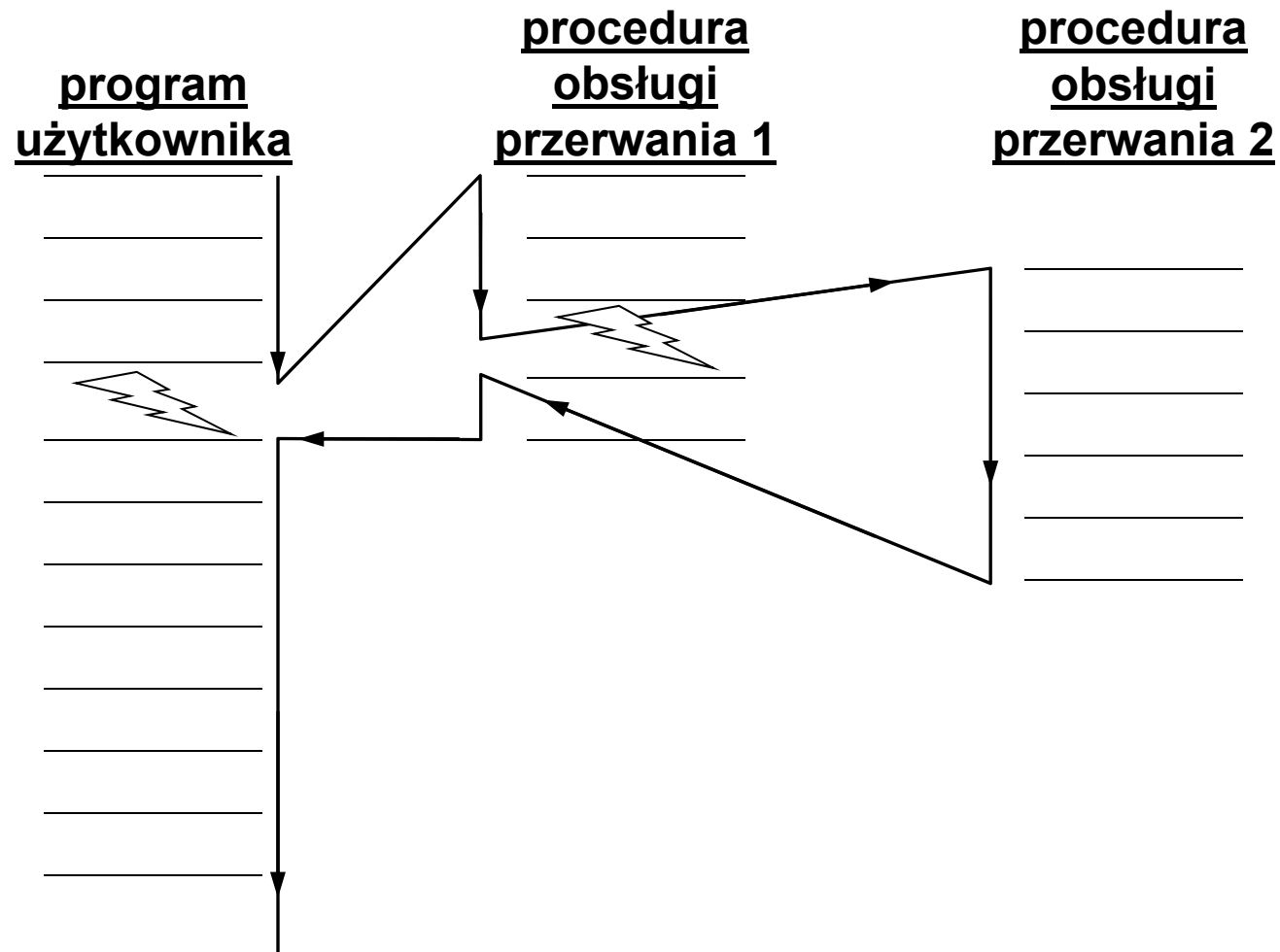


Sekwencyjna obsługa przerw — przepływ sterowania



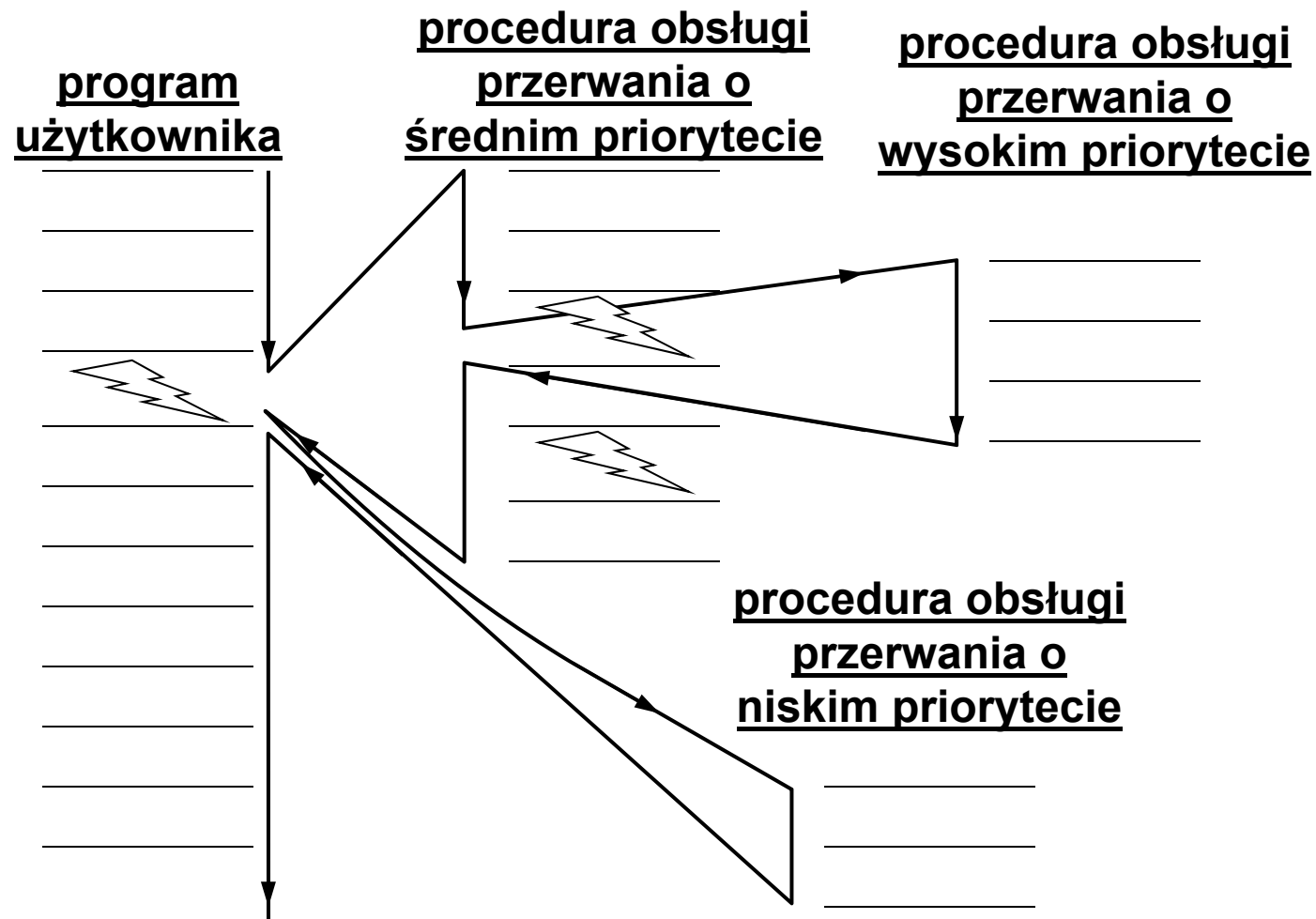


Zagnieżdżona obsługa przerwań — przepływ sterowania





Priorytetowa obsługa przerw — przepływ sterowania





Problemy współbieżnej obsługi wielu urządzeń

- Problem identyfikacji źródła przerwania — zidentyfikowanie urządzenia, które poprzez zgłoszenie przerwania wymusiło przekazanie sterowania do procedury obsługi przerwania.
- Problem priorytetów — zagwarantowanie określonej kolejności wyboru urządzeń w przypadku deklaracji gotowości kilku z nich w tym samym czasie.



Sposoby identyfikacji źródła przerwania

- Wiele linii przerwań — doprowadzenie do procesora osobnej linii przerwania dla każdego urządzenia
- Odpytywanie programowe — odczyt rejestru stanu i sprawdzanie bitu gotowości każdego urządzenia
- Odczyt wektora — odczytanie z magistrali informacji o adresie lub numerze urządzenia, które zgłosiło przerwanie, lub też numerze zgłoszonego przerwania



Poziomy obsługi przerwań

- poziom 0 — wszystkie przerwania odblokowane
- poziom 1 — wyłączenie zadań okresowych szeregowanych za pomocą czasomierzy
- poziom 2 — wyłączenie przetwarzania protokołu sieciowego
- poziom 3 — wyłączenie przerwań od terminala
- poziom 4 — wyłączenie przerwań od dysku
- poziom 5 — wyłączenie przerwań od urządzeń sieciowych
- poziom 6 — wyłączenie przerwań od zegara sprzętowego
- poziom 7 — wyłączenie wszystkich przerwań



Efektywność interakcji procesora ze sterownikiem — oznaczenia

- T_c — czas przetwarzania (obliczeń) przez procesor
- T_d — czas realizacji operacji wejścia-wyjścia przez sterowniki urządzeń
- T_o — skumulowany narzut czasowy, wynikający ze zwłoki pomiędzy zgłoszeniem gotowości przez sterownik, a reakcją przez procesor
- T_t — całkowity czas realizacji przetwarzania



Narzut czasowy

- W trybie odpytywania narzut czasowy $T_o = T_p$, gdzie
 - T_p — skumulowane opóźnienie w pętli odpytywania pomiędzy ustawieniem bitu gotowości, a odczytaniem rejestru stanu.
- W trybie sterowania przerwami narzut czasowy $T_o = T_b + T_h + T_r$, gdzie:
 - T_b — skumulowany czas oczekiwania na zwolnienie urządzenia,
 - T_h — skumulowany czas obsługi przerw,
 - T_r — skumulowany czas oczekiwania na przydział procesora po zakończeniu operacji wejścia-wyjścia.



Porównanie efektywności przetwarzania w systemie jednozadaniowym

- W systemie jednozadaniowym czas cyklu przetwarzania
$$T_t = T_c + T_d + T_o$$
- Ponieważ $T_p < T_h + T_b + T_r$
 - w systemie jednozadaniowym odpytywanie zapewnia większą efektywność nawet przy założeniu, że $T_r = 0$ i $T_b = 0$,
 - z punktu widzenia pojedynczego procesu odpytywanie zapewnia większą efektywność



Czas przetwarzania w systemie wielozadaniowym ⁽¹⁾

- Rozważmy zbiór procesów współbieżnych $P_1, \dots, P_n$.
- Niech $T_{t,i}$, $T_{d,i}$, $T_{c,i}$, $T_{p,i}$, $T_{h,i}$, $T_{b,i}$ i $T_{r,i}$ oznaczają odpowiednie parametry czasowe procesu P_i .
- Całkowity czas przetwarzania zbioru procesów w trybie odpytywania:

$$\sum_{i=1}^n T_{t,i} = \sum_{i=1}^n T_{c,i} + \sum_{i=1}^n T_{d,i} + \sum_{i=1}^n T_{p,i}$$



Czas przetwarzania w systemie wielozadaniowym (2)

- Całkowity czas przetwarzania w trybie sterowania przerwami, przy założeniu, że:

$$\sum_{i=1}^n T_{c,i} \geq \sum_{i=1}^n T_{d,i}$$

- wariant optymistyczny

$$\sum_{i=1}^n T_{t,i} = \sum_{i=1}^n T_{c,i} + \sum_{i=1}^n T_{h,i}$$

- wariant pesymistyczny

$$\sum_{i=1}^n T_{t,i} = \sum_{i=1}^n T_{c,i} + \sum_{i=1}^n T_{h,i} + \max_{i=1..n} T_{d,i}$$



Porównanie efektywności w systemie wielozadaniowym

- Ogólna efektywność przetwarzania w trybie sterowania przerwami w wariancie optymistycznym jest większa w porównaniu z trybem odpytywania wówczas, gdy

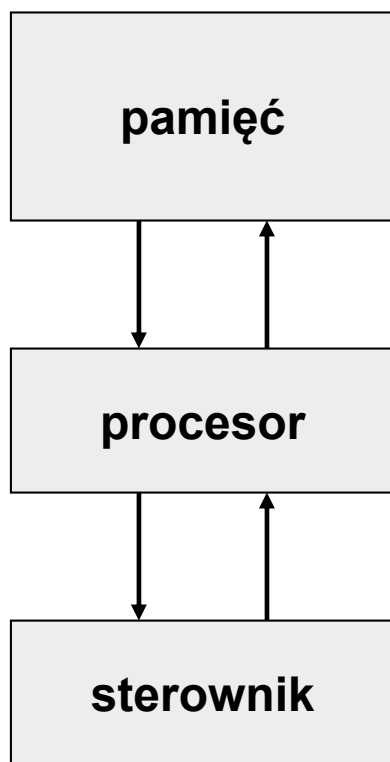
$$\sum_{i=1}^n T_{d,i} + \sum_{i=1}^n T_{p,i} \geq \sum_{i=1}^n T_{h,i}$$

- Ogólna efektywność przetwarzania w trybie sterowania przerwami w wariancie pesymistycznym jest większa w porównaniu z trybem odpytywania wówczas, gdy

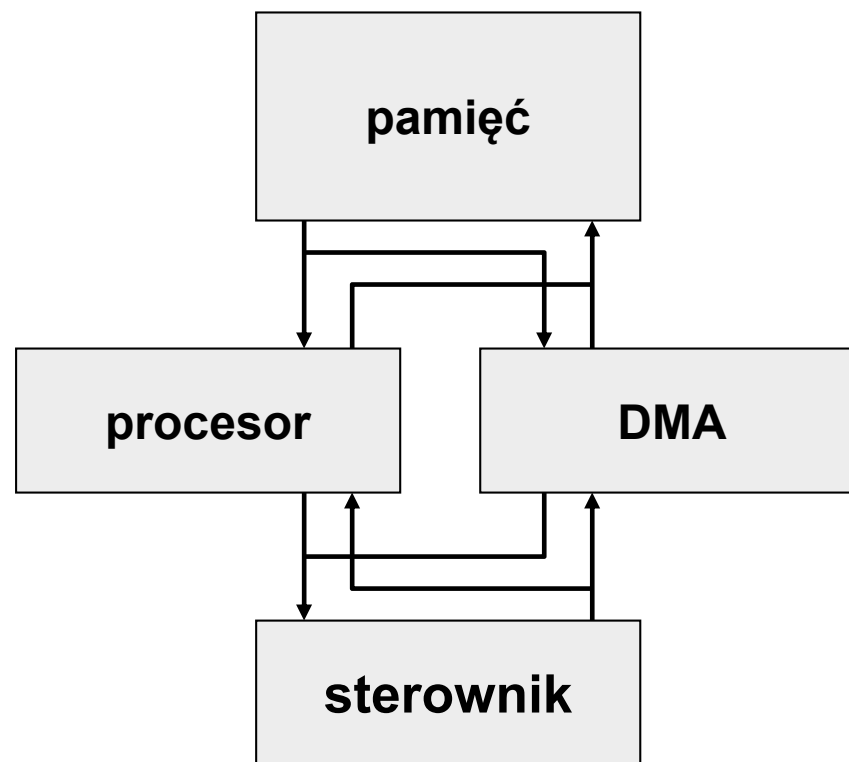
$$\sum_{i=1}^n T_{d,i} - \max_{i=1..n} T_{d,i} + \sum_{i=1}^n T_{p,i} \geq \sum_{i=1}^n T_{h,i}$$



Bezpośredni dostęp do pamięci



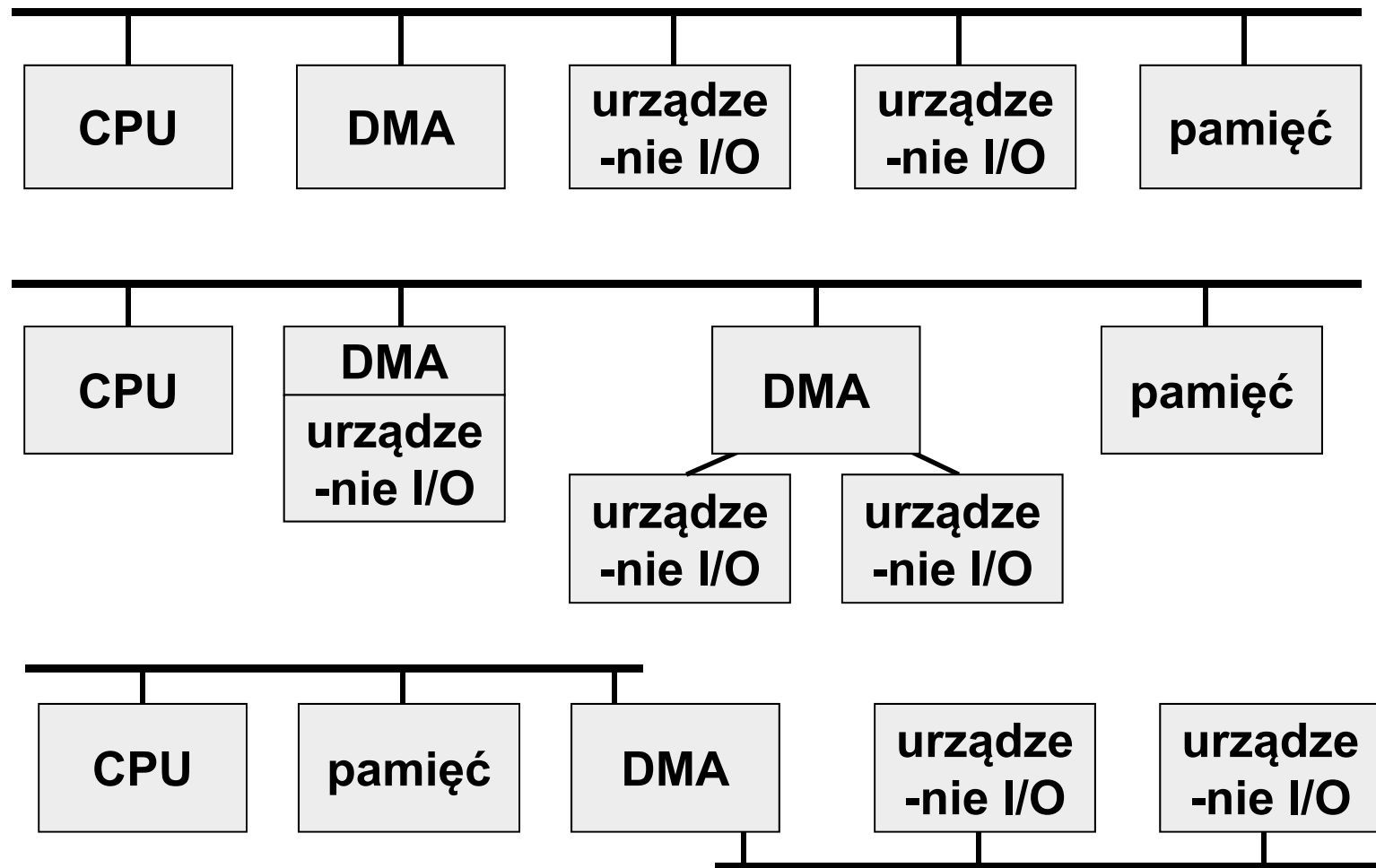
tradycyjne I/O



I/O z DMA



Organizacja wejścia-wyjścia z układem DMA





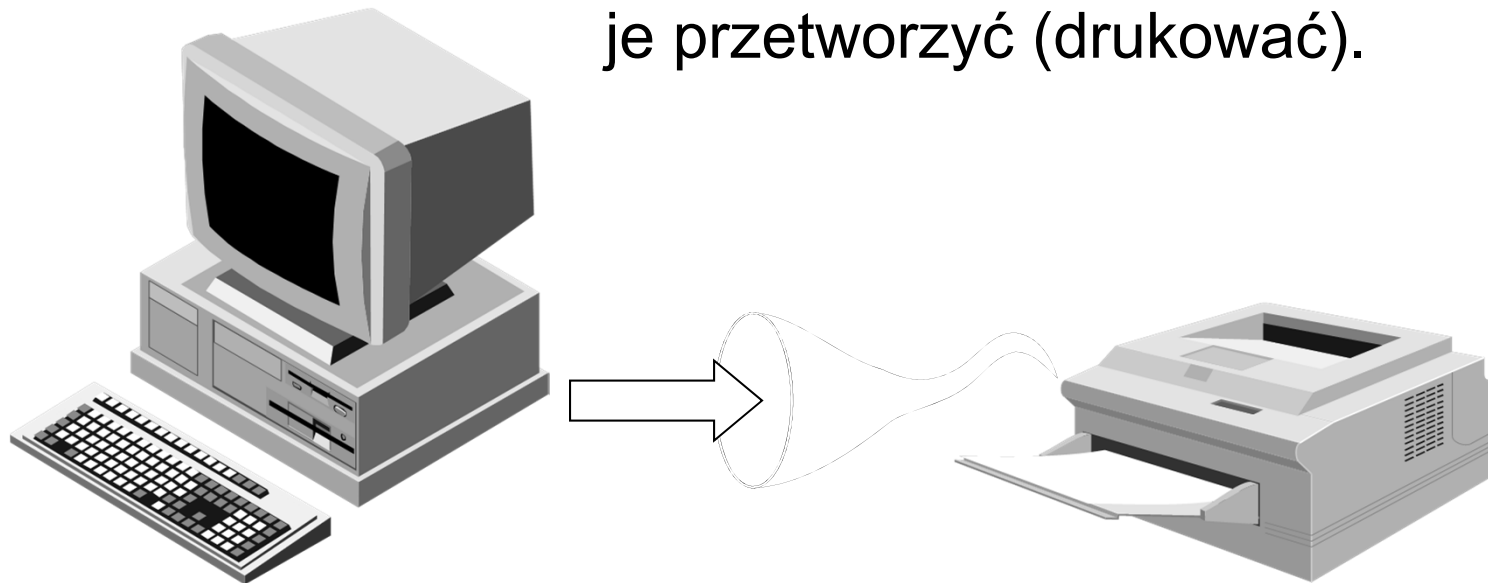
Buforowanie

- Dopasowanie urządzeń różniących się szybkością przekazywania danych — dopasowanie chwilowo szybszego producenta danych do możliwości konsumenta.
- Dopasowanie urządzeń różniących się podstawową jednostką transmisji danych — dopasowanie w celu efektywnego przekazywania danych urządzeń przesyłających mniejsze jednostki danych do urządzeń wymagających większych jednostek lub odwrotnie (fragmentowanie).
- Semantyka kopii — zagwarantowanie niezmienności danych w czasie wykonywania operacji wejścia-wyjścia.



Dopasowanie różnic szybkości

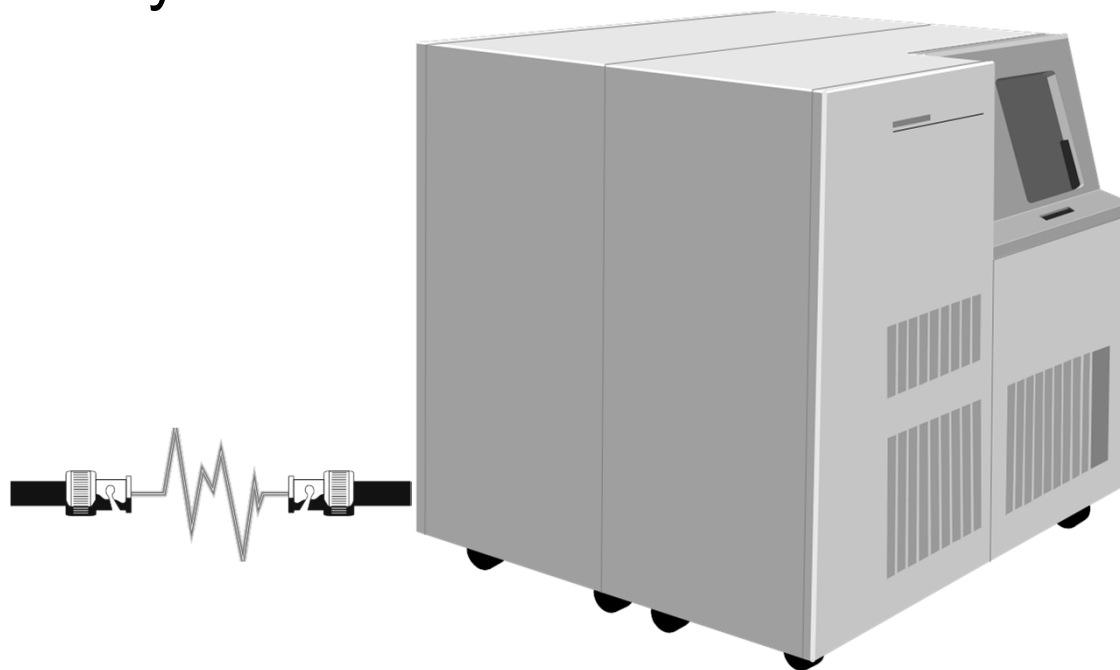
Przykład: komputer potrafi przekazać dane znacznie szybciej niż drukarka je przetworzyć (drukować).





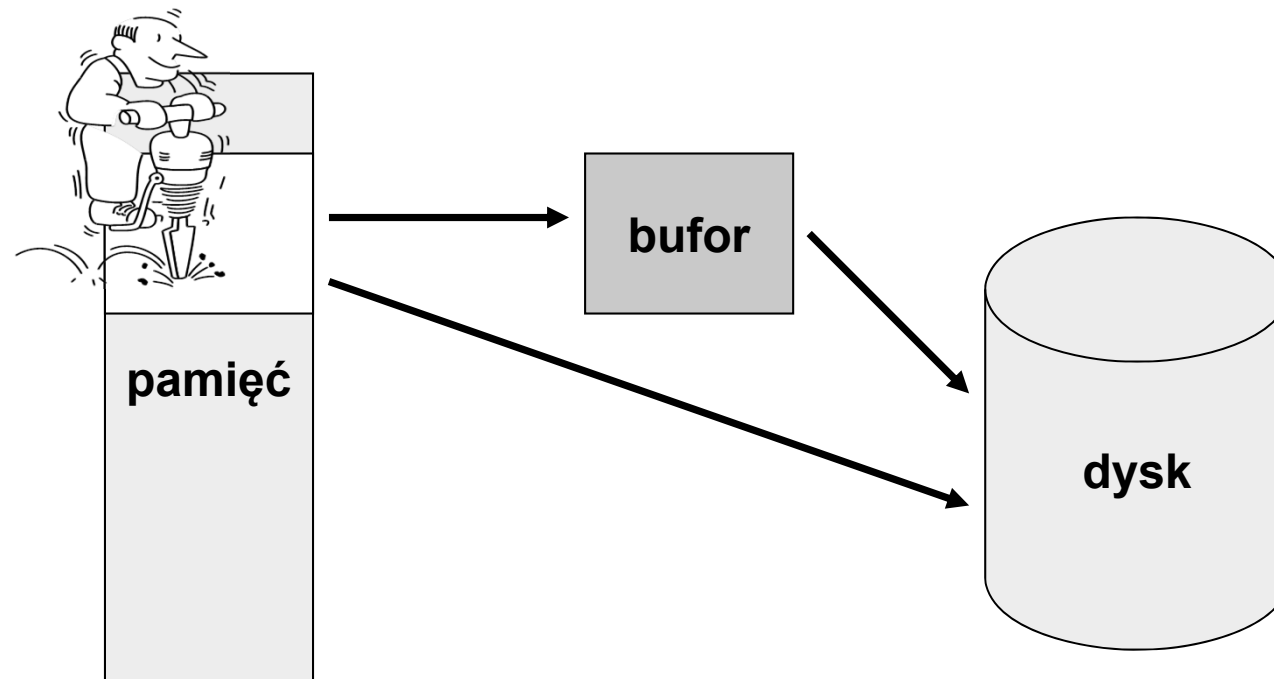
Dopasowanie jednostek transmisji

Przykład: zapis na dysku
danych odbieranych z
sieci.



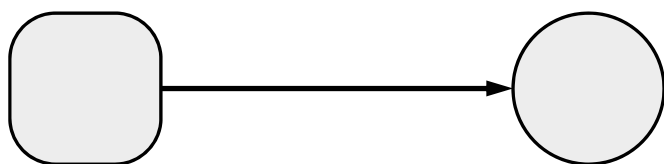


Semantyka kopii

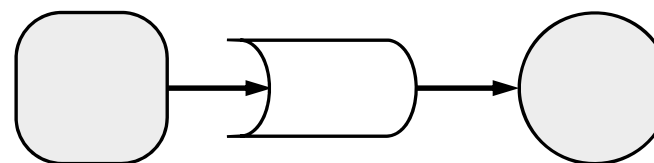




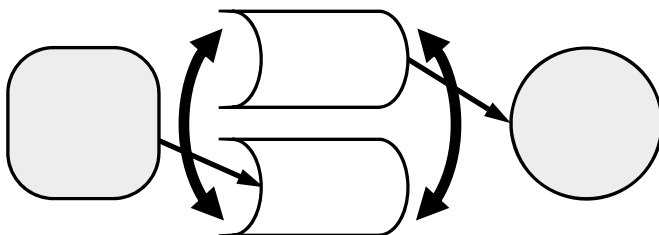
Realizacja buforowania



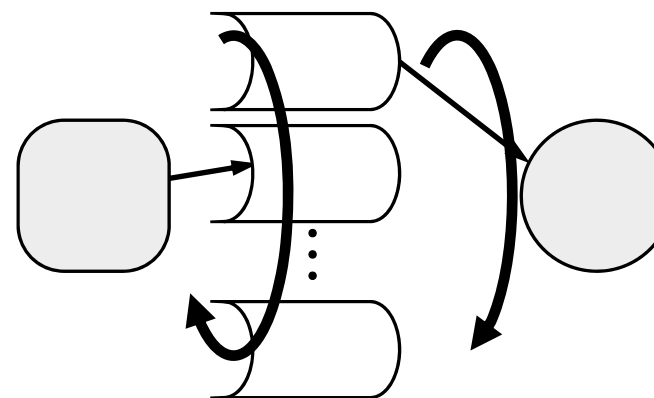
brak buforowania



bufor pojedynczy



bufor podwójny



bufor cykliczny

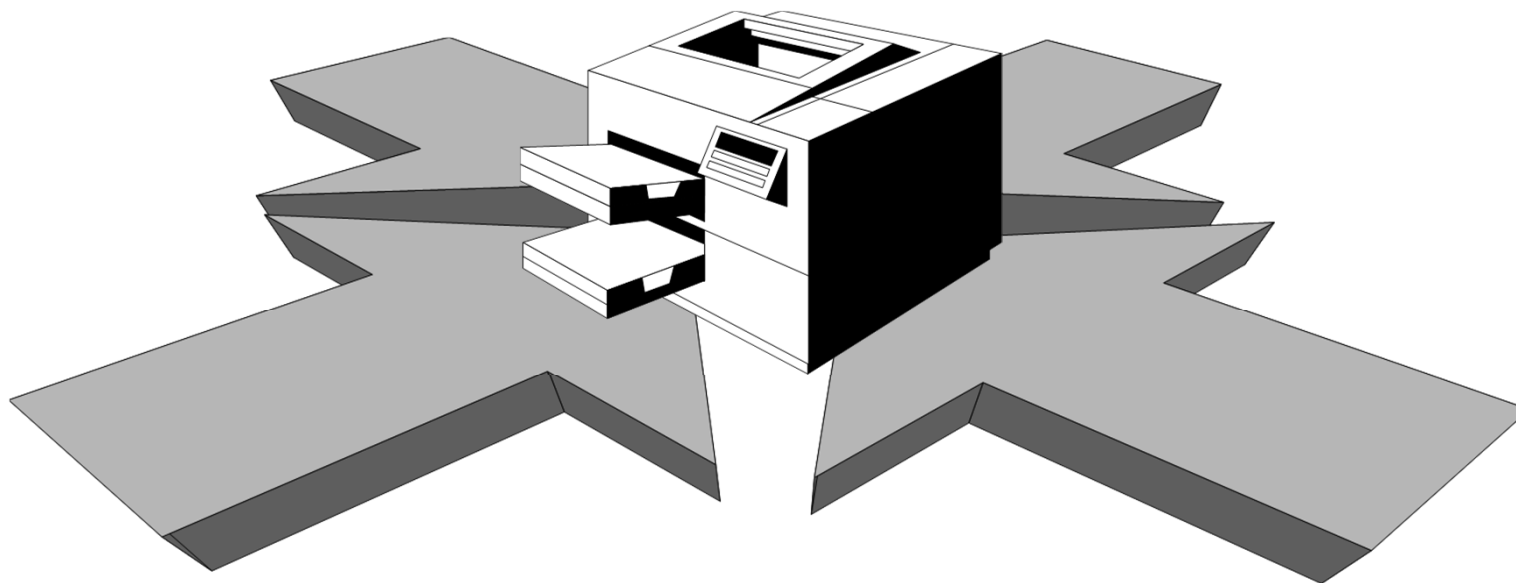


Przechowywanie podręczne

- Przechowywanie podręczne polega na gromadzeniu kopii danych w pamięci w celu poprawy efektywności ich przetwarzania.
- Przechowywanie podręczne w przypadku operacji wejścia zmniejsza czas dostępu.
- Przechowywanie podręczne w przypadku operacji wyjścia umożliwia skumulowanie wyników przetwarzania w dłuższym czasie i przekazanie ich na urządzenie zewnętrzne w wyniku jednej operacji wyjścia.



Spooling





Wirtualne wejście-wyjście

