



Univerzita Hradec Králové
Fakulta informatiky a managementu

Přidělování paměti II

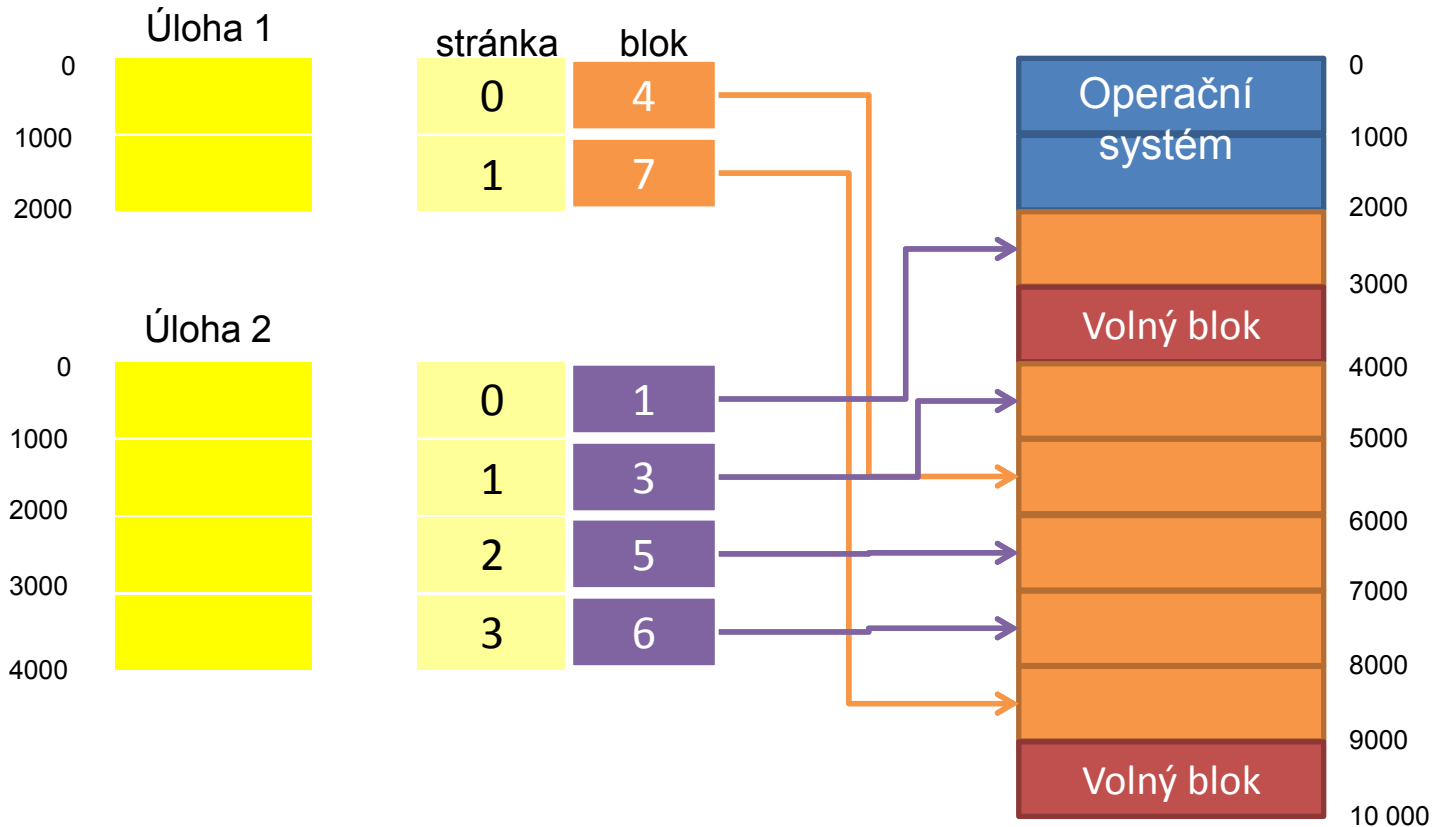
Mgr. Josef Horálek



- = *Přidělování jediné souvislé oblasti paměti*
- = *Přidělování paměti po sekcích*
- = *Dynamické přemístování sekcí*
- = **Stránkování**
- = Stránkování na žádost
- = Segmentace
- = Segmentace a stránkování na žádost

- = Na stejné úseky se rozdělí
 - = adresový prostor – stránky;
 - = operační paměť – fyzické stránky, stránkové rámy;
 - = každou stránku pak lze vložit do libovolného bloku;
 - = stránky jsou logicky souvislé vzhledem k uživatelskému programu;
 - = odpovídající bloky na sebe nemusí navazovat;
 - = transformace žádný vliv na uživatelskou úlohu;
 - = pro každou stránku musí být vyhrazen samostatný registr;
 - = efektivnost celé technologie značně ovlivňuje velikost stránky;
 - = stránkování řeší problém fragmentace bez fyzických přesunů;

- = Čtyři funkce modulu přidělování paměti
 - = sledování stavu se provádí pomocí tabulek;
 - = rozhodování o přidělování paměti provádí plánovač úloh;
 - = přidělení paměti;
 - = stránky se zavedou do přiřazených bloků;
 - = aktualizují se záznamy v tabulce stránek a tabulce bloku;
 - = uvolnění paměti;
 - = v tabulce bloku se odpovídající záznamy uvedou do stavu volný;



- = OS se stará o 3 základní tabulky:
 - = tabulka úloh (Job Table);
 - = tabulka bloku (Memory block table);
 - = tabulka stránek (Page Memory Table);

- = Výhody:
 - = eliminuje fragmentaci;
 - = umožňuje zvýšit počet úloh, pro které lze vytvořit paměťové prostory současně;
 - = lepší využití procesoru a operační paměti;
 - = eliminuje se režie zhušťování;
- = Nevýhody:
 - = technické prostředky pro transformaci stránek zvyšují obvykle cenu výpočetního systému a současně snižují jeho rychlost;
 - = uchování tabulek zabírá část operační paměti;
 - = současně se zvyšuje režie (čas procesoru) o dobu potřebnou k udržování těchto tabulek;

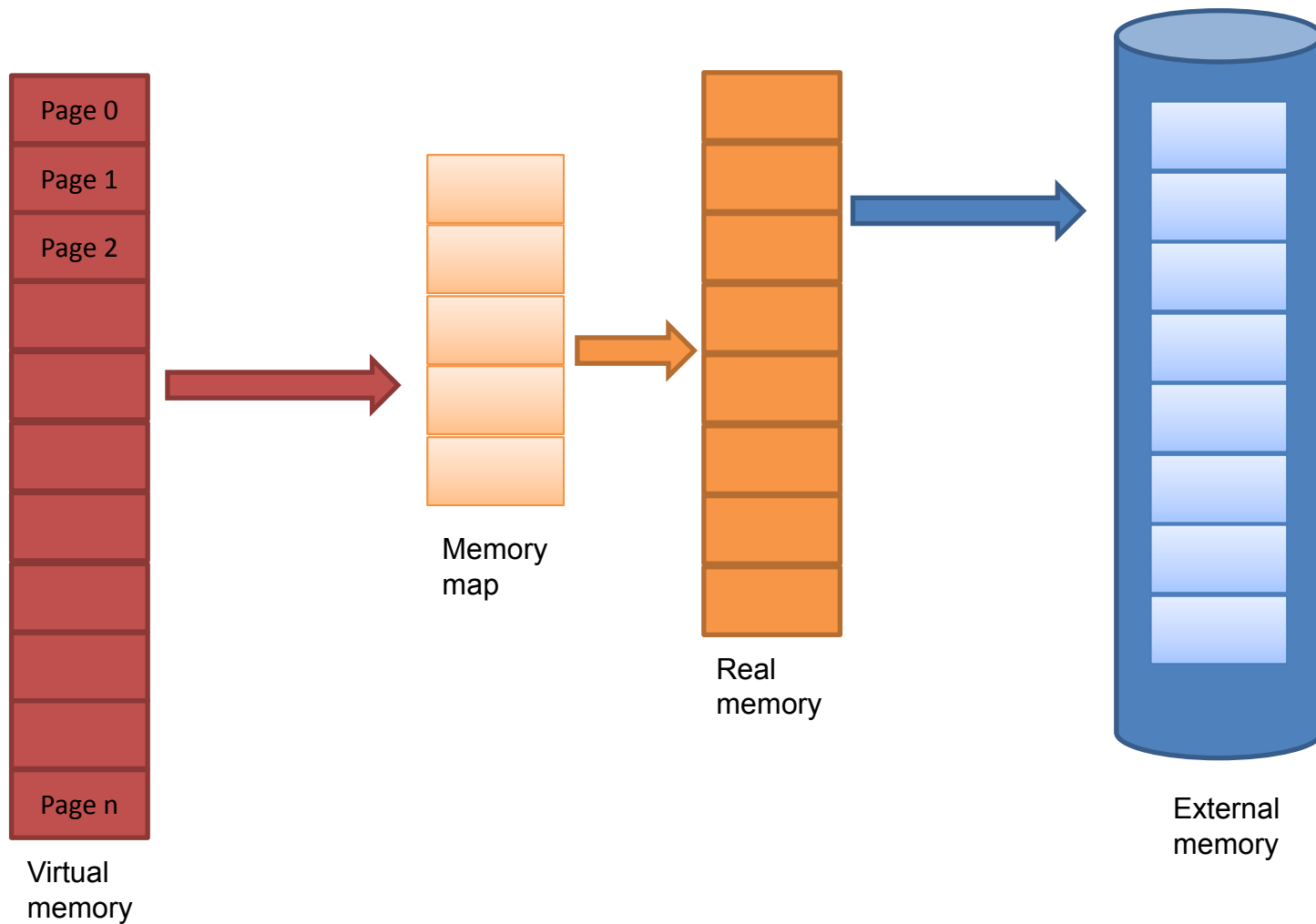
- = *Přidělování jediné souvislé oblasti paměti*
- = *Přidělování paměti po sekcích*
- = *Dynamické přemístování sekcí*
- = *Stránkování*
- = **Stránkování na žádost**
- = Segmentace
- = Segmentace a stránkování na žádost

= Dosud

- = úloha mohla být zpracována jen tehdy, když ji byla přidělena paměť pro celý adresový prostor;
- = časté nevyužité volné oblasti;

= Řešení

- = veliké paměti;
- = simulace rozsáhlé paměti
 - = virtuální paměť
 - = stránkování na žádost
 - = segmentace paměti.



=Většina programů během svého konkrétního průběhu využívá jen malou část svého adresového prostoru, protože:

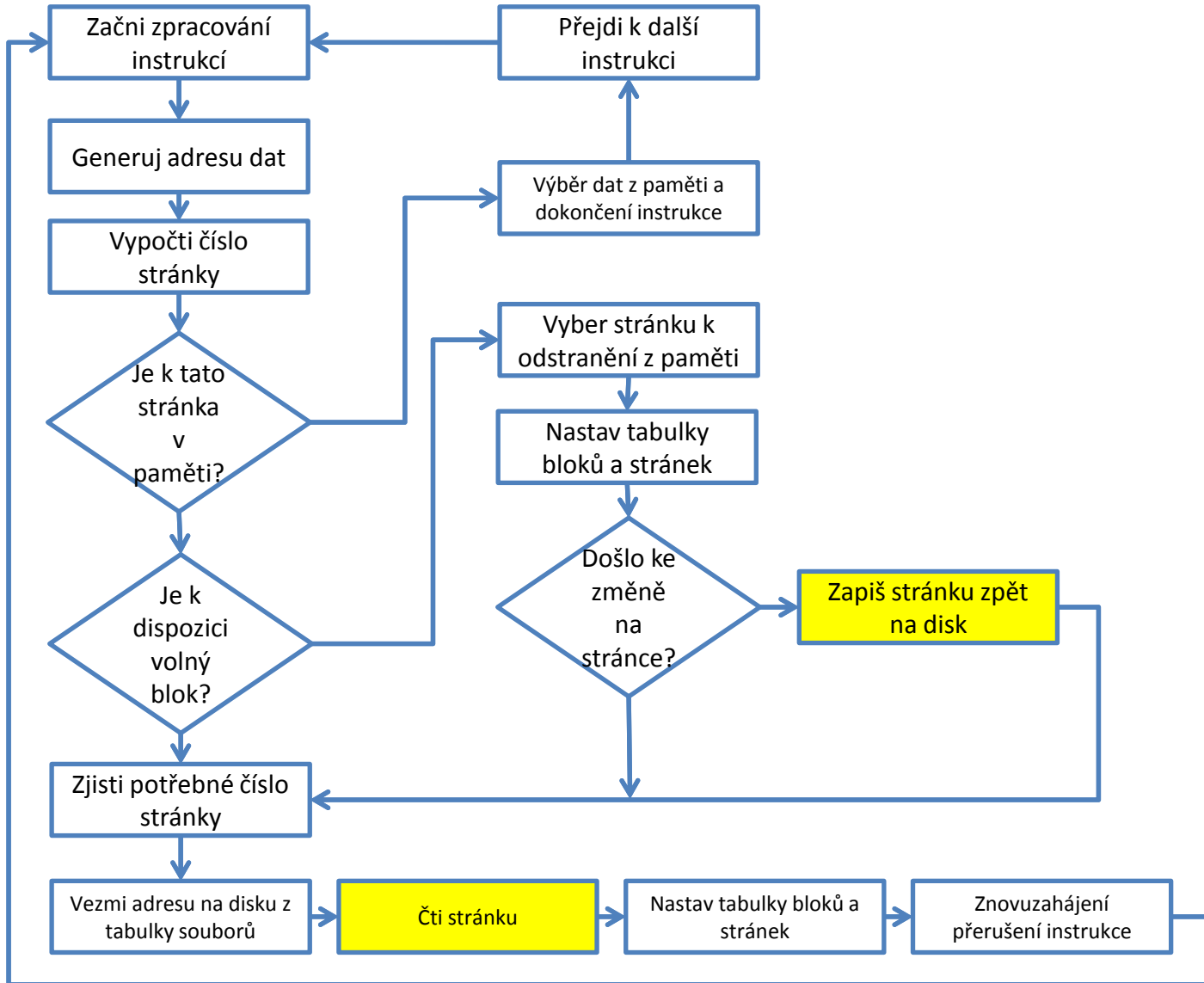
- = uživatelské programy pro ošetření chyb se užijí jen když k chybě dojde;
- = logické větve vylučují současný průběh alternativních částí programu;
- = mnoha tabulkám (statickým paměťovým strukturám) je přiděleno pevné množství adresového prostoru, které se ne vždy využije;
- = souběh mnoha podprogramů se časově vylučuje;

- = Možné problémy
 - = program se odvolává na oblast adresového prostoru, která není zavedena v operační paměti;
 - = strategie rozhodování, které stránky mají být uchovány v operační paměti;
- = Technické řešení
 - = rozšíření tabulky stránek o stavový bit;
 - = false - stránka je v operační paměti;
 - = true - stránka není v operační paměti;

- = Při naplánování zpracování úlohy je do operační paměti zavedena pouze první „startovací“ stránka;
 - = ostatní jsou pak zaváděny na žádost;
 - = nepotřebné stránky se do paměti vůbec nedostanou;
- = Možné komplikace
 - = co v okamžiku, kdy není paměťový prostor pro zavedení další stránky;
 - = řeší se pomocí výměny stránek;
 - = u stránkování na žádost je třeba rozšířit technické vybavení o tři důležité funkce:
 - = stavový bit v tabulce stránek;
 - = rozšíření mechanismu přerušení o výpadek stránky;
 - = záznam o používání stránky;

- = Čtyři funkce modulu přidělování paměti
 - = jsou při stránkování na žádost složitější a flexibilnější;
 - = sledování stavu paměti;
 - = sledování tabulek (PMTS, MBT, FMT);
 - = rozhodování o přidělení paměti;
 - = částečně provádí plánovač úloh;
 - = přidělování paměti
 - = je nutné najít vhodný blok paměti pro startovací stránku a změnit stavový bit bloku;
 - = uvolňování paměti
 - = není-li vhodný blok paměti k dispozici, musí být některý z obsazených uvolněn;

- = Ošetření výpadku stránky:
 - = paměť je přidělována a odebírána i během provádění úlohy pomocí mechanismu přerušení;
 - = při stránkování na žádost velmi úzce spolupracuje programové a technické vybavení;
 - = první část vývojového diagramu se provádí nejčastěji;
 - = implementována jako součást technického vybavení pro transformaci adres;
 - = druhá část je přerušovací modul v OS;

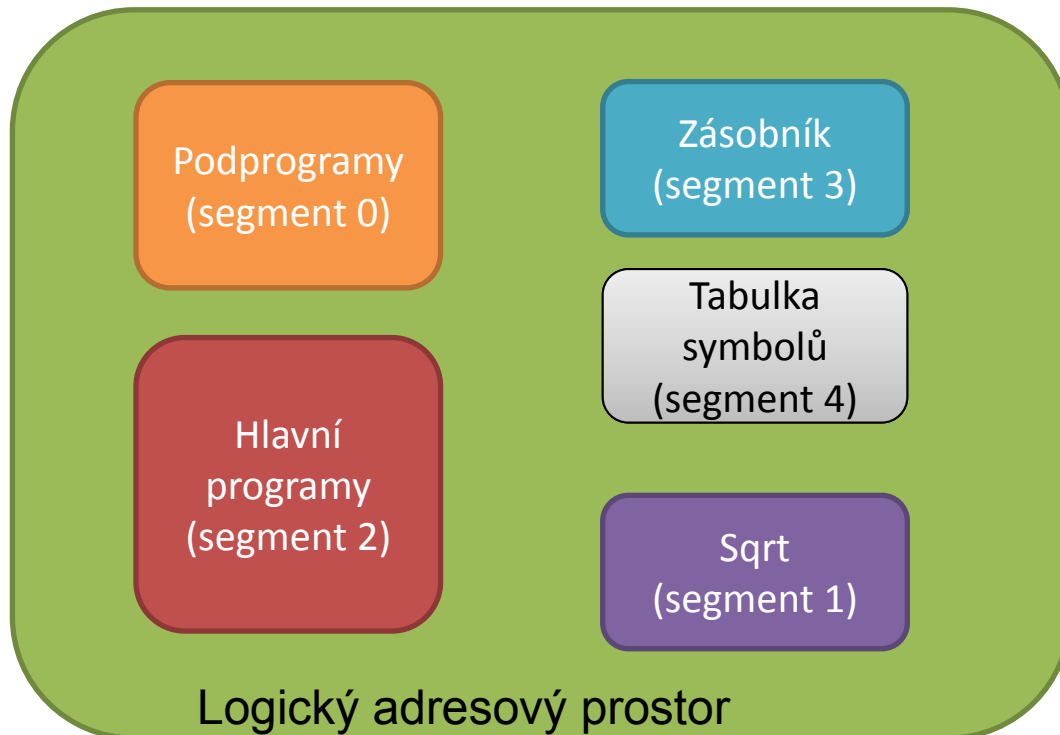


- = Přítomnost stránky v paměti se zjišťuje podle indikačního bitu v odpovídající položce tabulky stránek
 - = 0 - stránka je v paměti;
 - = 1 - stránka není v paměti.
 - = u systému s nepřímým adresováním na několika úrovních může nastat v podstatě neomezený počet výpadků stránek během vykonání jedné instrukce;
 - = při ošetření výpadku stránky mohou nastat až dvě I/O operace;
 - = během nich je možné přidělit procesor jiné úloze, ale vyžaduje to rozšířit stavy bloků o další položku přenos (in transit);

- = Řešení problémů „vyber stránku k odstranění z paměti“:
 - = nahradí se vždy stránka v bloku 3. (tj. první blok za OS);
 - = velmi jednoduché a velmi neefektivní;
 - = mohl by vést až k zahlceni systému;
 - = FIFO (First In - First Out)
 - = LRU (Least Recently Used)
 - = nejdéle nepoužita

- = *Přidělování jediné souvislé oblasti paměti*
- = *Přidělování paměti po sekcích*
- = *Dynamické přemístování sekcí*
- = *Stránkování*
- = *Stránkování na žádost*
- = **Segmentace**
- = Segmentace a stránkování na žádost

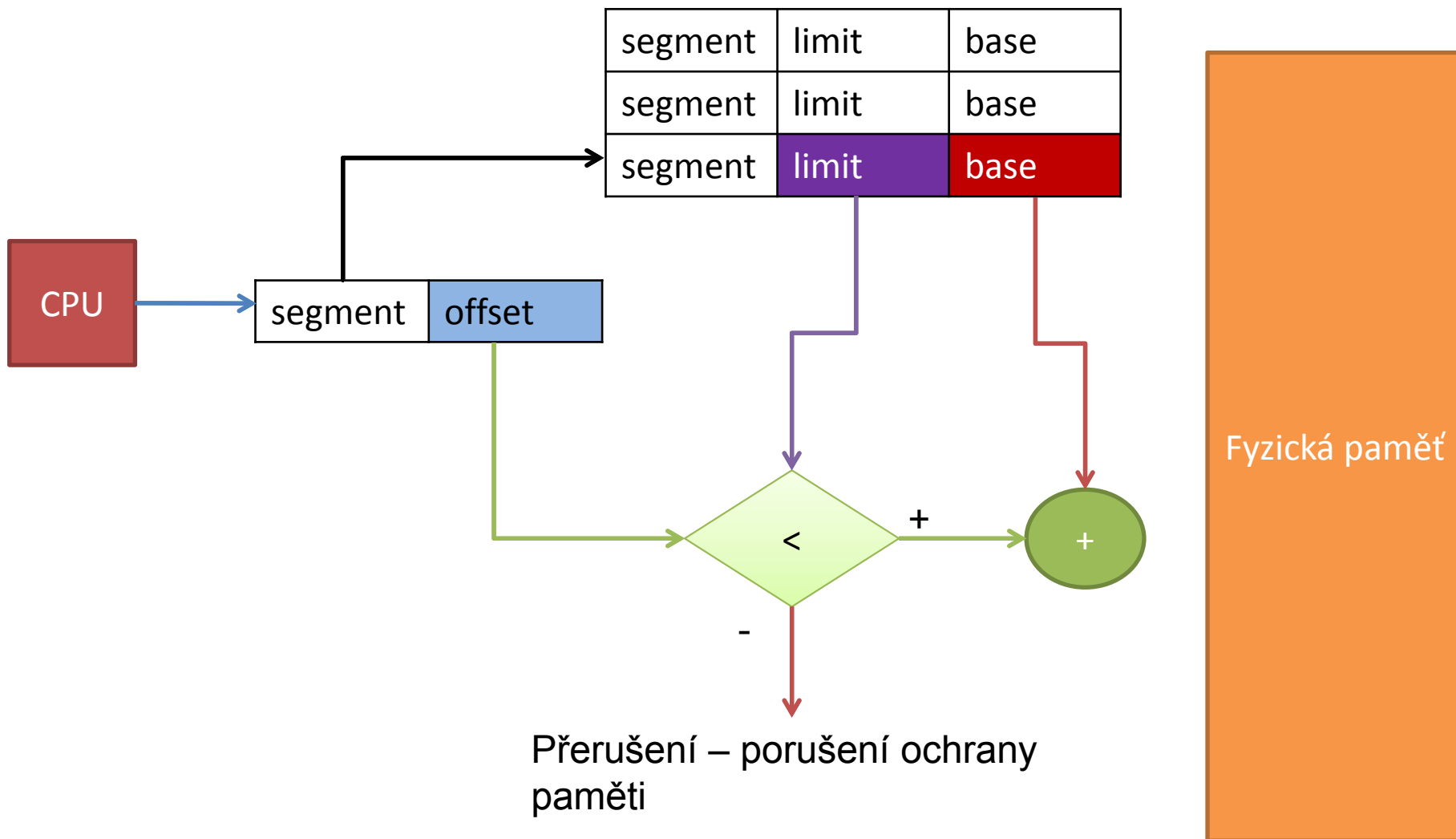
- = Je zcela nový přístup k paměťovému prostoru, který by paměť lépe využíval a zároveň i usnadňoval programování;
- = Jaké je vidění paměti uživatelem



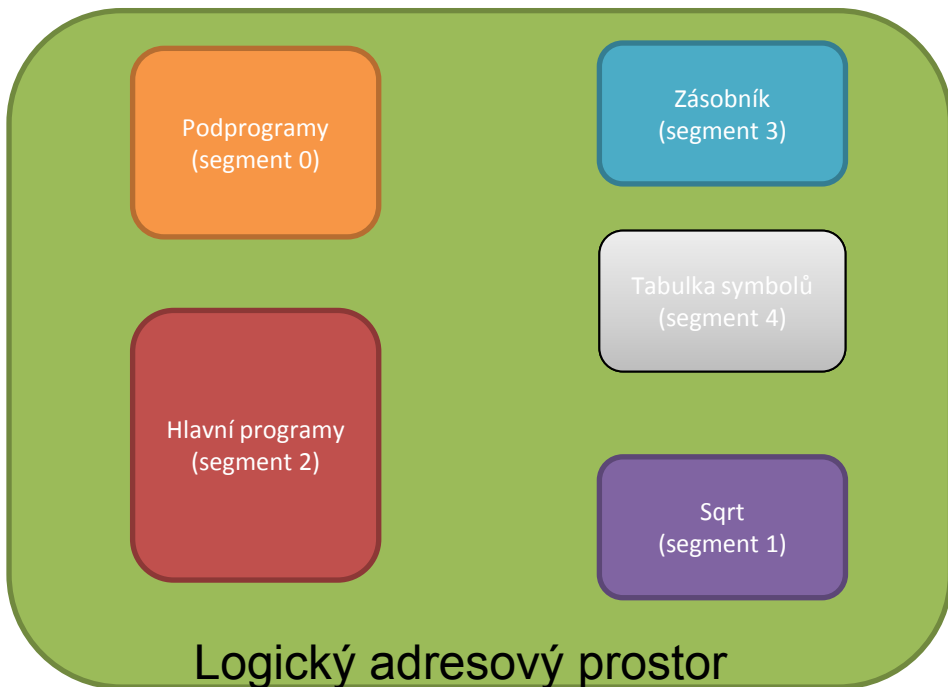
- = Segment = logické seskupení informací;
 - = adresový prostor každé úlohy je potom tvořen několika segmenty;
 - = technika přidělování paměti segmentům se nazývá segmentace;
- = Při segmentaci paměti musí každý odkaz do paměti obsahovat:
 - = číslo segmentu - (segment)
 - = paměťové místo v segmentu - (offset)

= Technické prostředky:

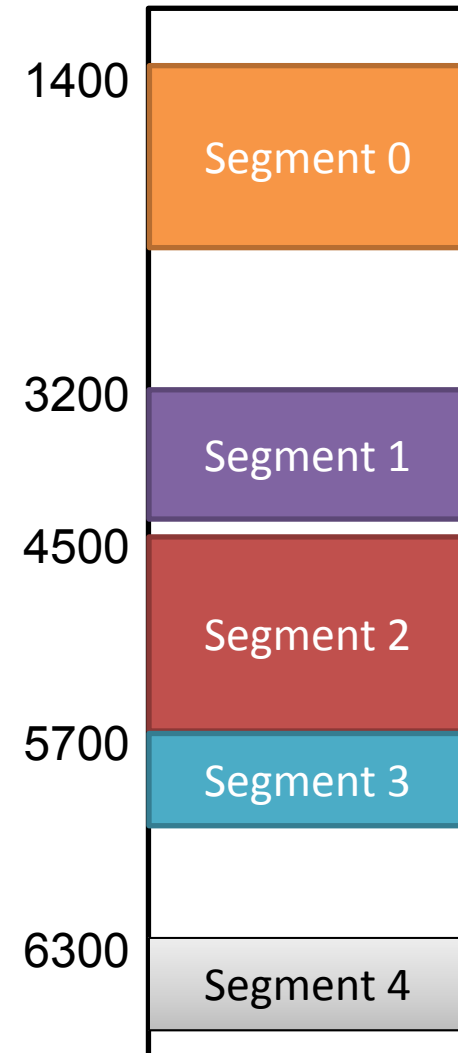
- = přemapování dvourozměrné adresy (segment, offset) na jednorozměrnou fyzickou adresu;
- = využívat převodní tabulku;
- = každý proces má svou tabulku segmentů;
- = logická adresa se skládá z čísla segmentu (segment) a relativní adresy slova v rámci segmentu (offset);
 - = segment se použije jako index v tabulce segmentů;
 - = offset, jakožto relativní adresa v rámci segmentu a nesmí být vyšší než limit (délka segmentu);
 - = pokud tomu tak je, nastává přerušení v důsledku porušení ochrany paměti;
 - = je-li vše v pořádku sečte se offset a base segment, čímž vznikne fyzická adresa, kterou je možno použít v paměti;



Segmentace paměti



limit	base
1000	1400
400	3200
1200	4500
300	5700
350	6300



- = stejně jako u stránkování mohou být tabulky segmentu uloženy buď v paměti nebo v rychlých registrech;
- = v případě velkého množství segmentů není uložení tabulek v registrech proveditelné a je nutno je ukládat do paměti;
- = segment-table length registr (STLR) obsahuje počet segmentů aktuální úlohy;

- = Ochrana a sdílení segmentu:
 - = každý segment je zabezpečen zvlášť;
 - = přístup k jednotlivým bytům segmentu je stejným jak pro segmenty obsahující instrukce, tak segmenty obsahující data;
 - = při uložení pole do segment memory-management hardware automaticky kontroluje indexy pole jestli jsou v pořádku a nepřekračují hranici segmentu;
 - = možnost sdílení (sharing) kódu nebo dat jednotlivými úlohami;

- = *Přidělování jediné souvislé oblasti paměti*
- = *Přidělování paměti po sekcích*
- = *Dynamické přemístování sekcí*
- = *Stránkování*
- = *Stránkování na žádost*
- = *Segmentace*
- = ***Segmentace a stránkování na žádost***

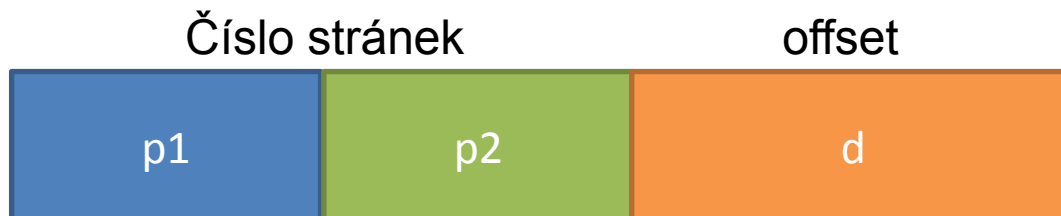
- = Kombinace segmentace a stránkování přináší vylepšení obou těchto technik;
- = Adresový prostor každého procesu je rozdělen na 2 části:
 - = 8192 segmentů může být privátních pro každý proces (Local Descriptor Table - LDT);
 - = 8192 segmentů může být sdílených (Global Descriptor Table - GDT);

- = Logická adresa je dvojice (selector, offset), kde selektorem je 16-bitové číslo:
 - = s - číslo segmentu (13 bitů);
 - = g - indikuje, je-li segment v GDT nebo LDT (1 bit);
 - = p - informace pro ochranu paměti 2 bity;
- = offset je 32-bitové číslo udávající relativní adresu slova v rámci daného segmentu;



- = Fyzická adresa je 32-bitové číslo
 - = select registr ukazuje na příslušnou položku v LDT nebo GDT;
 - = z ní jsou užity délka a počáteční adresa aktuálního segmentu pro vytvoření lineární adresy;
 - = lineární adresa = offset + počáteční adresa segmentu;
 - = každý segment je stránkován a velikost stránky je 4kB;
 - = tabulka stránek musí tedy obsahovat více než 1 mil. položek;
 - = každá položka zabírá 4B
 - = každý proces může požadovat až 4MB adresového prostoru pro svou tabulku stránek.
 - = složitá implementace

- = Lineární adresa je rozdělena na dvě části
 - = číslo stránky (20 b) a offset (12 b);
 - = číslo stránky je rozděleno na ukazatel do adresáře stránek (10 b) a ukazatel na stránku (10b);
 - = pro zvýšení rozsahu fyzické paměti může OS swapovat tabulky stránek na disk;
 - = jeden nepoužitý bit v položkách ukazatel do adresáře stránek je použit k indikaci, zda je daná položka na disku nebo v paměti;



- = Segmentace na žádost:
 - = stránkování na žádost je všeobecně uznávané jako nejlepší mechanismus pro vytvoření virtuální paměti;
 - = vyžaduje silnou hardwarovou podporu;
 - = pokud hardware není k dispozici, lze vytvořit virtuální paměť pomocí segmentace na žádost;
 - = systém využívá tabulek segmentu pro jednotlivé procesy, kde jsou informace o uložení, ochraně a velikosti segmentu;
 - = jelikož se segment ve fyzické paměti vyskytovat může i nemusí, má každá tabulka navíc indikační bit, který o tom informuje;

- = Je-li adresován segment, který není ve fyzické paměti, je vyvoláno přerušení (segment fault);
 - = není-li ve fyzické paměti dost místa pro natažení nového segmentu, je některý ze segmentů v paměti vyswapován na disk;
- = Je-li třeba vybrat segment pro odložení na disk, je vybrán ten na konci fronty;
 - = je-li pro natažení segmentu volné místo v paměti, je do něj segment natažen;
 - = aktualizují se záznamy v tabulce segmentu a identifikace segmentu je zařazena na vrchol fronty;



Univerzita Hradec Králové
Fakulta informatiky a managementu

Děkuji za pozornost...

