

Podstawy systemów operacyjnych

Wykład 6

Zarządzanie plikami w systemie Unix

Rodzaje plików:

- **Zwykły bądź typowy:** zawiera arbitralne dane w blokach danych. Zawiera informacje wprowadzane przez użytkownika lub aplikację. System plików nie narzuca żadnej wewnętrznej struktury, są one traktowane jako strumień bajtów.
- **Katalog:** Zawiera listę nazw plików wraz ze wskaźnikami do skojarzonych z nimi węzłów indeksowych (i-węzłów). Katalogi mają strukturę hierarchiczną. Pliki w katalogach to zwykłe pliki ze specjalnymi prawami zapisu, pozwalającymi tylko systemowi plików do nich pisać, natomiast programy użytkownika mogą odczytywać zapisane dane.

Rodzaje plików c.d.:

- **Specjalny:** Nie zawiera danych, ale oferuje mechanizm odwzorowania urządzeń fizycznych na nazwy plików. Nazwy te są używane by uzyskać dostęp do urządzeń peryferyjnych. Każde urządzenie we/wy jest skojarzone ze specjalnym plikiem.
- **Nazwane potoki:** Potok to usługa komunikacji między procesowej. Plik potokowy buforuje dane otrzymane w wyniku operacji wejścia tak, aby proces odczytujący dane z potoku wyjściowego otrzymaj je z wykorzystaniem strategii FIFO

Rodzaje plików c.d.:

- **Odnośniki:** Odnośnik to alternatywna nazwa pliku.
- **Łącze symboliczne:** Jest to plik zawierający skojarzoną z nim nazwę pliku.

i-węzły

- Wszystkie rodzaje plików w systemie Unix są kontrolowane przez system operacyjny za pomocą i-węzłów.
- i-węzeł (węzeł indeksowy) to struktura sterująca zawierająca kluczowe informacje potrzebne systemowi operacyjnemu dla danego pliku.
- Z jednym i-węzłem może być skojarzonych kilka nazw plików, ale aktywny i-węzeł jest skojarzony z dokładnie jednym plikiem.

Informacja w i-węźle rezydentnym

Tryb pliku	16-bitowa flaga przechowująca prawa dostępu i wykonania skojarzone z plikiem	
	12-14	Rodzaj pliku (typowy, katalog, znakowy, FIFO)
	9-11	Flagi wykonywania
	8	Prawo odczytu dla właściciela
	7	Prawo zapisu dla właściciela
	6	Prawo wykonywania dla właściciela
	5	Prawo odczytu dla grupy
	4	Prawo zapisu dla grupy
	3	Prawo wykonywania dla grupy
	2	Prawo odczytu dla pozostałych użytkowników
	1	Prawo zapisu dla pozostałych użytkowników
	0	Prawo wykonywania dla pozostałych użytkowników

Informacja w i-węźle rezydentnym c.d.

Licznik odwołań	Liczba odwołań katalogów do tego i-węzła
Identyfikator właściciela	Właściciel pliku
Identyfikator grupy	Właściciel grupy skojarzony z danym plikiem
Rozmiar pliku	Liczba bajtów pliku
Adresy pliku	39 bajtów informacji adresowych
Ostatni dostęp	Data ostatniego dostępu do pliku
Ostatnia modyfikacja	Data ostatniej modyfikacji pliku
Modyfikacja i-węzła	Data ostatniej modyfikacji i-węzła

Na dysku istnieje tablica i-węzła lub lista i-węzła zawierająca i-węzły wszystkich plików w systemie plików. Po otwarciu pliku jego i-węzeł jest wstawiany do pamięci operacyjnej i przechowywany w tablicy i-węzłów rezydującej w pamięci.

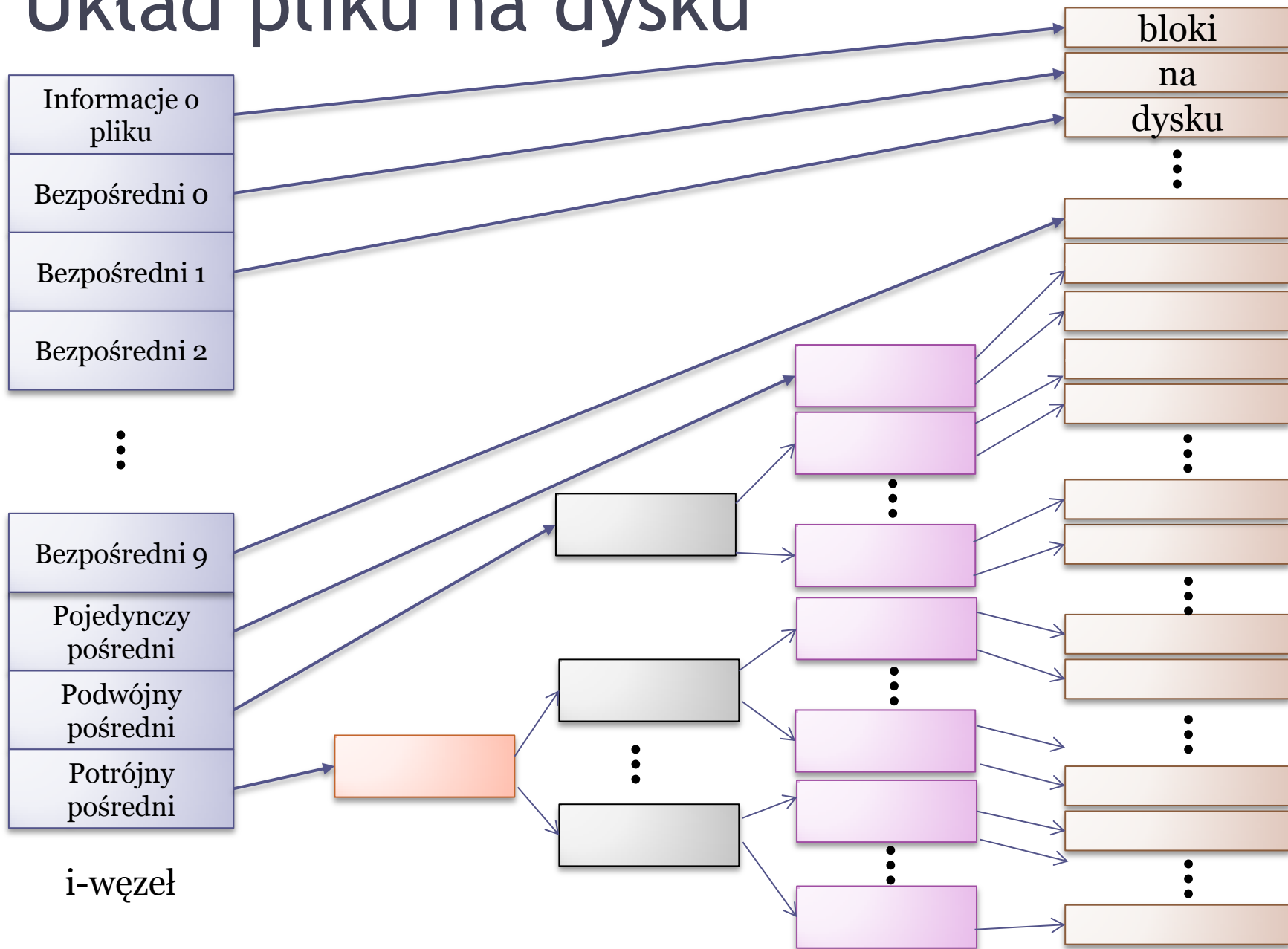
Alokacja plików

- Alokacja plików jest realizowana blokowo. Jest to alokacja dynamiczna realizowana w miarę potrzeb. Bloki plików na dysku nie muszą być blokami ciągłymi.
- Metoda indeksowania służy do śledzenia każdego pliku.
- Do przechowywania informacji o adresie stosuje się i-węzeł o rozmiarze **39-bajtów**, posiadający strukturę trzynastu 3-bajtowych adresów lub wskaźników.
- Pierwszych 10 adresów wskazuje na pierwszych 10 bloków danych pliku.

Alokacja plików c.d.

- Jeśli plik jest dłuższy niż 10 bloków stosuje się 3 poziomy pośrednie.:
 - **Jedenasty** adres w i-węźle odwołuje się do bloku na dysku zawierający kolejny fragment indeksu. Jest to tzw. pojedynczy blok pośredni, Zawiera on wskaźniki do kolejnych bloków pliku;
 - Jeśli plik zawiera więcej bloków, **dwunasty** adres w i-węźle wskazuje na podwójny blok pośredni, który zawiera listę pojedynczych bloków pośrednich. A w każdym z nich są wskaźniki do bloków pliku.
 - Jeśli plik zawiera więcej bloków, to **trzynasty** adres w i-węźle wskazuje na potrójny blok pośredni, będący trzecim poziomem indeksacji. Blok ten wskazuje na dodatkowe podwójne bloki pośrednie.

Układ pliku na dysku



Pojemność pliku

- Pierwszy wpis w i-węźle zawiera informacje o tym pliku lub katalogu. Pozostałe wpisy to omówione adresy.
- W Unix System V długość blok to 1KB
- Całkowita liczba bloków danych w pliku zależy od pojemności bloków o stałych rozmiarach.

Poziom	Liczba bloków	Liczba bajtów
Bezpośredni	10	10kB
Pojedynczy pośredni	256	256kB
Podwójny pośredni	$256 \times 256 = 65\text{kB}$	65MB
Potrójny pośredni	$256 \times 65\text{KB} = 16\text{MB}$	16 GB

Zalety schematu i-węzłów

- I-węzeł ma stały rozmiar i jest względnie nieduży, stad może być przechowywany w pamięci operacyjnej przez dłuższy czas.
- Dostęp do mniejszych plików może być realizowany za pomocą niewielkiego, bądź żadnego pośrednictwa, co zmniejsza czas przetwarzania oraz czas dostępu do pliku.
- Teoretyczny maksymalny rozmiar pliku jest wystarczająco duży, ba zaspokoić wszystkie wymagania.

Katalogi

- Katalogi mają strukturę drzewa hierarchicznego
- Katalog zawiera pliki i/lub katalogi
- Katalog przechowywany wewnątrz innego katalogu to podkatalog
- Katalog to prosty plik zawierający listę nazw plików oraz wskaźniki do skojarzonych i-węzłów.

Katalogi oraz i-węzły

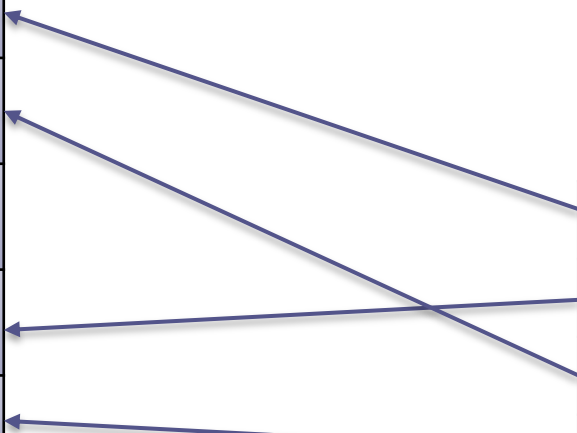
Tablica i-węzłów

⋮

Katalog

i1	Nazwa nr 1
i2	Nazwa nr 1
i3	Nazwa nr 1
i4	Nazwa nr 1

⋮



Struktura wolumenu

System plików Unix zajmuje jeden dysk logiczny lub partycję dyskową i składa się z następujących elementów:

- **Bloku startowego:** zawiera kod wymagany do uruchomienia systemu operacyjnego.
- **Superbloku:** zawiera atrybuty i informacje dotyczące systemu plików, takie jak rozmiar partycji oraz rozmiar tablicy i-węzła.
- **Tablicy i-węzła:** zestaw i-węzłów dla każdego pliku.
- **Bloków danych:** przestrzeń do przechowywania plików z danymi oraz podkatalogów.

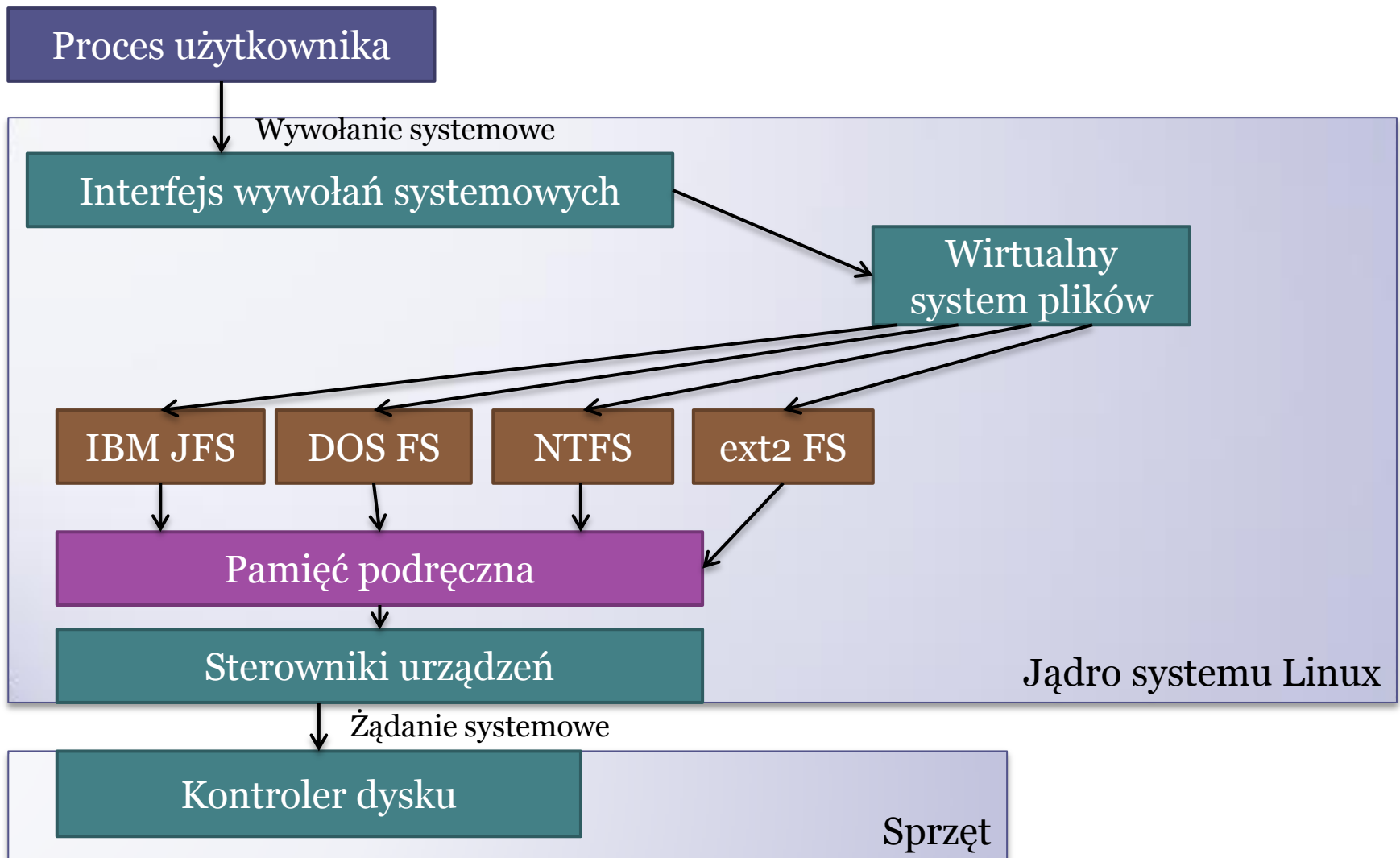


Linux

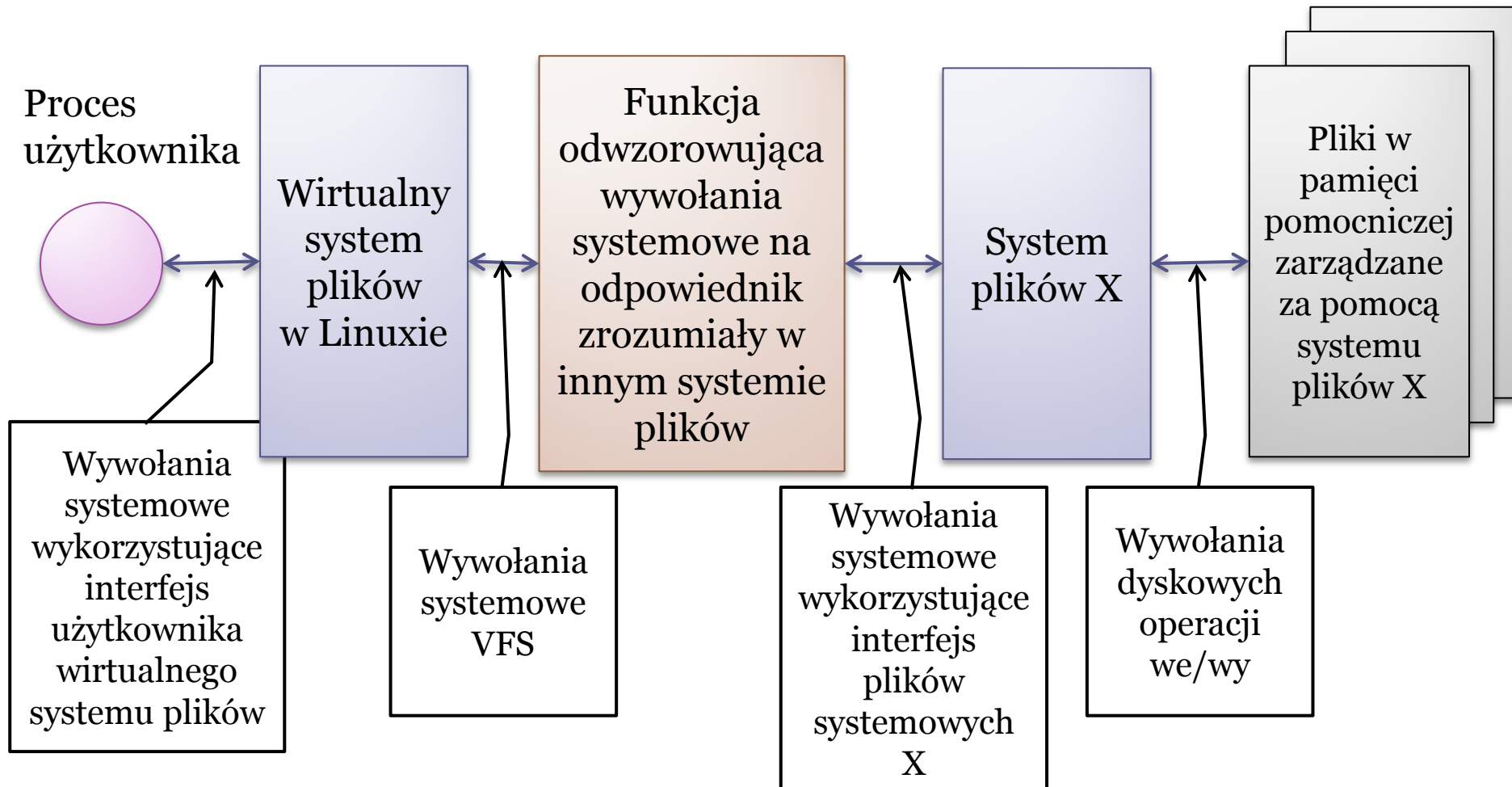
Wirtualny system plików w Linuxie

- Wirtualny system plików (VFS) zawiera ujednolicony interfejs systemowy dla procesów użytkownika
- VFS definiuje typowy model pliku, który może reprezentować każdą ogólną funkcję systemu plików oraz jej działanie
- W VFS przyjmuje się, że pliki to obiekty przechowywane w pamięci masowej komputera.
- pliki mają nazwy symboliczne, które umożliwiają ich identyfikację wewnątrz danego katalogu w danym systemie plików

Kontekst wirtualnego systemu plików



Idea wirtualnego systemu plików



Obiekty VFS

Cztery podstawowe rodzaje obiektów VFS, to:

- **Obiekt superbloku:** reprezentuje określony podłączony system plików.
- **Obiekt i-węzła:** Reprezentuje określony plik.
- **Obiekt wpisu do katalogu:** reprezentuje określony wpis do katalogu.
- **Obiekt pliku:** Reprezentuje otwarty plik skojarzony z danym procesem.

Mechanizm ten bazuje na schematach wykorzystywanych w systemie Unix

Obiekt superbloku

- Obiekt superbloku przechowuje informacje opisujące określony system plików, zazwyczaj odpowiada superblokowi systemu plików lub blokowi kontrolnemu systemu plików.
- Obiekt ten przechowuje wiele elementów danych np.:
 - Urządzenie na którym zainstalowano system plików
 - Podstawowy rozmiar bloku w systemie plików
 - Brudna flaga wskazująca, że superblok uległ zmianie, ale zmiany te nie zostały jeszcze zapisane na dysku.
 - Rodzaj pliku.
 - Wskaźnik do głównego katalogu w systemie pliku.
 - Semafore sterujące dostępem do plików.

Obiekt i-węzła

- I-węzeł jest skojarzony z każdym plikiem.
- Obiekt i-węzła przechowuje wszystkie informacje na temat nazwanego pliku z wyjątkiem jego nazwy oraz faktycznej zawartości pliku.
- Metody zdefiniowane dla obiektu i-węzła to:
 - **create**: tworzy nowy i-węzeł dla regularnego pliku skojarzonego z obiektem wpisu katalogowego w określonym katalogu.
 - **lookup**: przeszukuje katalog w celu znalezienia i-węzła odpowiadającego nazwie pliku.
 - **mkdir**: tworzy nowy i-węzeł dla katalogu skojarzonego z obiektem zapisu w określonym katalogu.

Obiekt zapisu w katalogu

- Obiekt zapisu w katalogu to specyficzny składnik ścieżki.
- Składnikami mogą być nazwa katalogu bądź nazwa pliku
- Obiekty zapisu w katalogu *przyspieszają* dostęp do plików oraz katalogów i są używane w pamięci podręcznej zapisów katalogowych.

Obiekt plikowy

- Obiekt plikowy służy do odzwierciedlenia pliku otwartego przez proces.
- Obiekt jest tworzony w odpowiedzi na wywołanie systemowe `open()` oraz niszczone w odpowiedzi na wywołanie systemowe `close()`.
- Składa się z określonej liczby elementów min:
 - Systemu plików zawierającego dany plik.
 - Licznika wykorzystania obiektów plikowych.
 - Identyfikatora użytkownika.
 - Identyfikatora grupy dla danego użytkownika.
 - Wskaźnika pliku, oznaczającego bieżącą pozycję w pliku, od której wykonana zostanie kolejna operacja.