System plików

Adam Jasiorowski, Piotr Makowski, Marcin Regulski, Sławomir Mazurek

1 I-węzły.

1.1 Definicja.

I-węzeł (węzeł indeksu) jest metryczką pliku. Opisuje jego atrybuty i położenie na dysku.

I-węzły istnieją w postaci statycznej na dysku, a jądro wczytuje je do pamięci, aby się nimi posługiwać. I-węzły opisują struktury `struct inode`, zdefiniowana w pliku `include/linux/fs.h` oraz `struct ext2_inode_info` z pliku `include/linux/ext2_fs_i.h`. I-węzeł na dysku zawiera pola z następującymi informacjami:

- Numer i-węzła w dyskowej tablicy i-węzłów.
- Identyfikator właściciela pliku i identyfikator grupy właściciela. Określa to zbiór użytkowników, którzy mają prawo dostępu do pliku.
- Typ pliku. Wyróżnia się pliki zwykłe, katalogi, specjalne pliki znakowe i blokowe oraz łącza.
- Prawa dostępu do pliku. Wyróżnia się trzy kategorie użytkowników: właściciela, użytkowników z grupy, do której należy właściciel i pozostałych użytkowników. Każdej z tych kategorii można nadać osobno prawo do czytania, pisania i zapisywania pliku.[1] Nadzorca systemu ma prawo do wszystkich czynności.
- Czasy dostępu do pliku: czas utworzenia pliku, czas ostatniej modyfikacji pliku, czas ostatniego dostępu do pliku i czas ostatniego zapisu i-węzła na dysk.
- Rozmiar pliku, maksymalnie 4 GB. Z punktu widzenia użytkownika plik jest strumieniem bajtów, indeksowanych od zera.
- Liczba bloków dyskowych zajmowanych przez plik[2] oraz tablica z numerami bloków zawierających dane pliku na dysku i pośrednich bloków wskazujących na bloki z danymi.
- Liczba dowiązań do pliku. Plik może być otwarty przez wielu użytkowników, a każdy użytkownik może kilka razy otworzyć ten sam plik.
- Wskaźnik do bloku identyfikacyjnego urządzenia, do którego należy i-węzeł.
- Liczba procesów odwołujących się do i-węzła.
- Liczba procesów mających prawo zapisu do pliku.
- Struktura z informacjami do blokowania pliku.

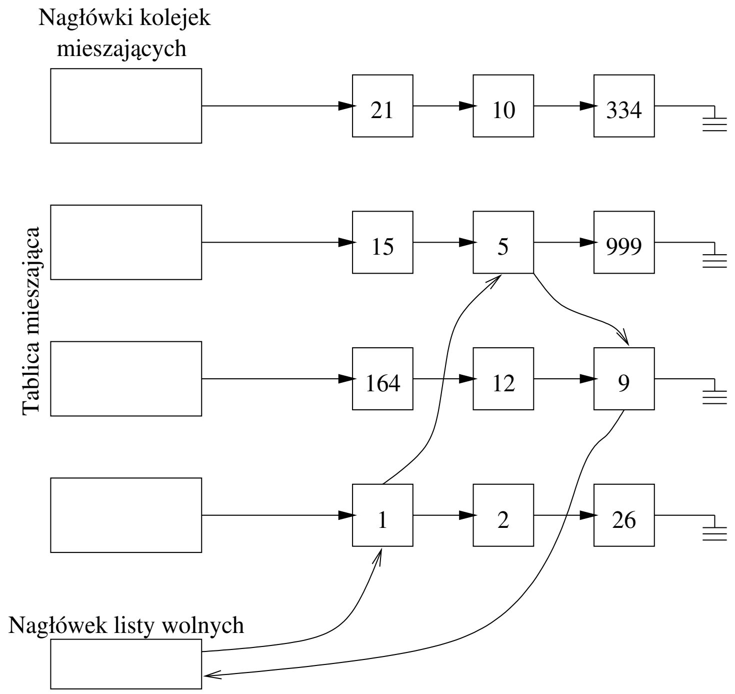
Sprowadzony do pamięci w wyniku działania funkcji `__iget` z pliku `fs/inode.c`, i-węzeł wzbogacony jest o następujące informacje:

- Stan i-węzła w pamięci wskazuje, czy:

---

[1] W przypadku katalogów prawo do wykonywania oznacza prawo do wejścia do katalogu.

[2] Albowiem plik jest przechowywany na dysku w wielu blokach.



Rysunek 1: I-węzły w kolejkach mieszających i na liście wolnych i-węzłów. Rozmiar tablicy mieszającej wynosi 512; tak definuje go stała `NR_IHASH` z pliku `fs/inode.c`.

  - i-węzeł ma założoną blokadę;
  - proces czeka na zdjęcie blokady z i-węzła;
  - reprezentacja i-węzła w pamięci rózni się od kopii dyskowej i-węzła z powodu zmiany danych w i-węźle.

- Logiczny numer urządzenia systemu plików zawierającego plik opisywany przez dany i-węzeł.
- Logiczny numer i-węzła.
- Wskaźniki do innych i-węzłów w pamięci, mianowicie utrzymujące i-węzeł na liście wolnych i-węzłów oraz w kolejce mieszającej służącej do przyspieszania dostępu do i-węzłów. Kolekę mieszającą wyznacza się przy użyciu numeru i-węzła i numeru urządzenia zawierającego i-węzeł. Jądro może zawierać w pamięci tylko jedną kopię i-węzła dyskowego, lecz i-węzły mogą być jednocześnie w kolejce mieszającej i na liście wolnych.
- Liczbę pokazującą, ile razy plik opisywany przez i-węzeł został otwarty.
- Wskaźniki do operacji na i-węzłach, różnych dla poszczególnych systemów plików.
- Semafor do blokowania dostępu do i-węzła. Używają go, między innymi, funkcje `open`, `mknod` i `get_empty_inode`.

W i-węźle nie przechowuje się nazw ścieżkowych używanych w celu dostępu do pliku; funkcja `namei` z pliku `fs/namei.c` zwraca i-węzeł pliku o podanej nazwie ścieżkowej, patrz rozdział 4.

1.2 Dostęp do i-węzłów.

Aby wykonać operacje na pliku, proces musi pobrać jego i-węzeł. Odbywa się to w ten sposób, że wyszukiwana jest w pamięci kopia szukanego i-węzła. Jądro, na podstawie numeru urządzenia i numeru i-węzła, wyznacza kolejkę mieszającą i próbuje pobrać z niej i-węzeł. Jeżeli nie zostanie on odnaleziony, pobierany jest i-węzeł z listy wolnych i zakłada się na niego blokadę, po czym wczytuje się do pamięci jego kopię dyskową. Pobieraniem i-węzłów zajmuje się funkcja `__iget` z pliku `fs/inode.c`.

```
Funkcja __iget

NAGŁÓWEK:
struct inode *__iget (struct super_block *sb, int nr, int crossmntp)

WYNIK:
W przypadku sukcesu, kopia z pamięci i-węzła o numerze nr z urządzenia
o bloku identyfikacyjnym sb.
W przypadku błędu, NULL (tylko gdy nie da się pobrać i-węzła z listy wolnych).

ALGORYTM:
znajdź odpowiednią kolejkę mieszającą;
jeżeli i-węzeł znajduje się w tej kolejce, to:
{

    1. Jeżeli znaleziony i-węzeł jest w trakcie operacji clear_inode, czyli
    uzdatniania go w celu ponownego użycia jako wolngo i-węzła, zaśnij
    w oczekiwaniu na zakończenie tej operacji; kiedy to nastąpi obudź się
    i zacznij poszukiwanie od samego początku.
    W przeciwnym przypadku zwiększ licznik dowiązań w i-węźle, zmniejsz o 1
    liczbę wolnych i-węzłów w systemie i jeśli na ten i-węzeł była założona
    jakaś blokada, to czekaj na jej zdjęcie.
    2. Dalej, jeżeli i-węzeł jest punktem zamontowania innego systemu plików,
    to zamień znaleziony i-węzeł na ten, który jest do niego zamontowany,
    czyli na główny i-węzeł zamontowanego systemu plików, ponieważ to w nim
    są dowiązania do katalogów i plików tam się znajdujących, a nie w tym,
    który uprzednio został odnaleziony. W tym przypadku zwróć pierwotnie
    odnaleziony i-węzeł do systemu i poczekaj na zdjęcie ewentualnych blokad
    z aktualnego.
    3. Na koniec czekaj na zdjęcie ewentualnej blokady z listy, na której
    znajduje się odnaleziony i-węzeł, żeby nie dopuścić do sytuacji, kiedy
    kilka procesów modyfikuje dowiązania do elementów listy sąsiadujących
    z tym i-węzłem (jest to lista dwukierunkowa).
    4. Wreszcie zwróć i-węzeł.
};

jeżeli początkowe poszukiwania nie dały rezultatu, to utwórz nowy i-węzeł
wykorzystując w tym celu jeden z nie używanych aktualnie przez żaden proces;
jeśli takich nie ma, to przeznacz nową stronę pamięci na i-węzły, umieść je
na liście dostępnych oraz pobierz i zainicjuj nowy i-węzeł;
po wykonaniu tej operacji rozpocznij poszukiwanie od początku, jeśli jednak
nie udało się przydzielć pamięci na nowe i-węzły to występuje błąd, więc
zwróć NULL;
```

Aby wczytać z dysku i-węzeł, jądro musi wyliczyć jego fizyczne położenie na dysku (urządzeniu logicznym). Ponieważ tablica i-węzłów jest przechowywana w pewnej ilości bloków dyskowych, jądro musi znaleźć blok, w którym znajduje się i-węzeł oraz jego numer w danym bloku (licząc od 0). Odbywa się to według nastepujących wzorów:

numer bloku = ((numer i-węzła – 1) div liczba i-węzłów w bloku) + numer bloku, od którego rozpoczyna się lista i-węzłów ;

położenie i-węzła w bloku = ((numer i-węzła – 1) mod liczba i-węzłów w bloku) * rozmiar i-węzła na dysku.

1.3 Zwalnianie i-węzłów.

Gdy proces zakończył operacje z plikiem, musi zwolnić kopię pamięciową jego i-węzła. Zmniejsza się o jeden liczbę otwartych plików opisywanych przez dany i-węzeł; gdy kopia i-węzła w pamięci różni się od kopii na dysku[3], uaktualnia się ją. I-węzeł zostaje umieszczony na liście wolnych, można także zwolnić wszystkie bloki danych pliku opisywanego przez i-węzeł oraz sam i-węzeł na dysku, gdy liczba dowiązań do pliku wynosi zero. Zwalnianiem i-węzłów zajmuje się funkcja `iput` z pliku `fs/inode.c`.

```
Funkcja iput

NAGŁÓWEK:
void iput (struct inode *inode)

WYNIK:
Brak.

ALGORYTM:
jeżeli nie ma blokady na i-węźle, to ją załóż;
jeżeli liczba procesów odwołujących się do i-węzła wynosi zero, wypisz
komunikat „Próba zwolnienia wolnego i-węzła” i zakończ działanie;
jeżeli i-węzeł opisywał łącze nienazwane, obudź wszystkie przerywalne
procesy oczekujące na dostęp to tego łącza;

(*)
jeżeli liczba procesów odwołujących się do i-węzła jest większa od 1,
to zmniejsz ją o jeden i zakończ działanie;

obudź procesy czekające na wolny i-węzeł;
jeżeli i-węzeł opisywał łącze nienazwane, zwolnij związaną z nim stronę
pamięci;
wykonaj funkcję put_inode,[4]
a jeżeli liczba dowiązań do pliku wynosi zero, zakończ działanie;
jeżeli i-węzeł należy zaktualizować, to to zrób, wykonaj test na blokadę i
GO TO (*);

jeżeli i-węzeł opisuje obiekt, do którego można zapisywać i zdefiniowano
limit dyskowy,[5] to go uaktualnij i GO TO (*);
zmniejsz o jeden liczbę procesów, które odwołują się do i-węzła;
zwiększ o jeden liczbę wolnych i-węzłów;
```

---

[3] Gdy zmieniła się zawartość pliku, czas ostatniego dostępu do pliku, jego właściciel lub prawa dostępu.

[4] Wykonuje ona czynności zwalniania i-węzła charakterysyczne dla poszczególnych systemów plików – czyli u nas dla systemu ext2.

[5] W ten sposób przetłumaczyłem angielski termin „disk quota”. Zwrot „kwota dyskowa” uważam za nieodpowiedni, kwota to suma pieniędzy. Ostatecznie zdecydowałem się na „limit dyskowy”.