Virtualizace RAM v PC – stránkování, swapování, účel a princip, výpočet adres v protected módu

Virtualizace RAM

Operační paměť (fyzická) je pro nainstalované programy téměř vždy nedostačující. Rozšiřuje se proto o tzv. virtuální paměť, do níž procesor odkládá momentálně nepotřebné stránky z OP.

Zásadní myšlenkou virtuálních pamětí je využít levnější sekundární paměti pro účely paměti primární. Sekundární pamětí je obvykle harddisk – rozdělený pro účely virtualizace RAM na rámce (frame), které jsou podle potřeby přesouvány (swapovány) z nebo do OP. V případě, že se počet volných stránek v OP blíží nule, procesor automaticky uvolní ty rámce (stránky), které nejdéle nepotřeboval. Tyto uvolněné rámce jsou poté uloženy na disk.

U unixových systémů je swapovací prostor tvořen samostatným log. oddílem bez souborového systému. Orientace na disku je jen podle čísel rámců. U Windows se vytváří na disku samostatný swapovací soubor. Swapovací soubor je sice pomalejší, ale jeho velikost je snadněji měnitelná.

Pro swapování je podmínkou zapnuté stránkování. Ve stránkovací tabulce najde OS stránky, které nejdéle nepoužíval a poté může najít volné místo ve swapovacím prostoru, kam celou stránku odloží. Do tabulky zapíše číslo rámce (swapnuté stránky) na disku a nastaví příznak „swapnuto“.

Pokud OS vyhledává data, podle příznaku „swapnuto“ zjistí, za stránka je swapnutá nebo není. Pokud je, najde volné místo v OP, opět zapíše do stránkovací tabulky číslo stránky a příznak zruší. Data jsou pak již připravena ke zpracování.

Stránkování

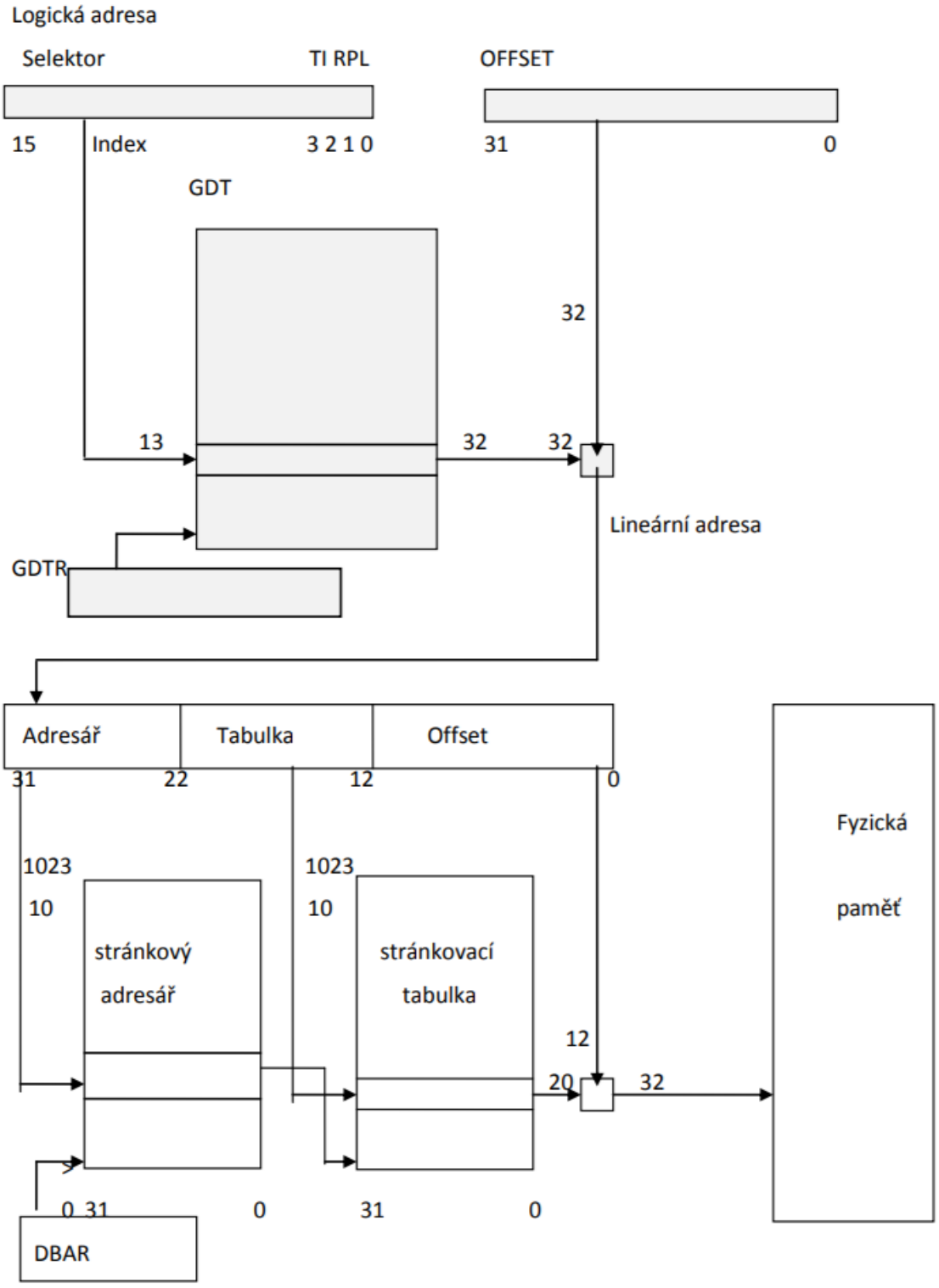
Po segmentaci v protected módu jde o další stupeň adresování operační paměti – na segmentaci navazující. Na rozdíl od segmentace, kdy má segment proměnnou délku, stránka má vždy konstantní délku, a to obvykle 4 kB. Pro práci se stránkou si pak stačí pouze zapamatovat číslo stránky.

Při stránkování je převáděna lineární adresa (výstup ze segmentace) podle dvoustupňových tabulek na fyzickou. Pokud tato tabulka převod neumožní, je vyvoláno přerušení, které swapuje stránku z disku za jinou, nepoužívanou stránku v paměti. Instrukce, která přerušení způsobila, je následně vyvolána znovu.

Není-li stránkování zapnuto, lineární adresa je pak považována za fyzickou. Dnes už však všechny procesory ve stránkovacím režimu pracují, a to s délkou stránky 4 kB, 8 kB nebo více.

Stránkování paměti s délkou stránky 4 kB

4 kB = 2^{12} B. Z toho plyne, že offset bude dlouhý 12 bitů. Výpočet lineární adresy je popsán níže v části „Výpočet adres v protected módu“).



Swapování

Swapování je odkládání momentálně nepotřebných dat do virtuální paměti (obvykle na vyhrazený prostor na HDD) a vrácení dat do operační paměti v momentě, kdy jsou data opět potřeba (*viz první část: Virtualizace RAM*).

Výpočet adres v protected módu

Segmentace = výpočet adres v protected módu; jde o první stupeň adresování operační paměti.

Protected mód se používá pro adresování paměti nad 1 MB. Paměť chrání proti zásahu jiných programů.

Protected mód byl původně určen pro procesory 80286, kdy byla zvětšena operační paměť na 16 MB a zejména kvůli umožnění multitaskingu (u kterého procesor musí zajistit izolovanost částí paměti určených pro jednotlivé úlohy, aby nedošlo k jejich promíchání; proto se tento režim také nazývá *protected* – chráněný).

Virtuální adresa je složena ze dvou složek (selektoru a offsetu). Celá adresa se v protected módu nazývá *virtuální adresa* (nebo *logická adresa*) a má obvykle 48 bitů.

a) **segment selector** – odpovídá segmentovému registru, 16 bitů
    a. **13 bitů** pro index – **určuje řádek v tabulce deskriptorů** (lokální nebo globální dle dalších tří bitů selektoru); 8 192 možných kombinací indexu
    b. **1 bit** určující, zda se jedná o globální nebo lokální prostor
    c. **2 bity** pro práva (RPL) (*viz konec dokumentu*)
b) **offset** v rámci registru – 32 bitů (u I80286 16 bitů)

Globální prostor je prostor, ve kterém jsou obvykle programy nebo proměnné přístupné více uživatelům. Ochranná funkce tedy zaručuje oddělení systémového a uživatelského software, kontrolu typu dat a oddělení jednotlivých úloh. Ochranné atributy jsou společné pro celý segment. Srovnání probíhá současně s překladem adresy.

Lokální prostor je obvykle určen jen pro umisťování jedinečných dat. Každý proces může mít svůj vlastní lokální prostor, v němž se pak řádky číslují opět od 0.

Globální tabulka deskriptorů

Tabulka viditelná v celém systému (její adresa je jediná globálně viditelná, tato adresa je uložená v registru GDTR – global descriptor table register[1]). Zbytek se adresuje pouze pomocí selektorů a deskriptorů.

Nachází se v ní globální kódové i datové segmenty a rovněž deskriptor popisující lokální tabulky deskriptorů, kdy každý proces může mít vlastní lokální tabulku.

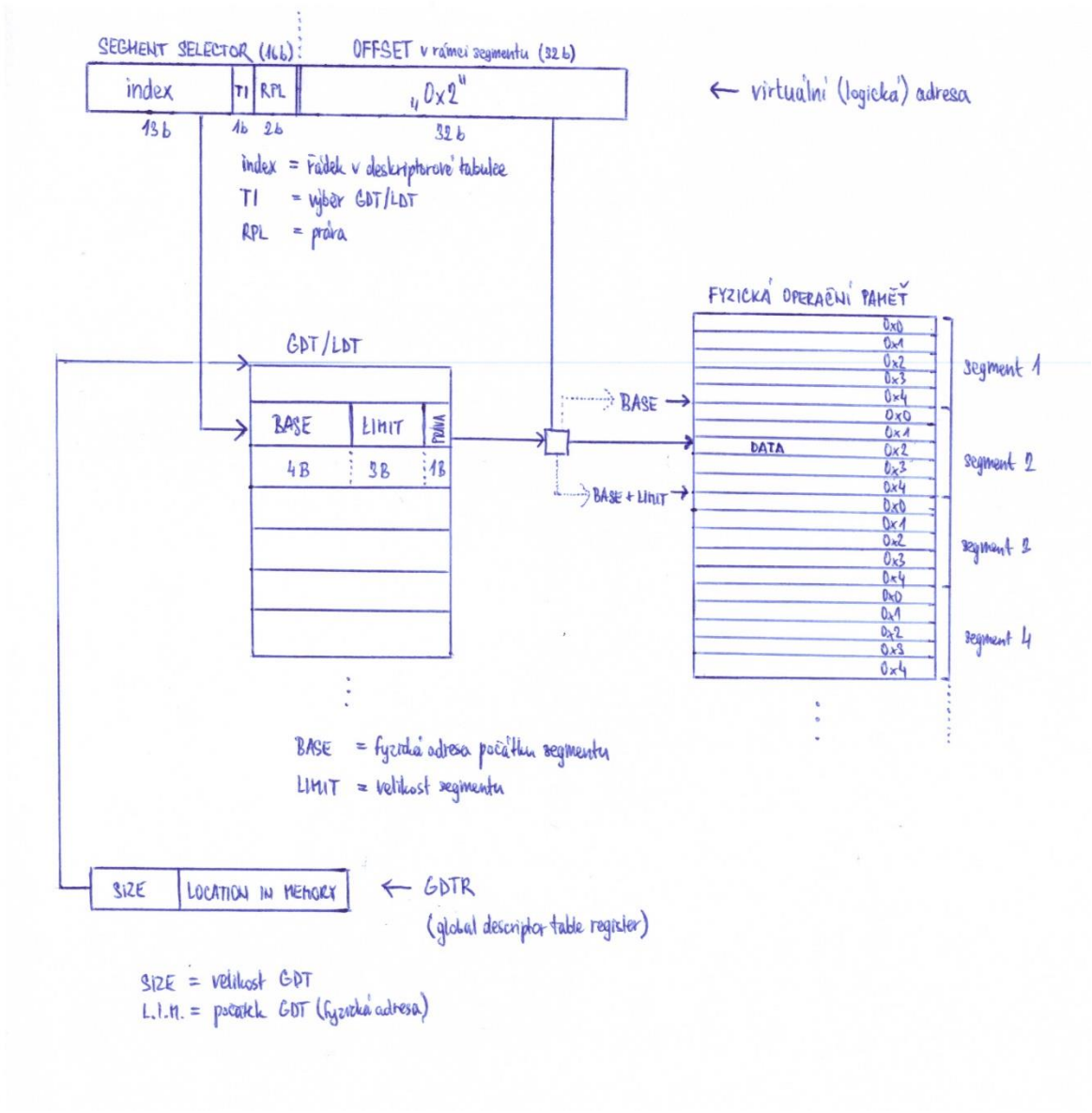
Postup při výpočtu adresy

Proces pracuje s virtuální (logickou adresou) skládající se ze segment selectoru a offsetu. Chce-li se proces dostat na danou adresu ve vlastní vyhrazené části OP, musí se vykonat následující:

1. ze *segment selectoru* se zjistí, zda se má vyhledávat segment procesu v GDT nebo LDT
2. v příslušné (GDT/LDT) tabulce se najde příslušný řádek – který konkrétní řádek se má najít se pozná z *indexu* ze *segment selectoru*

[1] v **GDTR** je uloženo 48 bitů – prvních 16 bitů označuje velikost GDT (SIZE), zbylých 32 bitů označuje počátek GDT v OP (obdobně jako BASE při výběru segmentu)

3. z nalezeného řádku v tabulce deskriptorů se dále zjistí, kde začíná segment, ve kterém jsou data, ke kterým se proces snaží dostat (první 4 B, tzv. BASE, značí adresu ve fyzické OP, kde segment začíná) a jak je segment velký (následující 3 B z řádku, tzv. LIMIT)
4. provede se kontrola, zda *offset* (druhá část virtuální adresy) není větší než LIMIT – to by totiž znamenalo, že se proces snaží přistoupit k datům z jiného segmentu, která mohou patřit jinému procesu – takový přístup by byl odmítnut
5. pokud kontrola proběhne v pořádku, v daném segmentu se nalezne příslušný řádek s daty (LIMIT ukazuje na 0x0, k této adrese se přičte *offset* → řádek ve fyzické OP)
6. bylo dosaženo fyzické operační paměti



RPL (Requester Privilege Level)

Určuje úroveň oprávnění přístupu k danému segmentu, má 2 bity.

a) **úroveň 0** – přístup k operačnímu systému, který řídí obvodové funkce mikroprocesoru a spravuje paměť; některé (privilegované) instrukce mohou pracovat pouze na této úrovni ochrany

b) **úroveň 1** – obsahuje rutiny pro správu systému pomocí OS
c) **úroveň 2** – pro zpracování knihoven, kartoték apod.
d) **úroveň 3** – přístupné jsou jen uživatelské programy

Ochrana paměti

Ochrana paměti je zajištěna mezním registrem na offset (procesor nedovolí použít offset větší, než je hodnota v mezním registru). Změna mezního registru a segment registru je privilegovaná instrukce (RPL úrovně 0), takže ji může provést pouze jádro operačního systému.

Běžící proces tedy může v neprivilegovaném stavu volně pracovat pouze s offsetem. Pokusí-li se proces běžící v neprivilegovaném režimu provést privilegovanou instrukci (tj. změnu segmentu nebo mezního registru), vyvolá procesor vnitřní přerušení, které může dotyčný proces okamžitě ukončit.