Akademia Nauk Stosowanych
Instytut Techniczny
Kierunek: Informatyka
studia I stopnia, semestr 3

AKADEMIA NAUK STOSOWANYCH W NOWYM SĄCZU

Systemy operacyjne

WYKŁAD 7 i 8

dr inż. Stanisława Plichta

splichta@ans-ns.edu.pl

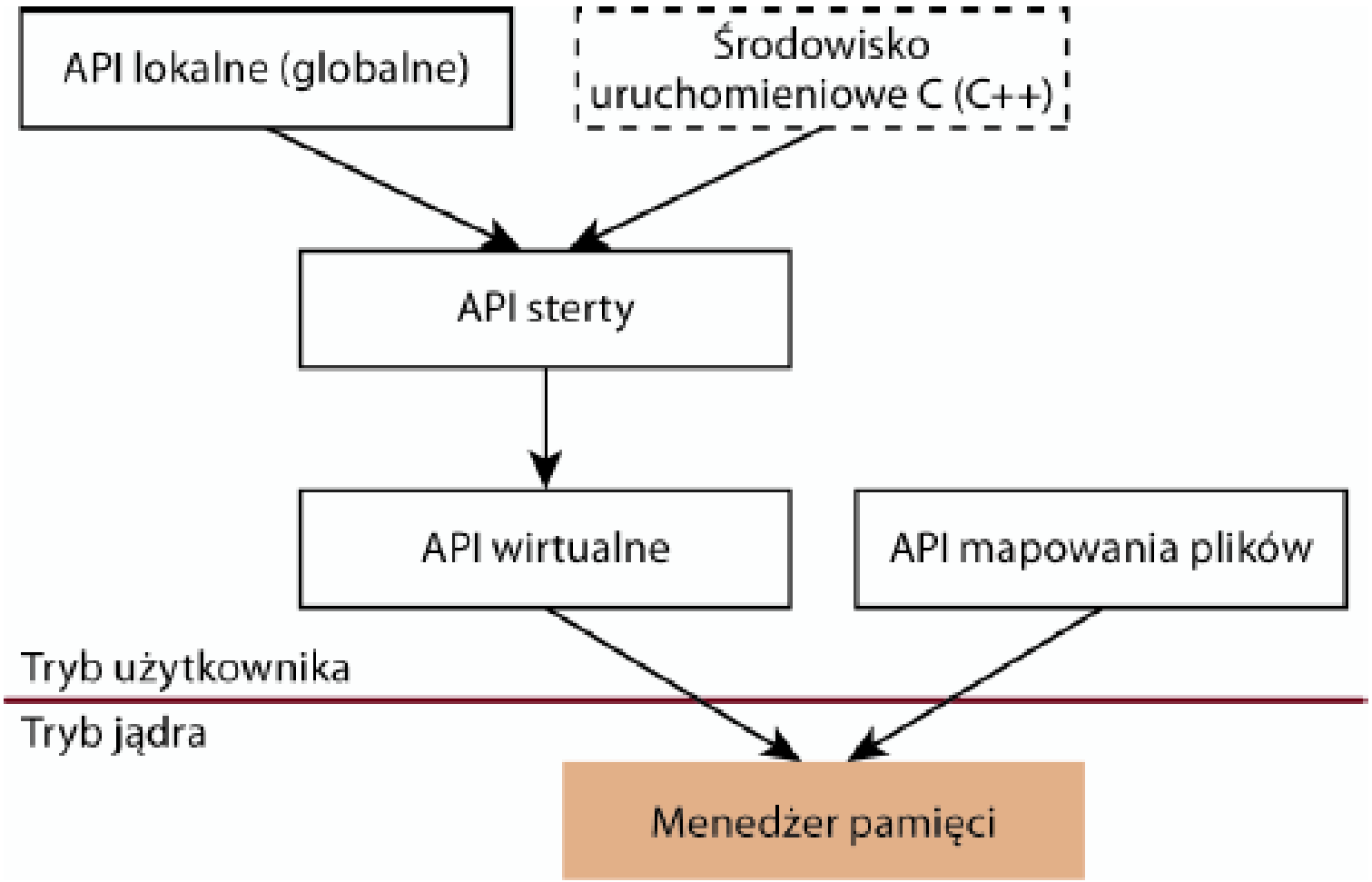
Komponenty Menedżera pamięci

- Menedżer pamięci systemu Windows - część centrum wykonawczego.
- Zestaw systemowych usług wykonawczych.
- Obsługa wykrytych sprzętowo wyjątków w zarządzaniu pamięcią.
- Procedury działające jako wątki trybu jądra w ramach procesu *System*.
  - Menedżer zestawu równowagi
  - Procedura wymiany procesu/stosu
  - Moduł zapisu stron zmodyfikowanych
  - Moduł zapisu stron zmapowanych
  - Wątek dereferencji segmentu
  - Wątek zerowania stron

Usługi menedżera pamięci

- Przydzielanie i zwalnianie pamięci wirtualnej.
- Rozdzielanie pamięci między procesy.
- Odwzorowanie plików do pamięci.
- Przenoszenie stron wirtualnych na dysk.
- Pobieranie informacji o zakresie stron wirtualnych.
- Zmienianie zabezpieczenia stron wirtualnych i blokowanie ich w pamięci.

Usługi menedżera pamięci

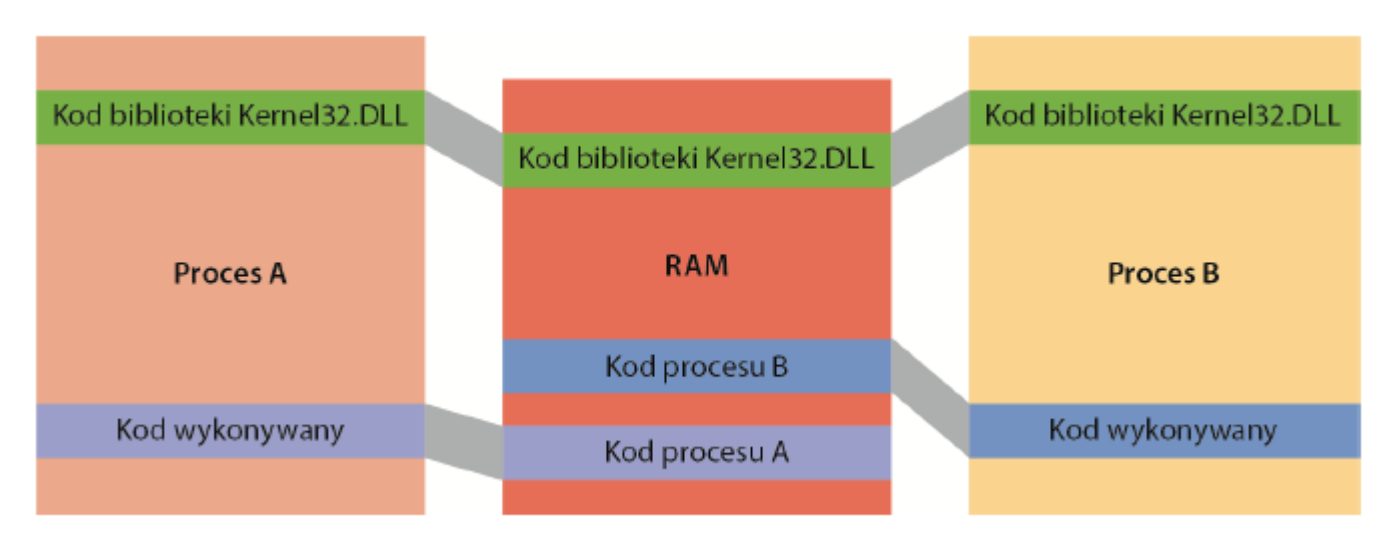


Stany stron i przydzielanie pamięci

- Strony w wirtualnej przestrzeni adresowej procesu mogą być:
  - wolne,
  - zarezerwowane,
  - zadeklarowane,
  - współdzielone.
- Strony prywatne są alokowane za pomocą funkcji menedżera pamięci.
- Strony współdzielone są zwykle mapowane na widok sekcji.
- Sekcje są widoczne w Windows API jako obiekty mapowania plików.

Pamięć współdzielona i pliki mapowane

- *Pamięć współdzieloną* (*shared memory*) to taka, która jest widoczna dla więcej niż jednego procesu lub jest obecna w wirtualnej przestrzeni adresowej więcej niż jednego procesu.
- Podstawowymi składnikami menedżera pamięci służącymi do implementowania współdzielenia pamięci są obiekty sekcji.



Ochrona pamięci

Ochrona jest realizowana na cztery sposoby:

- Wszystkie ogólnosystemowe struktury danych i pule pamięci używane przez komponenty systemu działające w trybie jądra są dostępne tylko w trybie jądra — wątki trybu użytkownika nie mają dostępu do tych stron.
- Każdy proces dysponuje odrębną przestrzenią adresową, zabezpieczoną przed dostępem wątku należącego do innego procesu.
- Sprzętowa ochrona pamięci.
- Dostęp do pamięci współdzielonej mają tylko procesy z odpowiednimi uprawnieniami.

Sterty w trybie jądra - systemowe pule pamięci

- Menedżer pamięci tworzy dwa typy stert, używanych przez komponenty trybu jądra do przydzielania pamięci systemowej:
  - pula niestronicowana - składa się z części systemowych adresów wirtualnych, które zawsze rezydują w pamięci fizycznej i dzięki temu są w każdej chwili dostępne.
  - pula stronicowana - obszar pamięci wirtualnej w przestrzeni systemu, który może być stronicowany.

Listy asocjacyjne

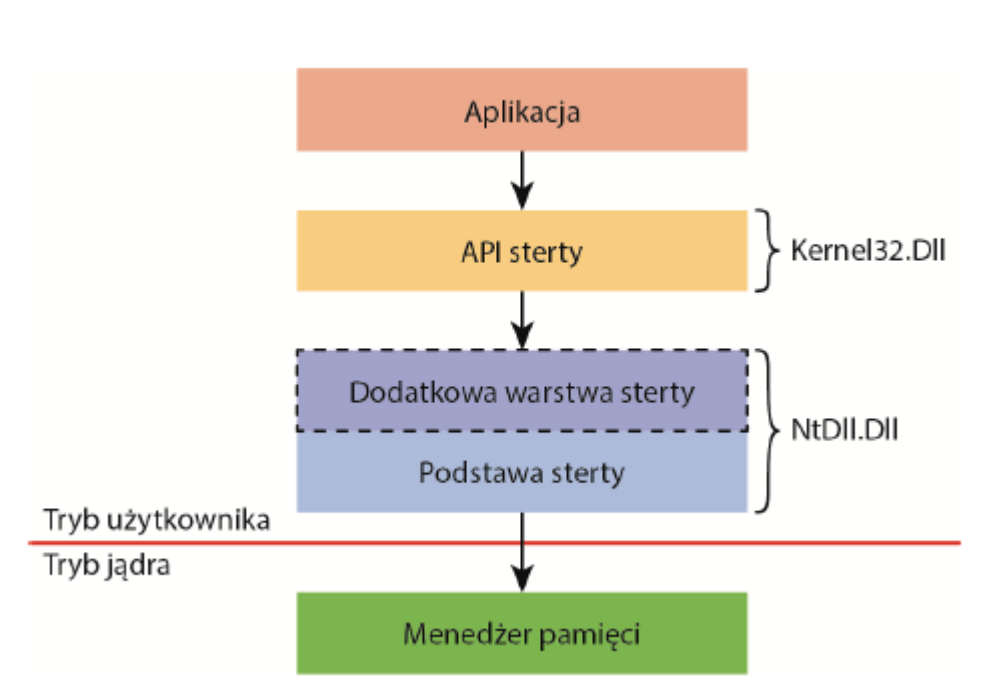
- System Windows ma wbudowany mechanizm - *listy asocjacyjne* (*look-aside lists*), służący do szybkiego przydzielania pamięci.
- Listy asocjacyjne zawierają tylko bloki o ustalonym rozmiarze, natomiast przydziały pul mogą mieć różne rozmiary.
- Pule są bardziej elastyczne pod względem zastosowań, ale listy asocjacyjne są szybsze, gdyż nie używają żadnych blokad.

Sterty procesu

- Każdy proces ma co najmniej jedną stertę — domyślną stertę procesu o rozmiarze 1 MB, którą można powiększyć.
- Sterta domyślna może być wykorzystywana bezpośrednio przez program lub pośrednio przez niektóre wewnętrzne funkcje systemu.

Systemowe pule pamięci

Do wersji systemu Windows 10 i Server 2016 istniał tylko jeden typ sterty - *sterta NT*

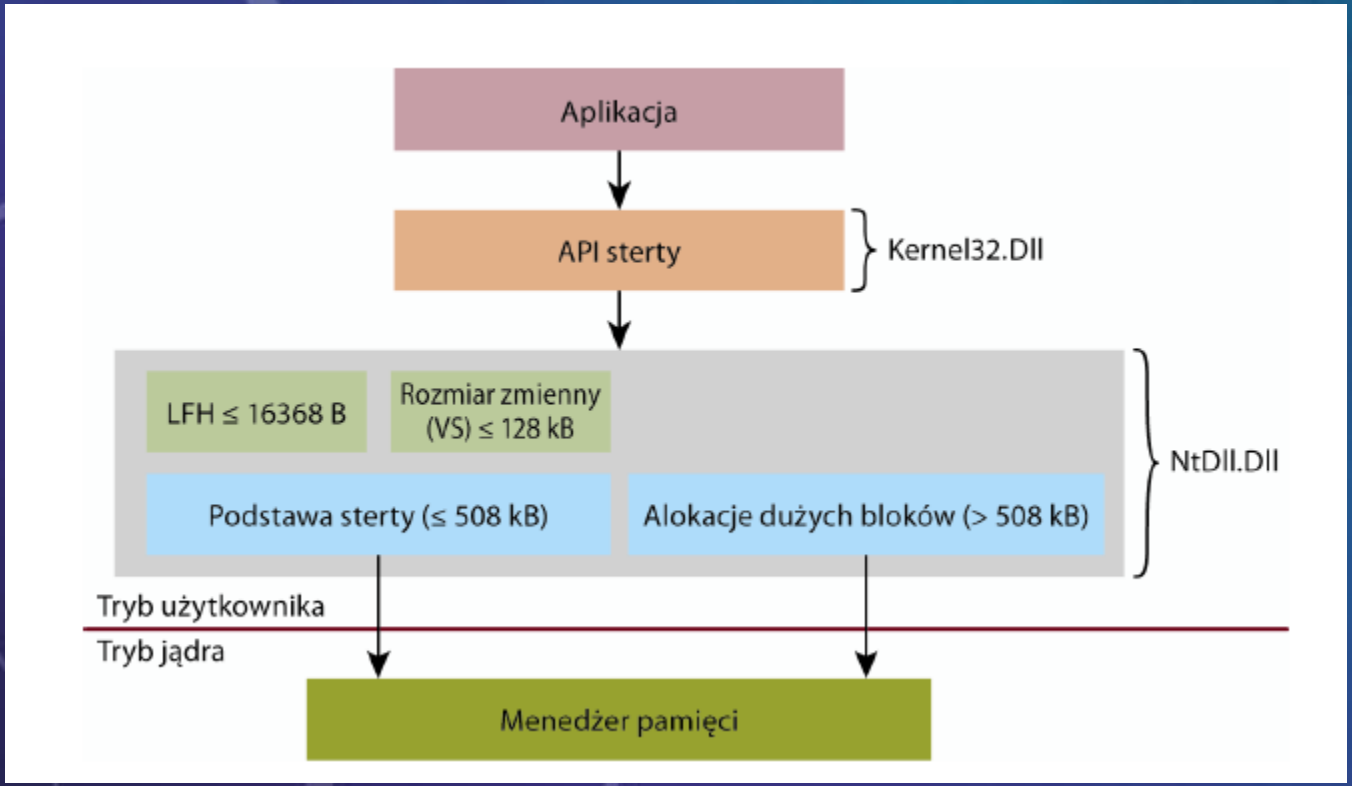


Sterta o małej fragmentacji

- Wiele aplikacji działających w systemie Windows ma stosunkowo małe zapotrzebowanie na pamięć — zwykle mniej niż 1 MB.
- Algorytm najlepszego dopasowania, jaki stosuje menedżer sterty, pomaga w utrzymaniu małej zajętości pamięci przez każdy proces.

Systemowe pule pamięci

W wersji Windows 10 i Server 2016 wprowadzono stertę segmentacji



Bezpieczeństwo stert

W miarę rozwijania menedżera sterty większy nacisk położono na:

- wczesne wykrywanie błędów wykorzystywania sterty,
- łagodzenie skutków potencjalnych zagrożeń związanych ze stertami.

- Osłabianie zagrożeń wynikających z braku odpowiednich zabezpieczeń w aplikacjach.

Układy wirtualnej przestrzeni adresowej

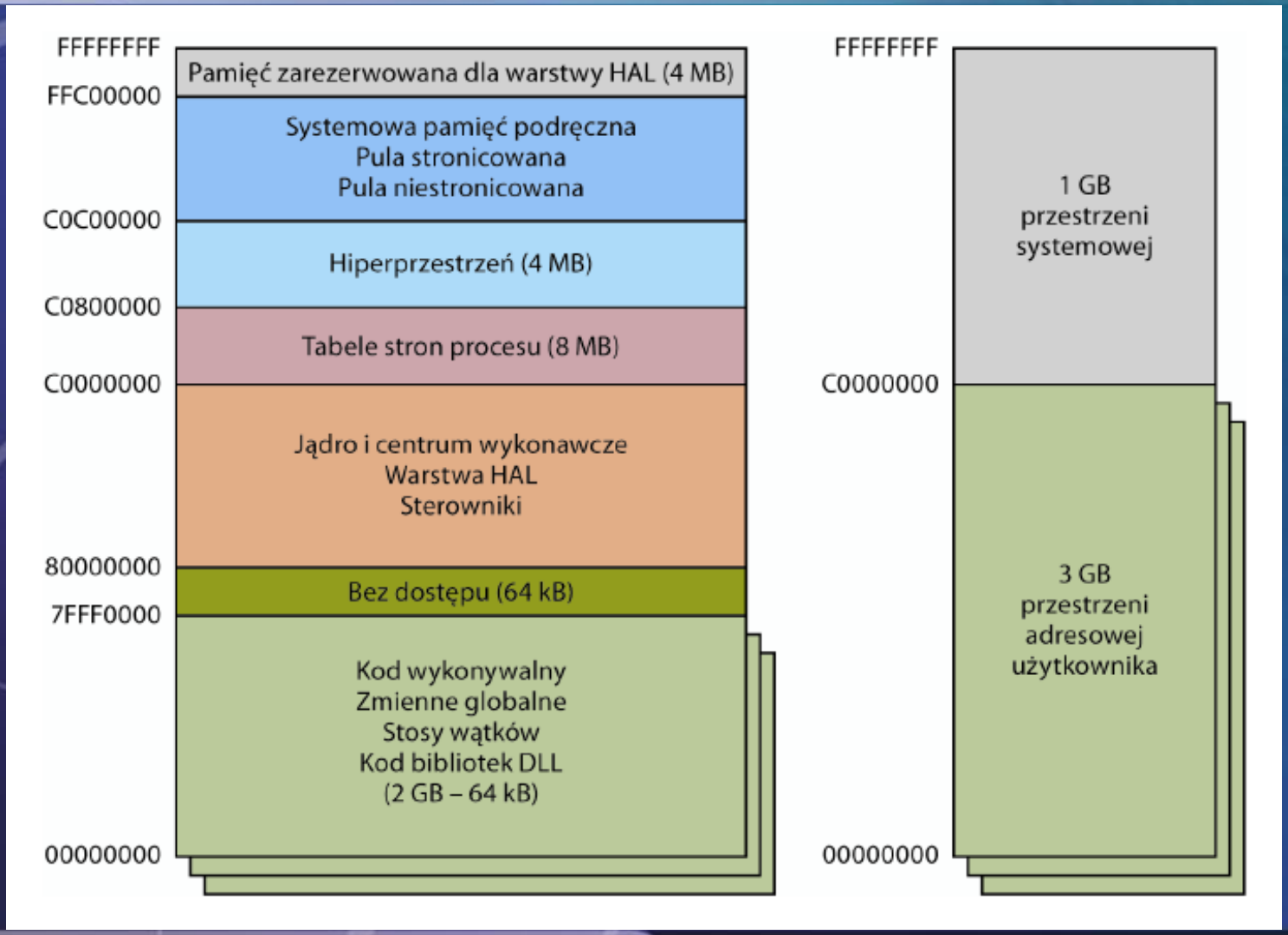
W systemie Windows do wirtualnej przestrzeni adresowej mapowane są następujące typy danych:

- Prywatne dane i kod poszczególnych procesów.
- Ogólnosesyjne dane i kod.
- Ogólnosystemowe dane i kod.

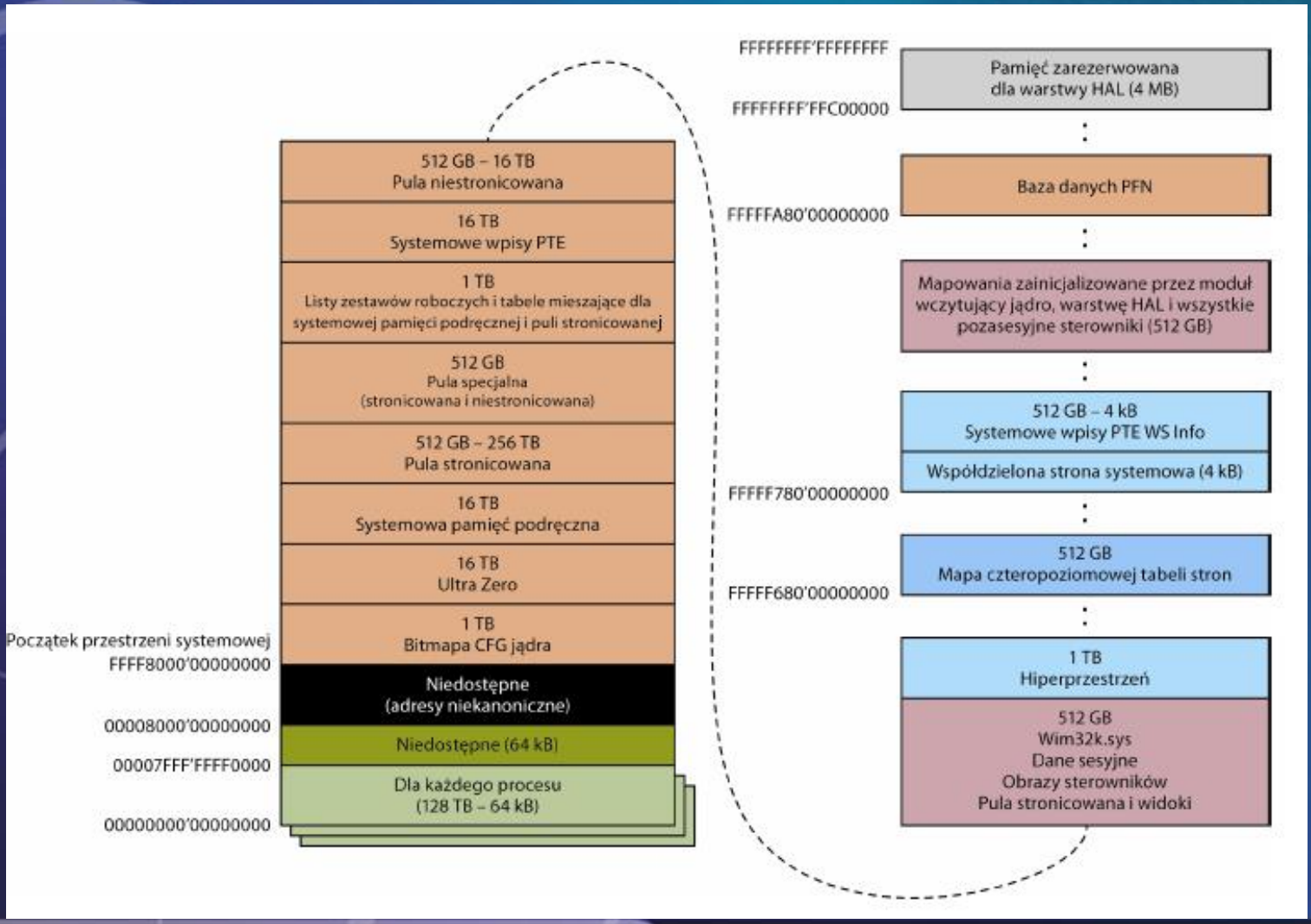
Wirtualna przestrzeń adresowa dzieli się na dwie połowy:

- dolne 128 TB jest przeznaczone na prywatny użytek procesów użytkownika
- górne 128 TB stanowi przestrzeń systemową

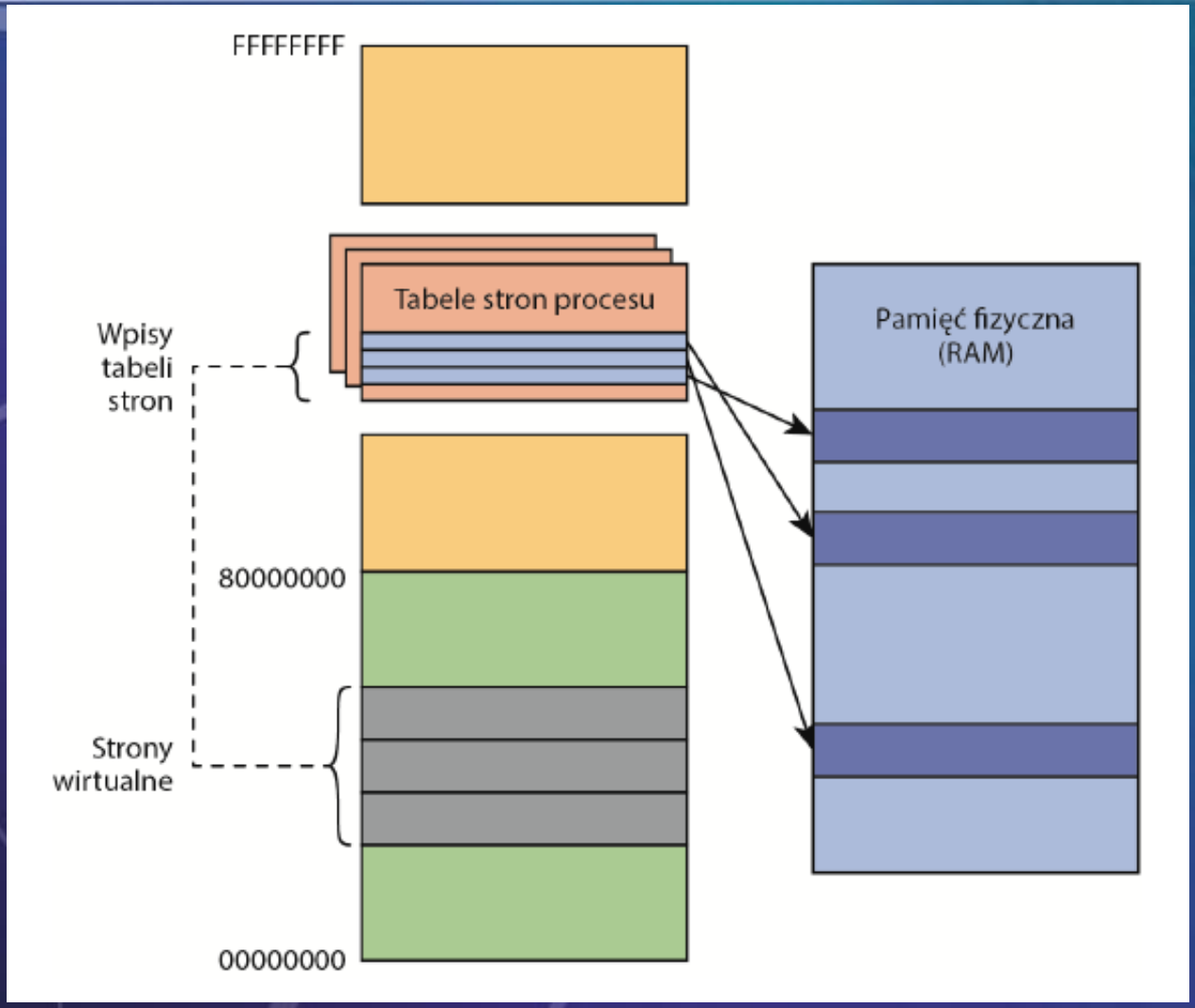
Układy przestrzeni adresowej na platformie x86



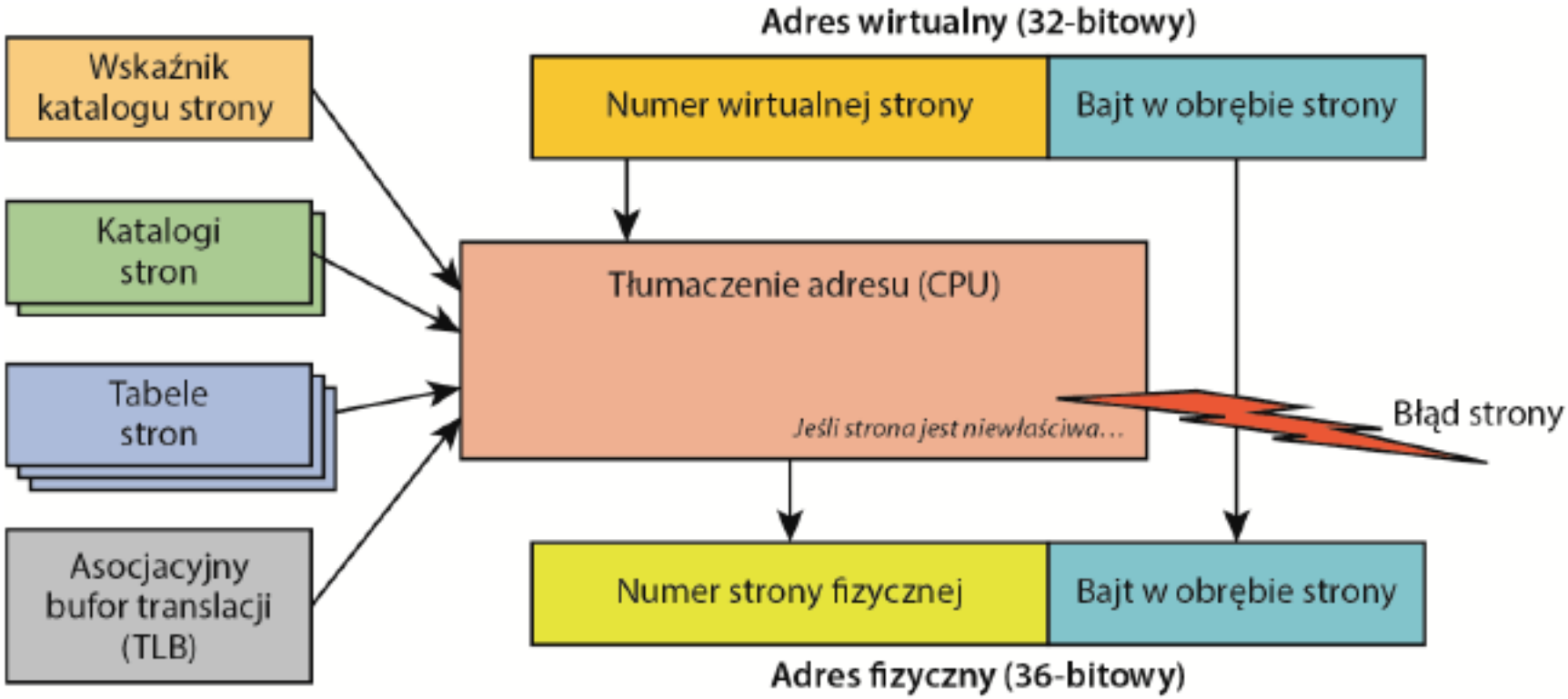
Układy wirtualnej przestrzeni adresowej x64



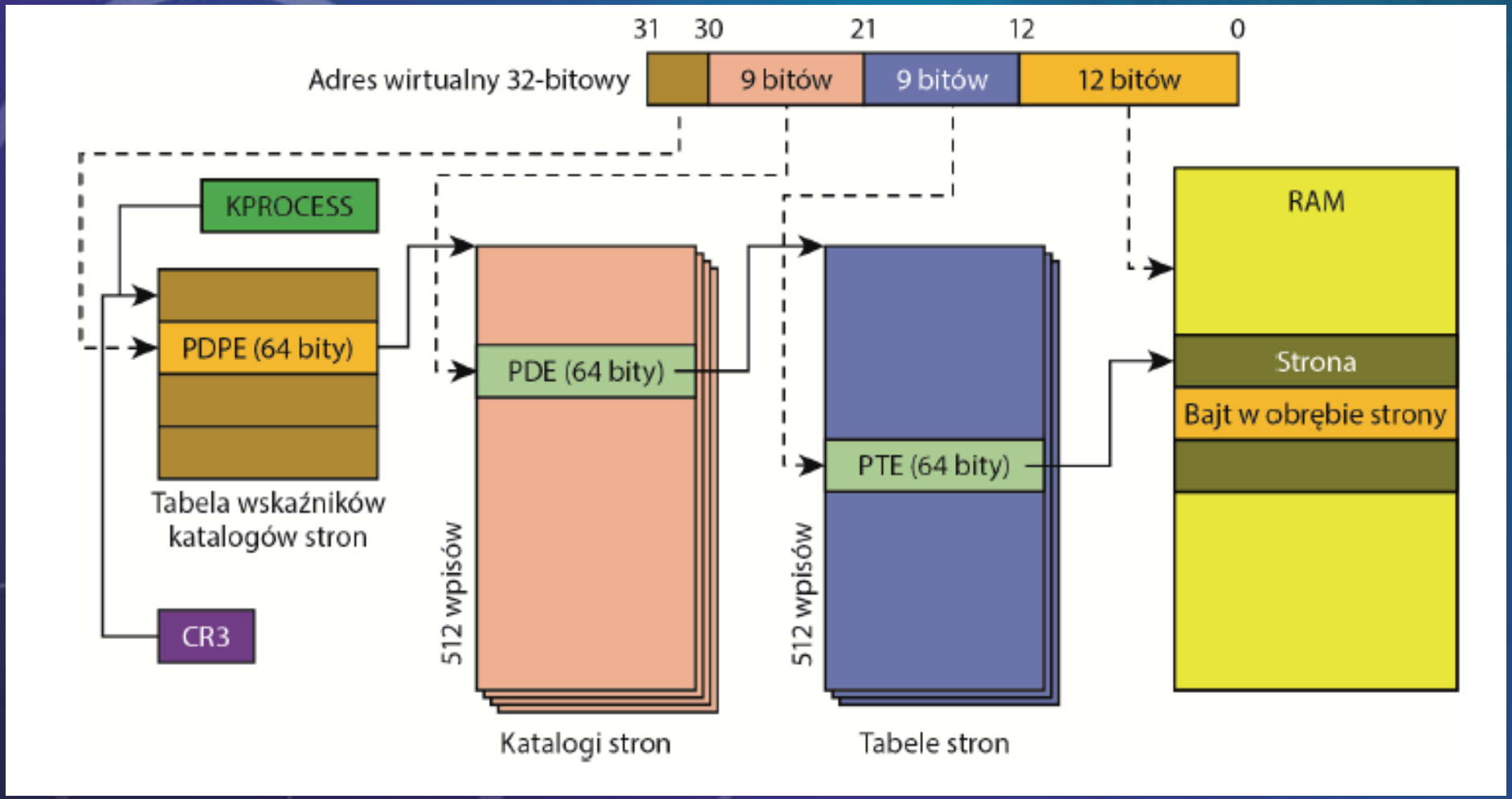
Mapowanie adresów wirtualnych na pamięć fizyczną (x86)



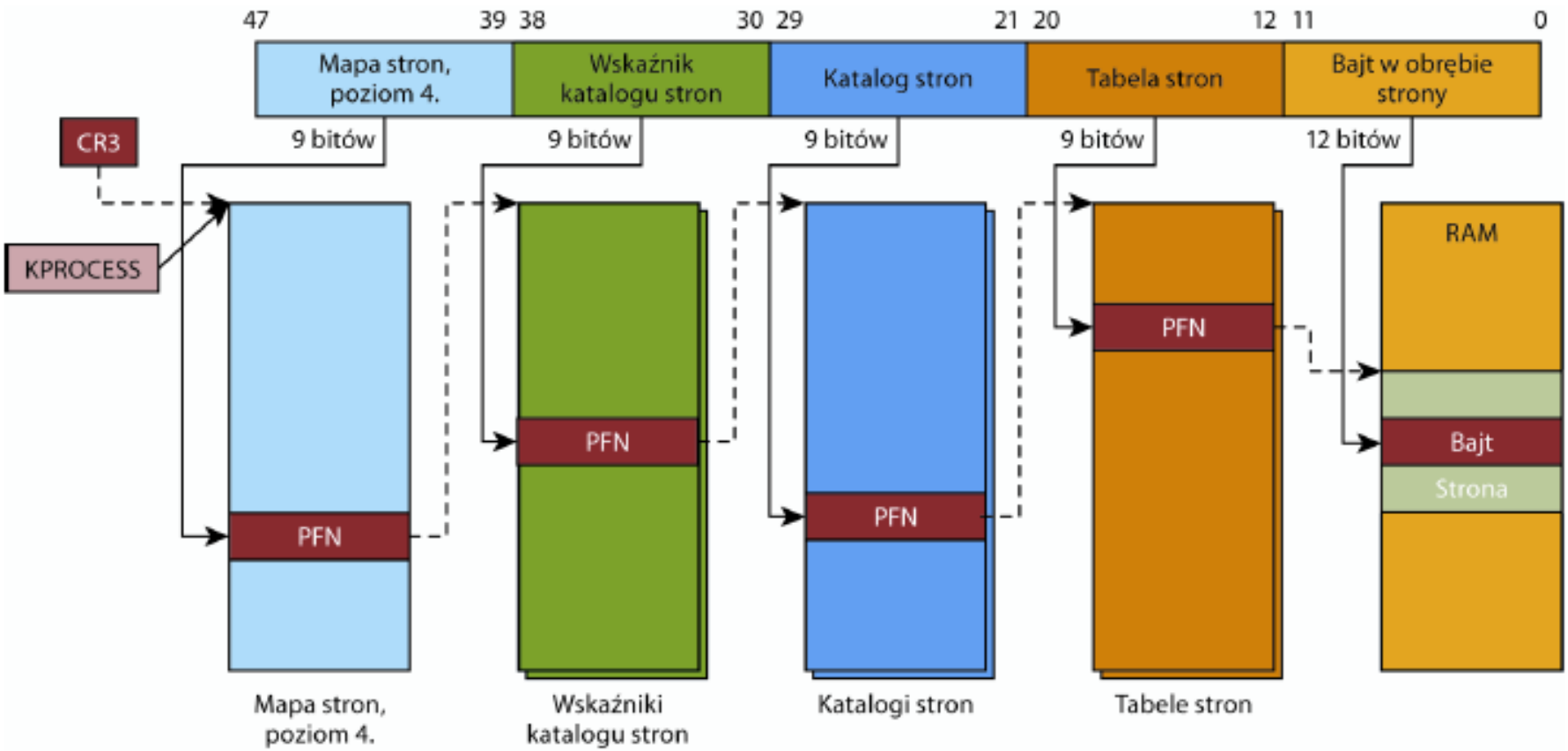
Ogólny schemat procesu tłumaczenia adresów



Tłumaczenie adresu wirtualnego w architekturze x86



Tłumaczenie adresu na platformie x64



Deskryptory adresów wirtualnych

- Menedżer pamięci korzysta z algorytmu stronicowania na żądanie.
- Stronicowanie na żądanie jest formą *wartościowania leniwego* (ang. *lazy evaluation*) — oczekiwania z wykonaniem zadania do momentu, gdy jest ono wymagane.

Deskryptory adresów wirtualnych

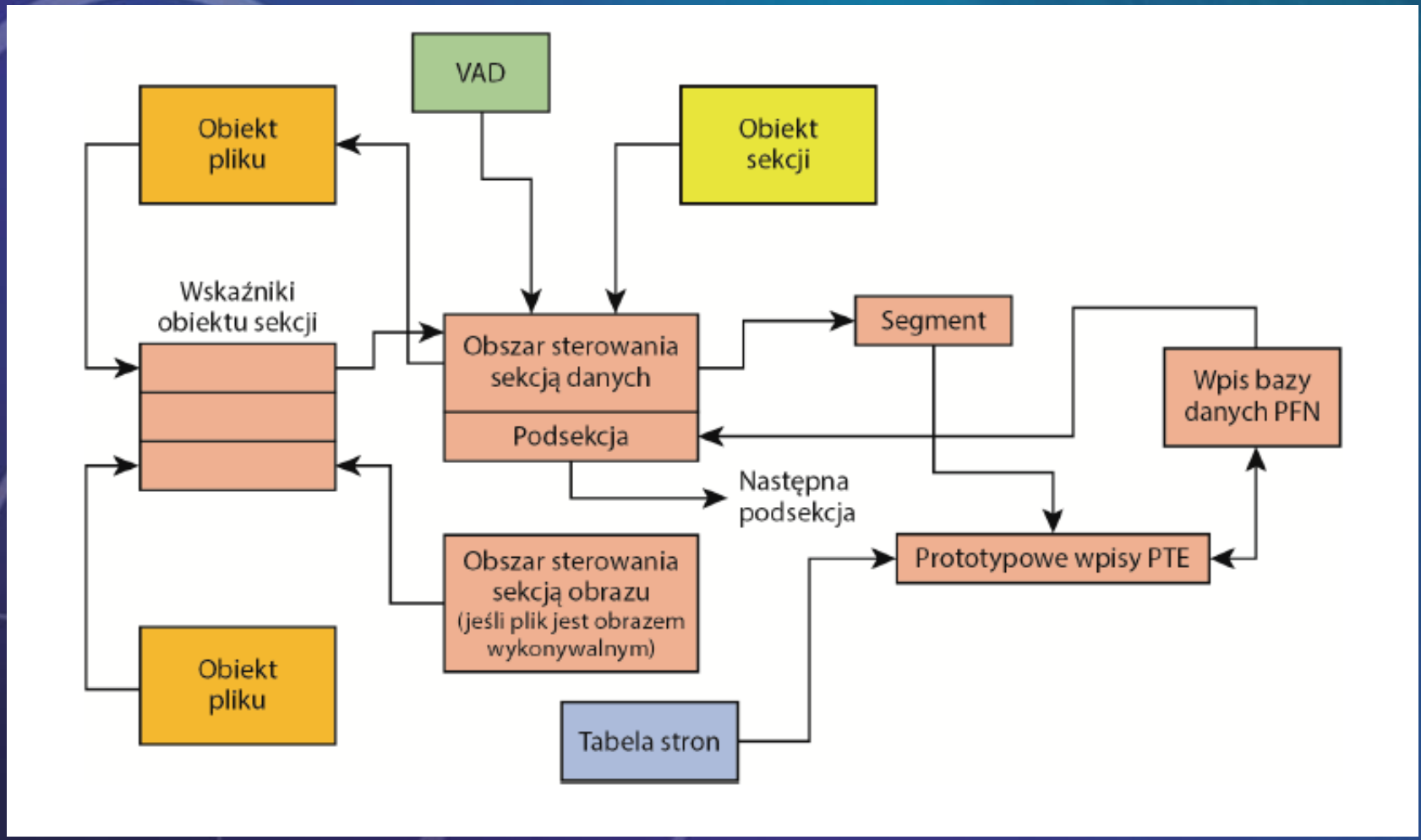
- Dla każdego procesu menedżer pamięci utrzymuje zestaw deskryptorów adresu wirtualnego.
- Są one zorganizowane w formie zrównoważonego drzewa AVL (Adelsona-Velskiego i Landisa).
- Operacje wstawiania, przeszukiwania i usuwania są bardzo szybkie.
- Liczba porównań niezbędnych do wyszukania deskryptora odpowiadającego adresowi wirtualnemu jest minimalna.

Obiekt sekcji

- Obiekt sekcji nazywany jest w podsystemie Windows obiektem mapowania pliku
- Obiekty sekcji, podobnie jak inne obiekty, mogą być przydzielane i zwalniane przez menedżer obiektów.

Typ obiektu	Sekcja
Atrybuty w treści obiektu	Rozmiar maksymalny Zabezpieczenie strony Plik stronicowania lub plik mapowany Osadzona lub nieosadzona
Usługi	Tworzenie sekcji Otwieranie sekcji Rozszerzanie sekcji Mapowanie lub usuwanie mapowania widoku Odpytywanie sekcji

Wewnętrzne struktury sekcji



Zarządzanie pamięcią

- Menedżer pamięci w systemie Windows ładuje strony do pamięci zgodnie z algorytmem stronicowania na żądanie i łączenia w klastry.
- W przypadku błędów braku stron w plikach wykonywalnych klastry składają się z 3 stron, a w pozostałych przypadkach z 7 stron.
- W celu zoptymalizowania rozruchu procesu wprowadzono moduł inteligentnego ładowania stron z wyprzedzeniem.

Zarządzanie pamięcią

Strategia wymiany

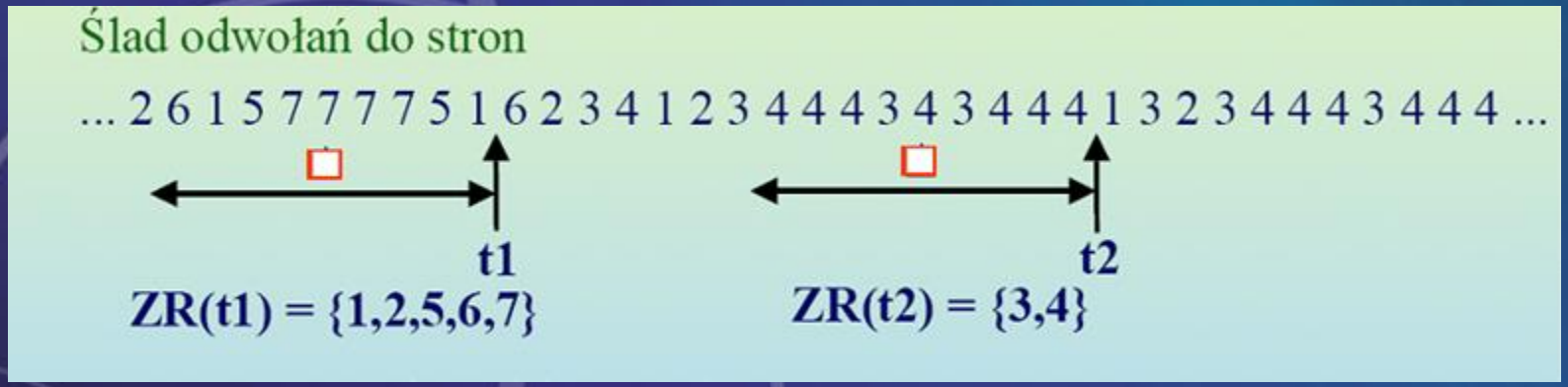
- Algorytmy LRU (*Least Recently Used)* - wymagane jest, aby system pamięci wirtualnej śledził, kiedy strona jest używana - gdy pojawia się żądanie, usuwana jest strona, która najdłużej nie była używana.
- FIFO (*First In,First Out)* - usuwana jest strona, która najdłużej była w pamięci fizycznej.

Strategie wymiany – charakterystyka

- globalne
- lokalne

Model zbioru roboczego

- ***Model zbioru roboczego*** opiera się na założeniu, że program ma charakterystykę strefową (lokalność odwołań).
- ***Okno zbioru roboczego*** □ to ustalona liczba odwołań do stron.
- ***Zbiór roboczy*** to zbiór stron do których nastąpiło □ ostatnich odwołań.



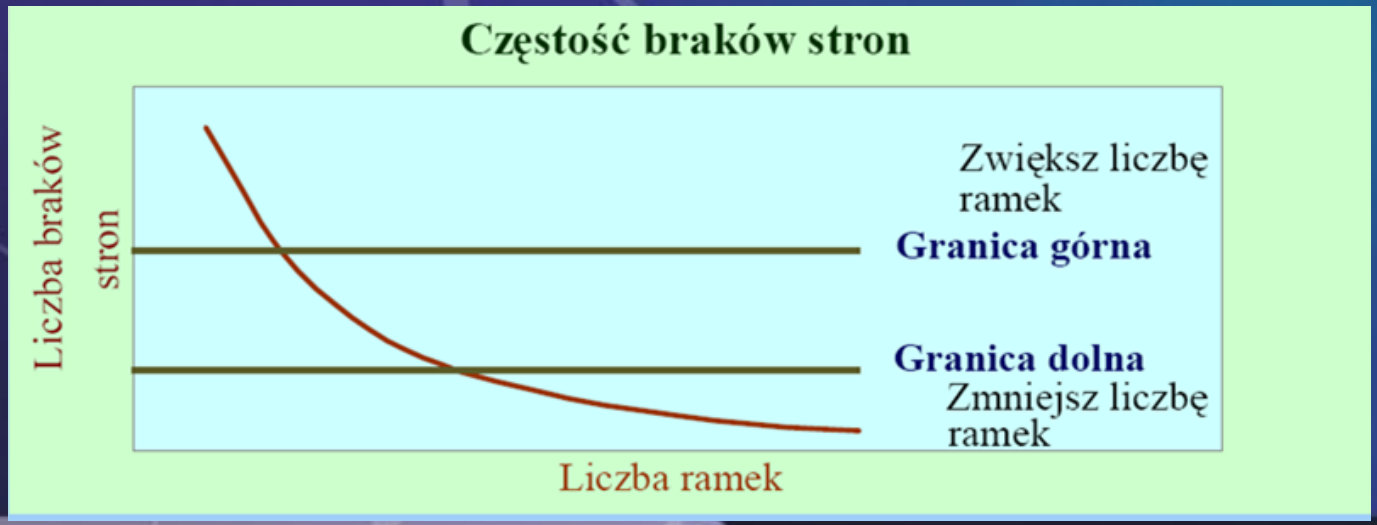
RZR_i – rozmiar zbioru roboczego i-tego procesu

Z – całkowite zapotrzebowanie na ramki $Z=\Sigma RZR_i$

Szamotanie powstaje gdy Z> liczba dostępnych ramek

Częstość braków stron

- Model zbioru roboczego daje dobre rezultaty, jednak nie jest wygodną metodą nadzorowania szamotania.
- Prostszym sposobem jest mierzenie częstości braków stron.
  - Ustala się dolną i górną granicę częstości braków stron.
  - Jeśli proces przekracza górną granicę, przydziela mu się dodatkową ramkę (w przypadku niedoboru ramek można wstrzymać jakiś proces).
  - Jeżeli częstość braku stron procesu spada poniżej dolnej granicy, odbiera mu się ramkę.



Zarządzanie zestawami roboczymi

- Każdy proces rozpoczyna działanie z domyślnym zbiorem roboczym, który może zawierać od 50 do 345 stron.
- Menedżer pamięci może zezwolić procesowi na rozszerzanie zbioru roboczego poza górną granicę lub zmniejszenie go poniżej dolnej granicy.
- Menedżer zbiorów roboczych, przed usunięciem strony, sprawdza bit używalności w tablicy stron - jeśli jest wyzerowany, strona jest uważana za kandydatkę do postarzenia – na podstawie określonego wieku (zwiększenie licznika) następuje kwalifikowanie stron do usunięcia ze zbioru roboczego.

Systemowe zbiory robocze

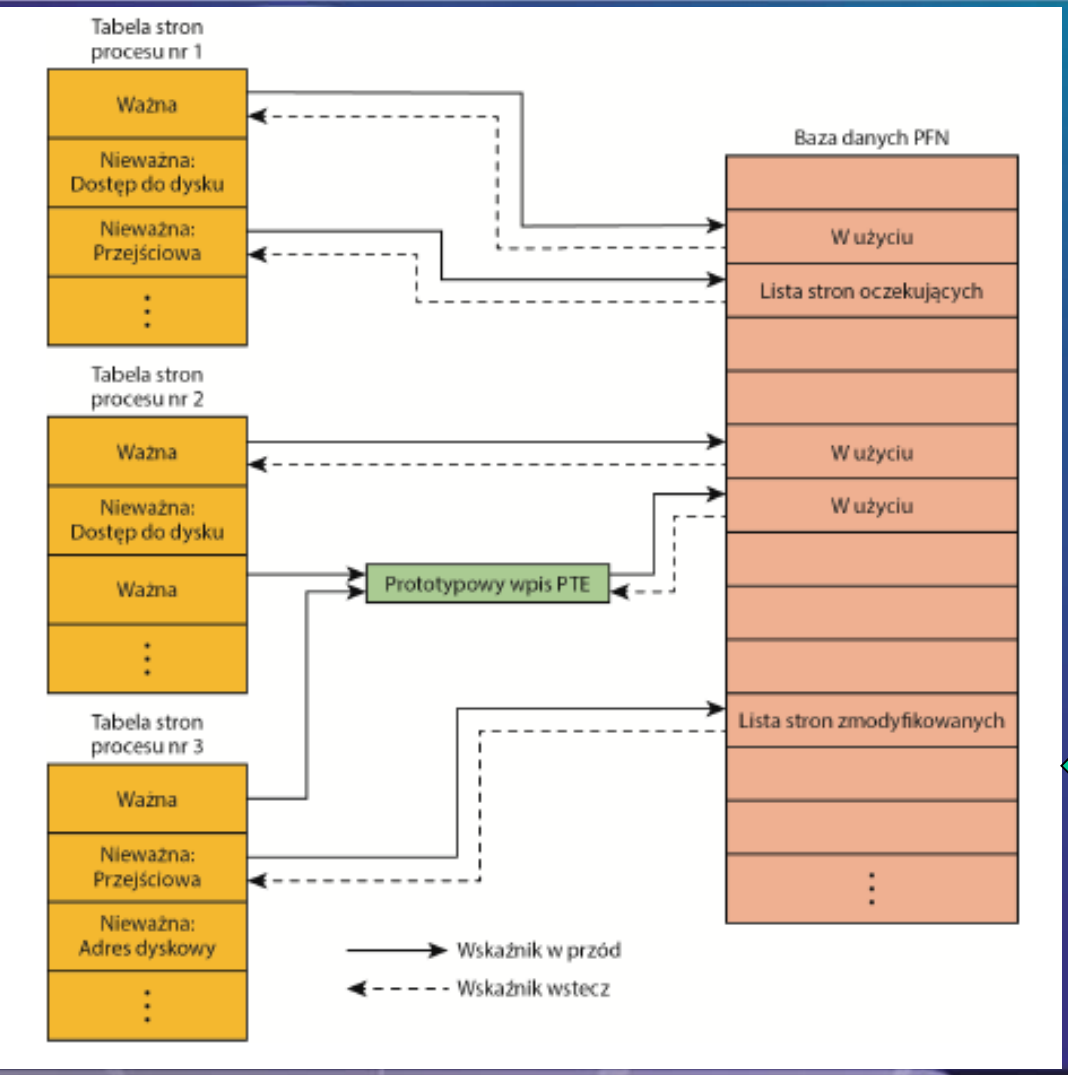
- Kod i dane możliwe do stronicowania są zarządzane przez trzy globalne zbiory robocze nazywane ogólnie systemowymi zbiorami roboczymi. Są to:
  - **Zbiór roboczy systemowej pamięci podręcznej -** zawiera strony rezydujące w systemowej pamięci podręcznej.
  - **Zbiór roboczy puli stronicowanej** - zawiera strony rezydujące w puli stronicowanej.
  - **Zbiór roboczy wpisów PTE -** zawiera kod i dane z załadowanych sterowników, które są stronicowane oraz obraz jądra i strony z sekcji zmapowanych na przestrzeń systemową.

Baza danych numerów stron pamięci

Stany stron w pamięci fizycznej:

- Aktywna (ważna)
- Przejściowa
- Oczekująca
- Zmodyfikowana
- Zmodyfikowana bez zapisu
- Wolna
- Wyzerowana
- ROM
- Wadliwa

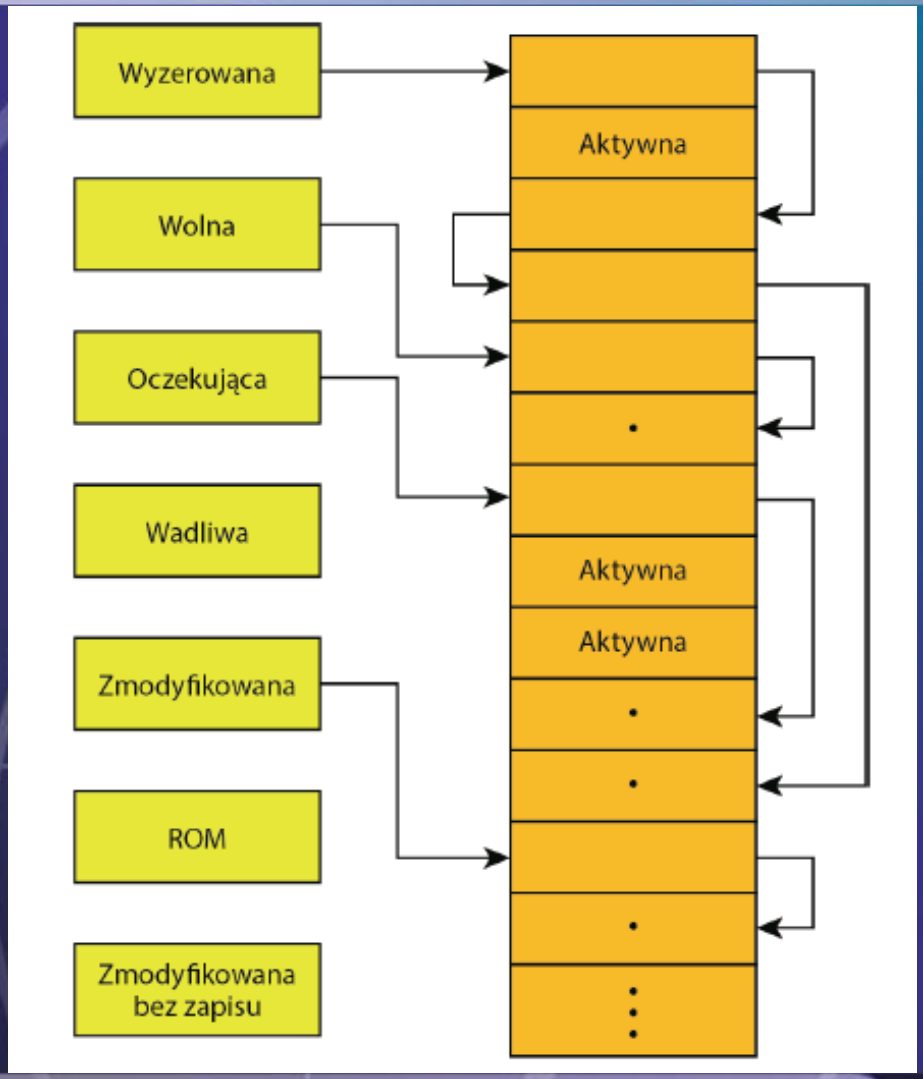
Baza danych numerów stron pamięci



- Baza danych PFN składa się z tablicy struktur reprezentujących poszczególne strony pamięci fizycznej.

Powiązania bazy z tabelami stron

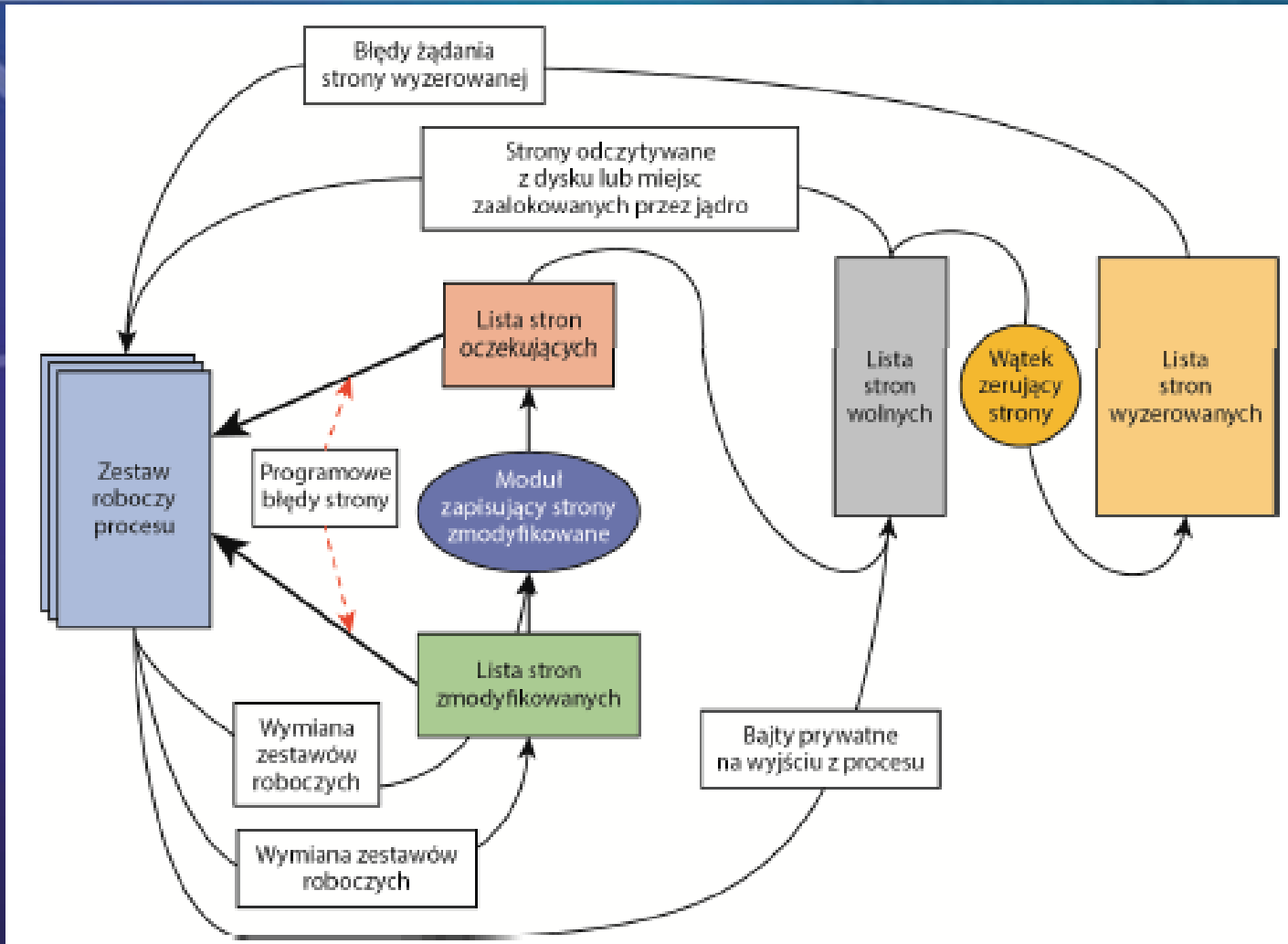
Baza danych numerów stron pamięci



Spośród stanów stron sześć jest zorganizowanych w połączone listy, dzięki czemu menedżer pamięci może szybko zlokalizować strony określonego typu.

Dynamika list stron

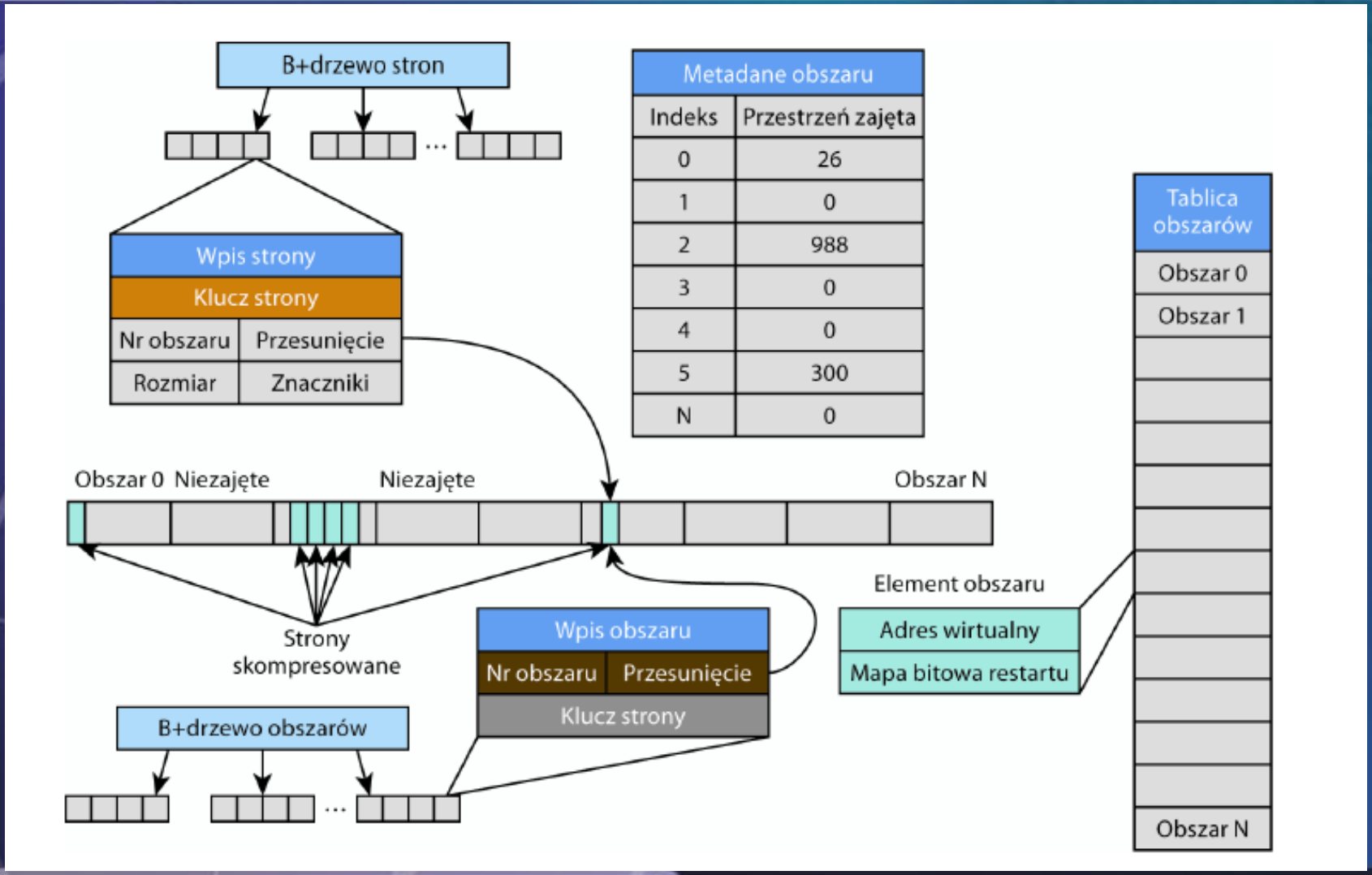
Przemieszczanie ramek stron między listami stronicowania



Priorytet stronicowy

- Każda strona pamięci fizycznej ma w systemie priorytet stronicowy przypisany przez menedżer pamięci - liczba z zakresu od 0 do 7.
- Priorytet pozwala określić kolejność pobierania stron z listy stron oczekujących.
- Menedżer pamięci dzieli listę na osiem części - każda zawiera strony o takim samym priorytecie.
- Pobierana jest strona z podlisty o najniższym priorytecie.
- Priorytety stronicowe są przypisywane także wszystkim wątkom oraz procesom - strona otrzymuje priorytet taki sam jak priorytet stronicowy wątku.
- Domyślnie procesy otrzymują priorytet stronicowy o wartości 5.

Architektura kompresji



Proaktywne zarządzanie pamięcią - SuperFetch

- W wersjach klienckich systemu Windows wprowadzono mechanizm SuperFetch - usprawnienie w zarządzaniu pamięcią fizyczną.
- Do zastępowania najdawniej używanej strony dodana została informacja na temat historii dostępu do pliku.
- Mechanizm SuperFetch zawiera następujące komponenty:
    - moduł śledzenia,
    - moduły zbierania i przetwarzania danych śledzenia,
    - agenty,
    - menedżer scenariuszy.

Proaktywne zarządzanie pamięcią - SuperFetch

- Jednym z aspektów mechanizmu SuperFetch jest obsługa scenariuszy - specjalne działania mające na celu zwiększenie wygody użytkownika:
  - Hibernacja
  - Czuwanie
  - Szybkie przełączanie użytkowników

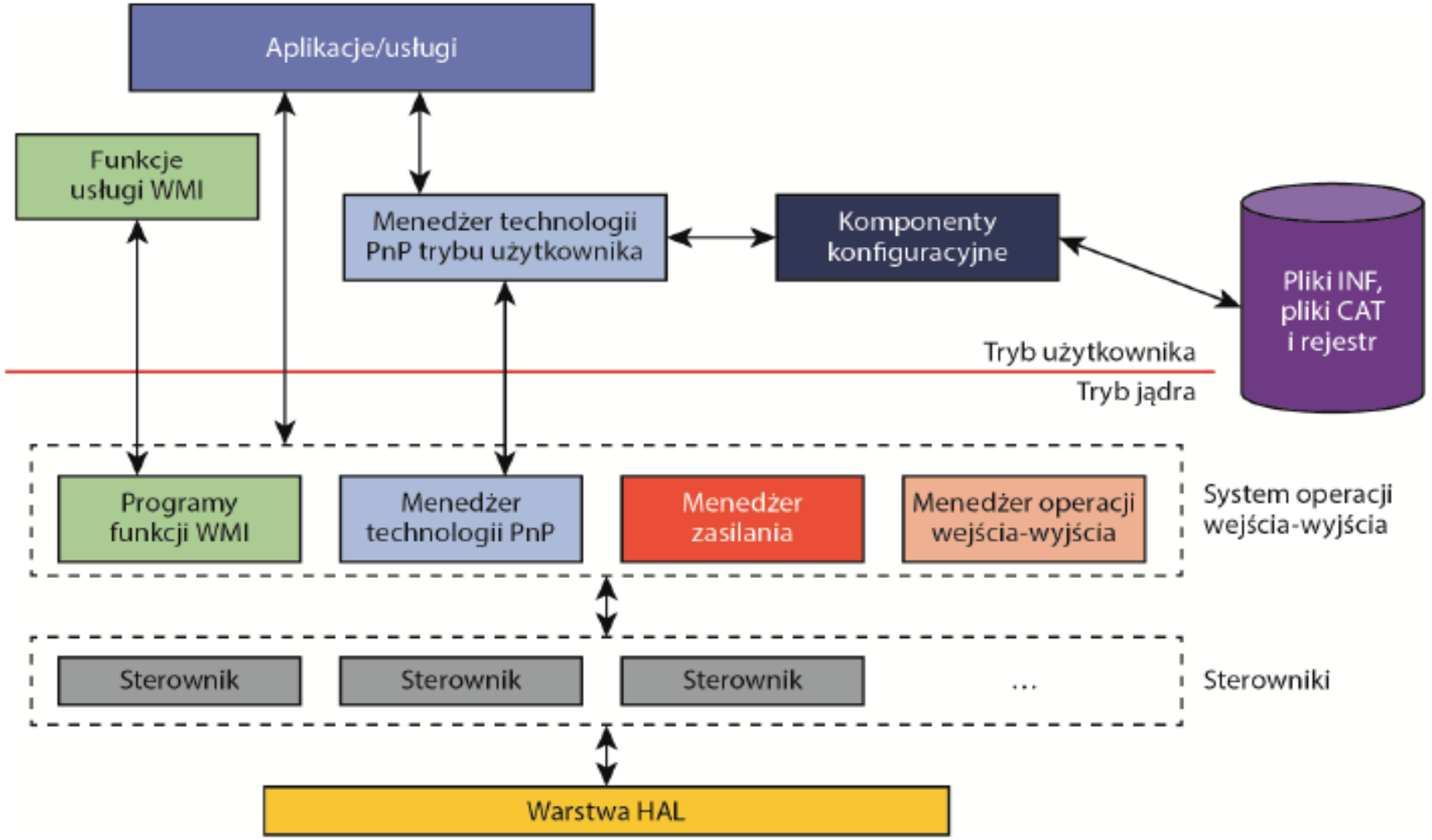
Mechanizm - SuperFetch

- Mechanizm SuperFetch podejmuje większość decyzji na podstawie informacji, które zostały uzyskane przez scalanie, analizowanie i obrabianie nieprzetworzonych śladów oraz wpisów dziennika.
- Moduł śledzenia dodatkowo korzysta z tradycyjnych mechanizmów „postarzania” stron, które są wbudowane w menedżer pamięci.
- Funkcja SuperFetch stale pobiera dane śledzenia z systemu:
  - wykorzystanie stron,
  - dostęp do stron za pośrednictwem mechanizmów menedżera pamięci,
  - kontrola bitu dostępu
  - powiększanie wieku zestawu roboczego.

System WE/WY

- System Windows zapewnia aplikacjom abstrakcję urządzeń, zarówno fizycznych, jak i programowych.
  - Jednakowe zabezpieczenia i nazewnictwo.
  - Asynchroniczne, pakietowe operacje wejścia-wyjścia o dużej wydajności.
  - Usługi pozwalające na tworzenie sterowników.
  - Dynamiczne ładowanie sterowników urządzeń i usuwanie ich z pamięci.
  - Obsługa technologii Plug and Play.
  - Obsługa zarządzania zasilaniem.
  - Obsługa wielu możliwych do zainstalowania systemów plików.

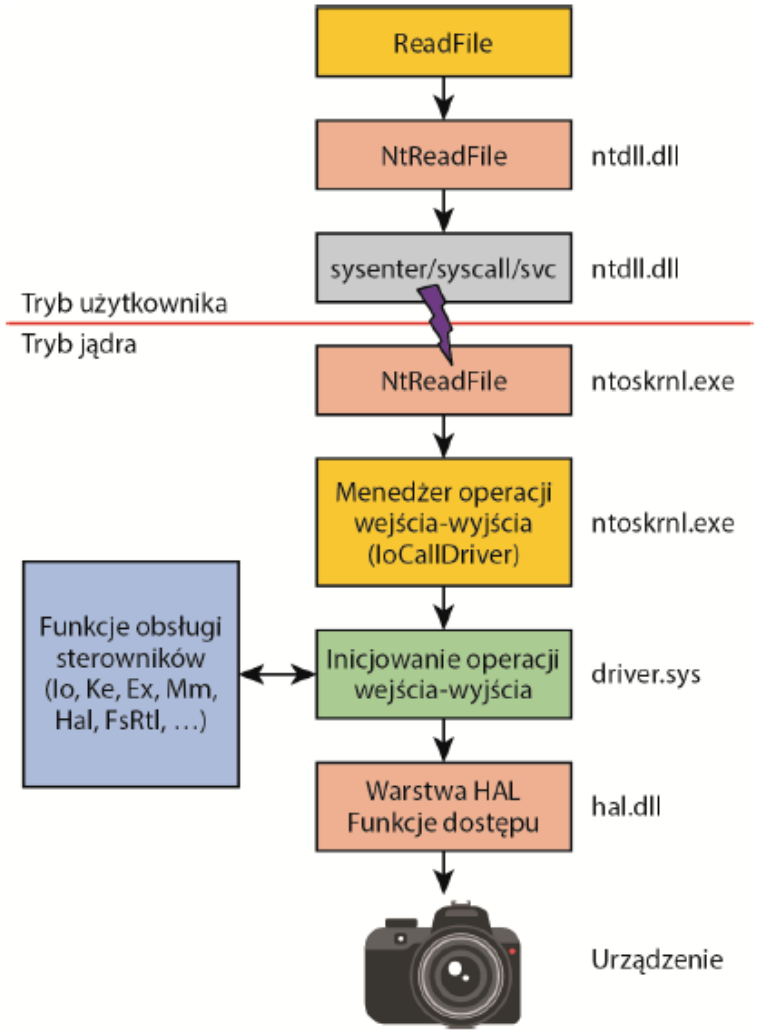
Komponenty systemu operacji wejścia-wyjścia



Menedżer operacji wejścia-wyjścia

- Menedżer operacji wejścia-wyjścia to jądro systemu operacji wejścia-wyjścia.
- Działanie systemu operacji wejścia-wyjścia bazuje na pakietach.
- Menedżer operacji wejścia-wyjścia tworzy pakiet IRP w pamięci w celu reprezentowania operacji wejścia-wyjścia.
- Sterownik odbiera pakiet IRP - wykonuje określoną w nim operację i przekazuje pakiet z powrotem do menedżera operacji wejścia-wyjścia.
- Systemu Windows udostępnia kilka zaawansowanych funkcji, takich jak asynchroniczne, bezpośrednie i buforowane operacje wejścia-wyjścia.

Przepływ typowego żądania operacji WE/WY



- System operacyjny dokonuje abstrakcji wszystkich żądań operacji wejścia-wyjścia jako operacji względem pliku wirtualnego, ponieważ menedżer operacji wejścia-wyjścia dysponuje informacjami wyłącznie o plikach.
- Na sterowniku spoczywa odpowiedzialność za dokonanie translacji komentarzy zorientowanych plikowo do postaci poleceń konkretnych urządzeń.

Poziomy żądań przerwania

Poziom IRQL ma dwa trochę różne znaczenia:

1. Poziom IRQL to priorytet przypisywany do źródła przerwania pochodzącemu z urządzenia - numer priorytetu ustawiany jest przez warstwę HAL.

2. Każdy procesor ma własną wartość poziomu IRQL.