

Implementacja pamięci wirtualnej w Windows NT

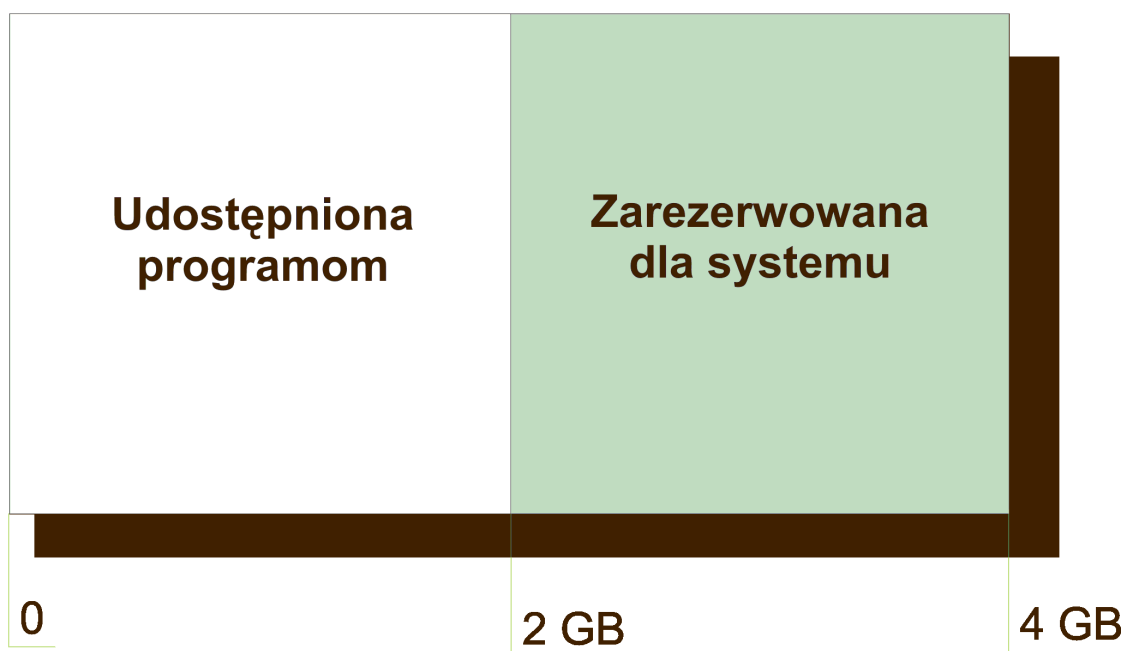
autor: Bolek Szewczyk

<http://rainbow.mimuw.edu.pl/~bs181296/studia/systemy-operacyjne/Windows-NT-VMM.pdf>

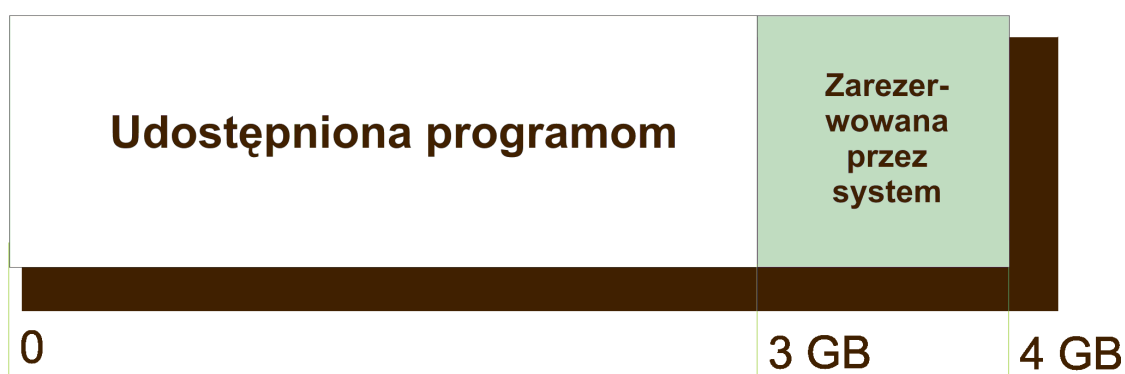
Wprowadzenie

Zarządzanie pamięcią wirtualną w Windows NT w niczym nie przypomina zarządzania pamięcią wirtualną poprzednich wersji Windows. Poprzednie wersje systemu adresują pamięć za pomocą pary liczb, segmentu i przesunięcia, podczas gdy Windows NT używa adresów 32-bitowych. Na takich adresach można łatwo wykonywać operacje arytmetyczne – jest to adresowanie ciągłe.

32-bitowa przestrzeń adresowa Windows NT

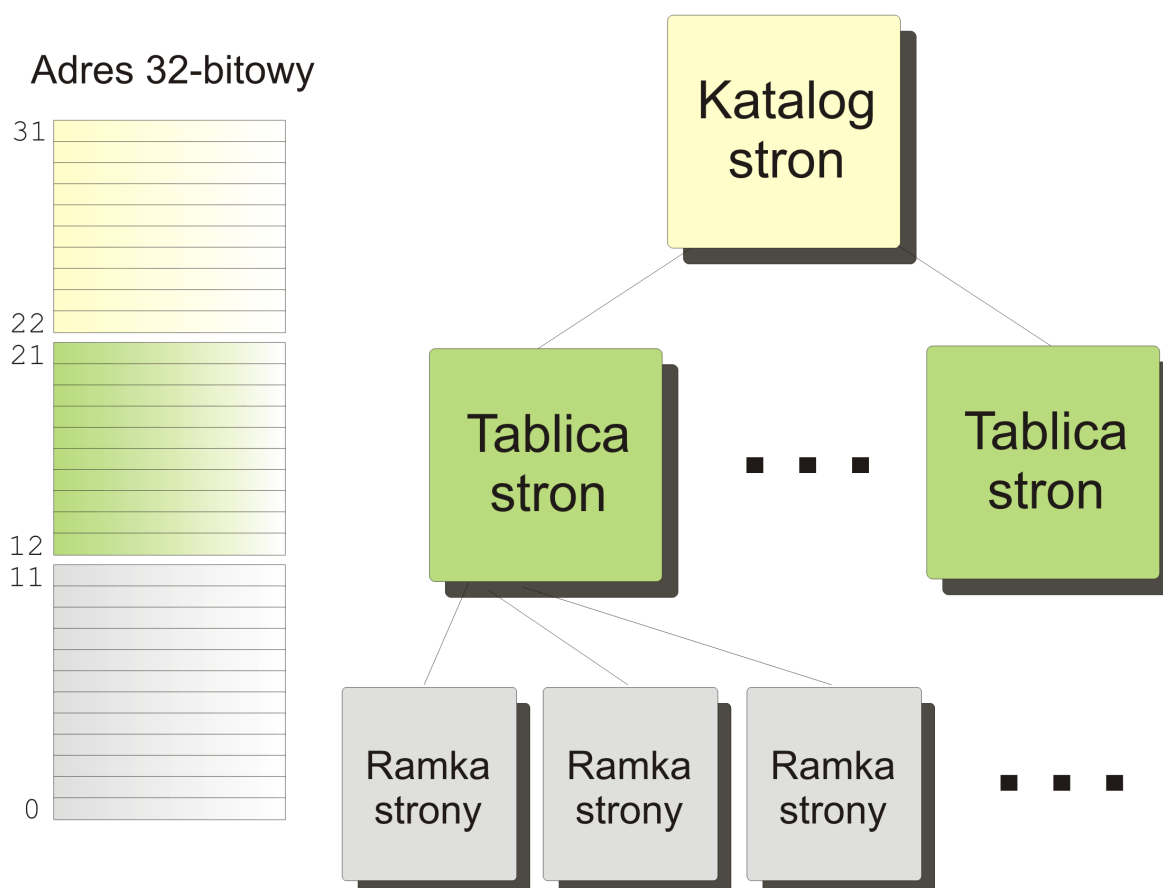


Przestrzeń adresowa Windows NT 4.0



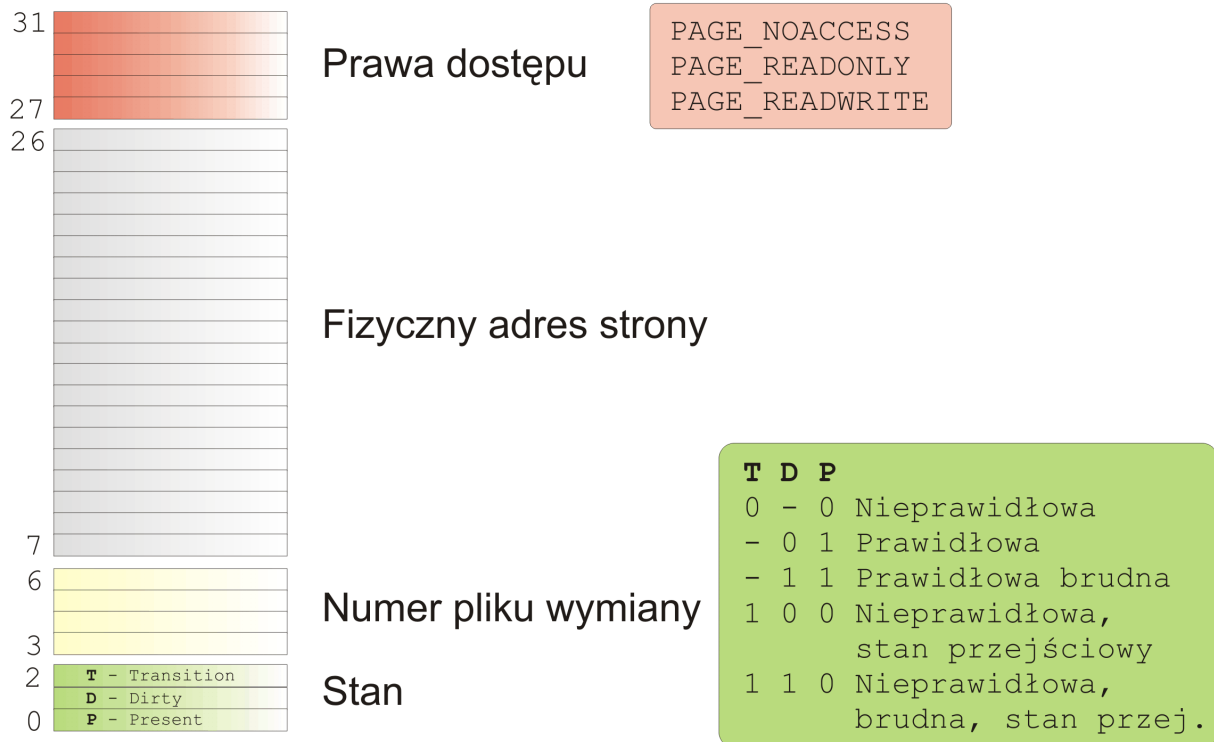
Programy w Windows NT mają dostęp do 32-bitowej liniowej przestrzeni adresowej. Każdy adres w pamięci jest słowem 32-bitowym. Przestrzeń adresowa każdego procesu działającego w Windows NT jest podzielona na dwie części: zarezerwowaną przez system operacyjny oraz dostępną dla aplikacji. Adresy dostępne dla aplikacji mają zgaszony najwyższy bit, czyli są mniejsze niż 2 gigabajty. W części zarezerwowanej znajdują się wspólne dla wszystkich procesów procedury i dane systemu operacyjnego. W nowszych systemach operacyjnych firmy Microsoft, np. Windows NT 4.0 Enterprise Edition, część zarezerwowana przez system operacyjny stanowi już tylko jedną czwartą 32-bitowej przestrzeni adresowej.

32-bitowe adresowanie wirtualne



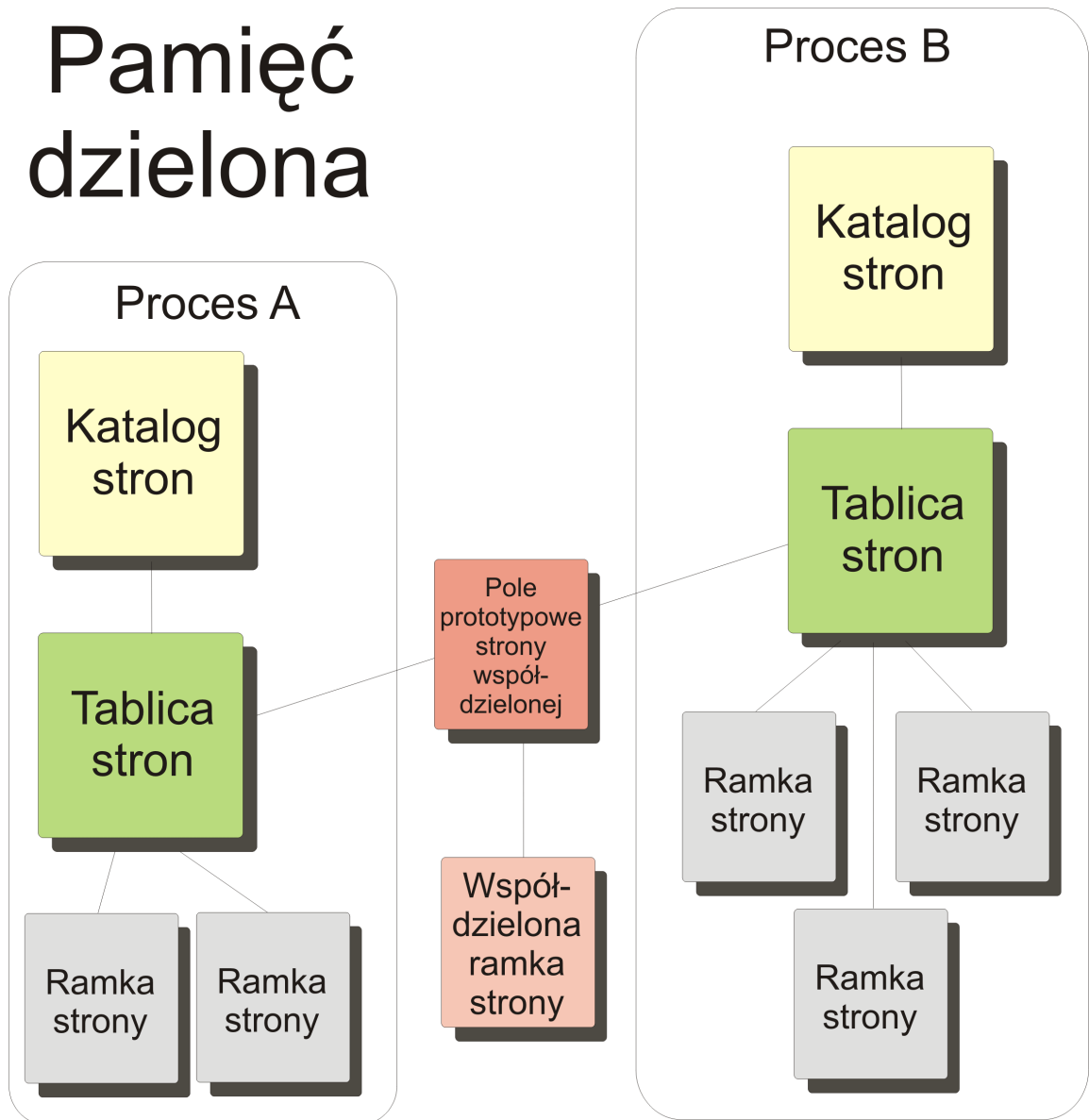
Pamięć jest podzielona na strony o rozmiarze 4 kilobajtów (nie dotyczy to Windows na Alpha i innych platformach, gdzie strona może mieć rozmiar np. 8 kilobajtów). Każda strona jest albo katalogiem stron, albo tablicą stron, albo ramką. 32-bitowy adres w Windows NT składa się z trzech grup bitów umożliwiających znalezienie faktycznego adresu fizycznego. 10 najważniejszych bitów adresu używanych jest jako indeks w katalogu stron, z którego pobierana jest 32-bitowa wartość. Katalog stron jest przypisany do procesu. Wartość pobrana z katalogu służy do identyfikacji tablicy stron. Kolejne 10 bitów adresu zostaje użyte jako indeks w tablicy stron. Liczba pobrana z tablicy stron służy do identyfikacji ramki. Pozostałe 12 bitów stanowi przesunięcie względem początku strony fizycznej, opisywanej przez wybraną ramkę. Tłumaczenie 32-bitowego adresu wirtualnego na adres fizyczny przypomina przechodzenie B-drzewa – wykonywana jest stała ilość odwołań do pamięci, równa głębokości drzewa. Jak łatwo policzyć, tablice stron potrzebne do zaadresowania całych 4 GB pamięci zajęłyby 4 MB pamięci, dlatego są one tworzone w razie potrzeby. Znajdowanie adresu fizycznego strony jest zoptymalizowane przy użyciu sprzętowego bufora TLB (Translation Lookaside Buffer).

Pole opisu strony



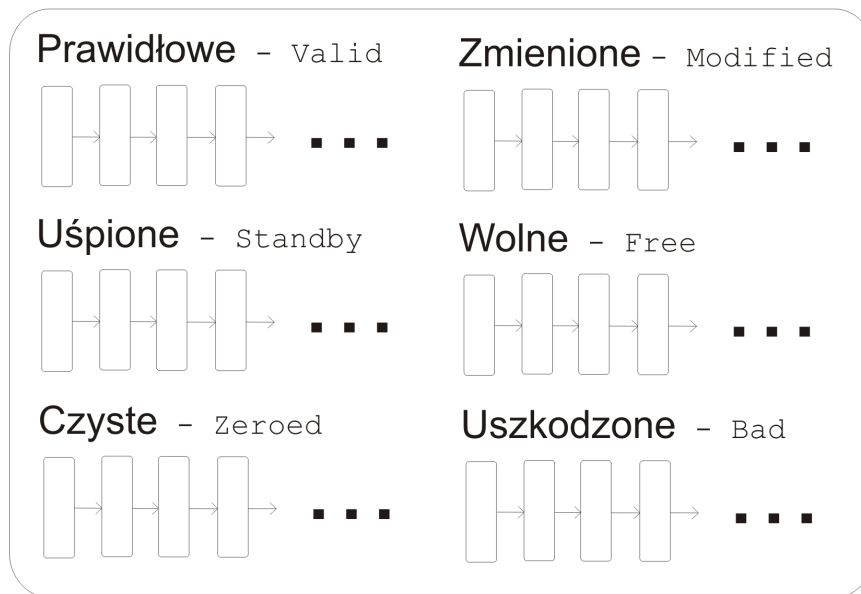
Podczas procesu tłumaczenia adresu z tablicy stron pobierana jest 32-bitowa liczba – jest to pole opisu strony. Pole opisu strony zawiera na pięciu najwyższych bitach maskę praw dostępu. API Windows definiuje trzy możliwe wartości dla tej maski - `PAGE_NOACCESS` oznacza brak dostępu do strony, `PAGE_READONLY` dostęp tylko do odczytu, a `PAGE_READWRITE` dostęp i do odczytu i do zapisu. Kolejne 20 bitów określa adres strony w pamięci fizycznej – 20 bitów wystarczy do zaadresowania dowolnych 4KB w pamięci rozmiaru 4GB. Jeżeli strona jest na dysku, owe 20 bitów zostaje użyte jako adres strony wewnątrz pliku wymiany. Kolejne 4 bity wyznaczają jeden z 16 możliwych plików wymiany, który odpowiada za przechowywanie tej strony na dysku. Dalej następują bity stanu strony: zapalony bit transition oznacza stronę w stanie przejściowym. Zapalony bit dirty oznacza stronę, której zawartość została zmieniona i powinna zostać zapisana. Włączony bit present implikuje stronę znajdującą się w pamięci. Jest tylko pięć możliwych znaczeń bitów stanu, ponieważ dla strony obecnej w pamięci (present) bit transition nie jest brany pod uwagę, a dla stron ze zgaszonymi bitami transition i present bit dirty nie ma znaczenia.

Pamięć dzielona



Dla realizacji pamięci dzielonej wprowadzono kolejny typ strony, tak zwane pole prototypowe strony współdzielonej. Każda strona współdzielona ma przypisane jedno pole prototypowe. Na wszystkie pola prototypowe w systemie przeznaczony jest 8MB pamięci, przydzielanej w razie potrzeby. Proces, który współdzieli stronę pamięci z innym procesem ma w odpowiedniej tablicy stron wskazanie na pole prototypowe (zamiast na zwykłą ramkę strony).

Baza danych ramek stron



Proces menadżera pamięci korzysta z bazy danych stron, przechowywanej w specjalnym obszarze pamięci niestronicowanej – nigdy na dysku, zawsze w fizycznej pamięci operacyjnej. Baza przechowuje informację o wszystkich ramkach pamięci fizycznej. Dane pogrupowane są w 6 list, po jednej liście dla każdego możliwego stanu strony. Elementy list zawierają odniesienia do informacji o stronie (czyli do tablicy stron wraz z indeksem w tablicy, lub do pola prototypowego). Strony prawidłowe znajdują się w pamięci i są w użyciu. Strony zmienione są w użyciu, znajdują się w pamięci i nie zostały jeszcze zapisane na dysk. Element listy uśpione to ramki z przypisaną stroną oczekująca na przejście do zbioru wolnych ramek. Strony wolne to strony nie przypisane do żadnego procesu. Strony wyzerowane są gotowe do przypisania do dowolnego procesu. Strony uszkodzone nie będą użyte przez żaden proces.