Увод у организацију и архитектуру рачунара 2

Александар Картељ

kartelj@matf.bg.ac.rs

Напомена: садржај ових слајдова је преузет од проф. Саше Малкова

Виртуелна меморија

Увод

Улога виртуелне меморије

- Концепт виртуелне меморије омогућава да програми у рачунарском систему употребљавају већи меморијски простор него што је стварна величина физички присутне меморије

Претходници

- Пре развоја технике виртуелне меморије
  - или је програм са подацима морао да стане цео у меморију
  - или је програмер морао експлицитно да управља тзв. механизмом преклапања (енгл. *overlay*)

- Техника виртуелне меморије аутоматизује старање о меморији и ослобађа програмера сувишног старања о физичким ресурсима
  - аутоматски се подаци и делови програма пребацују из физичке меморије на диск и обратно

Основне функције

- Две основне функције које остварује техника виртуелне меморије су:
  - Премештање
    - сваки програм користи свој виртуални адресни простор
    - током извршавања тај адресни простор се може пресликавати у различите физичке меморијске локације
    - детаљи управљања овим пресликавањем немају никакве везе са имплементацијом програма
    - сву бригу око пресликавања воде процесор и оперативни систем
  - Заштита
    - раздвојеност адресних простора програма пружа међусобну изолованост програма и заштиту података и кода

ВМ и кеш

- Виртуелна меморија и кеш деле неке концепте и претпоставке
  - успешност примене виртуалне меморије и кеша почива на локалности простора и времена
  - као што је кеш мањи и бржи од главне меморије, тако је главна меморија бржа и мања од диска
- Разлике
  - различита мотивација и намена имају већи број последица
  - различит ниво перформанси омогућава да се део технике ВМ имплементира у софтверу, док се све у вези кеша имплементира искључиво у хардверу

Виртуелна меморија

Страничење и пресликавање виртуелних адреса

Основи концепта ВМ

- Основи концепта виртуелне меморије су:
  - организација меморије по страницама
  - пресликавање виртуалних и физичких адреса
  - страничење помоћу диска

Странице меморије

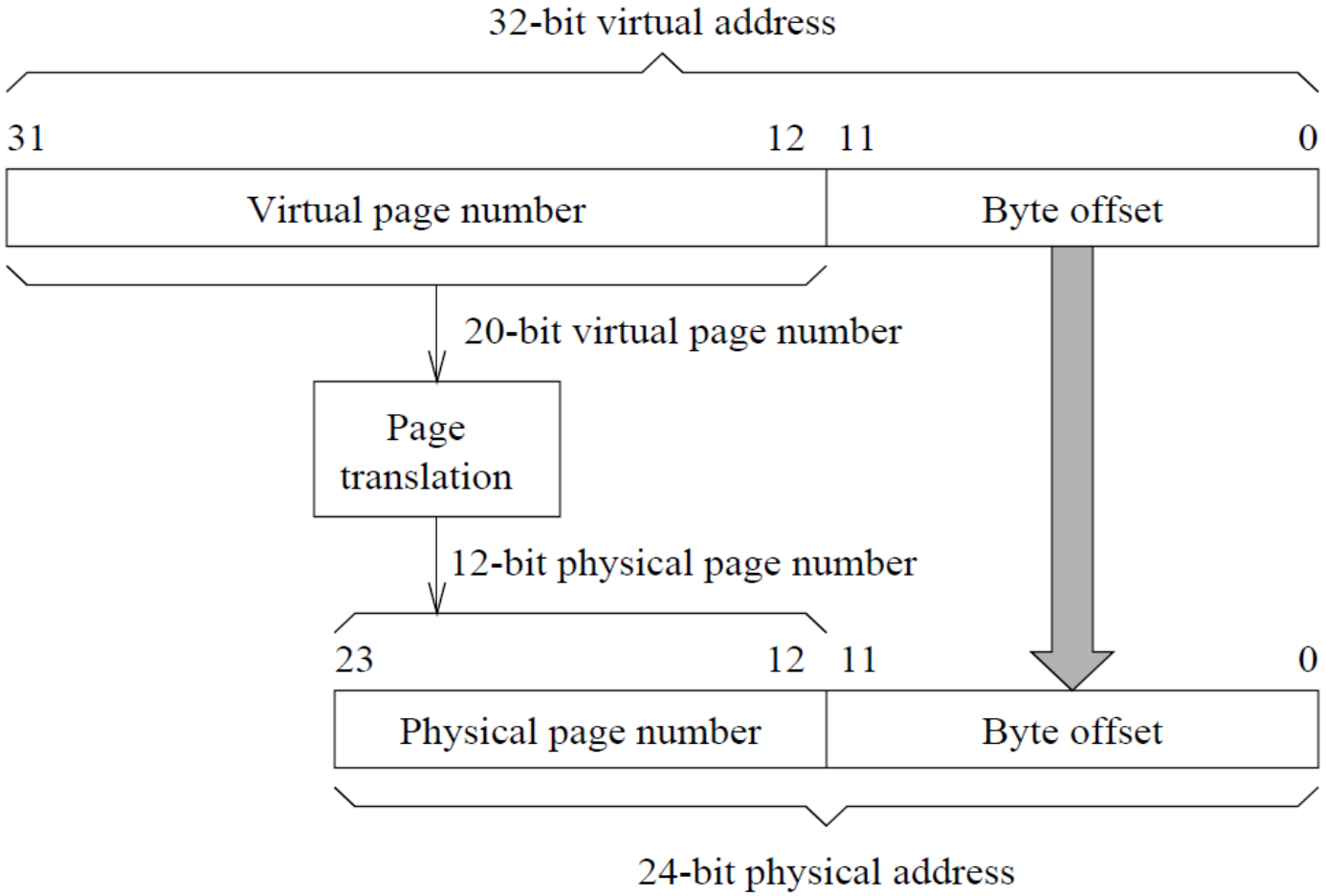
- Целокупан виртуелан адресни простор се дели на *виртуелне странице*
  - Битови виртуелне адресе се деле на *број виртуелне странице* и *адресу у страници*
- Слично томе, физичка меморија се дели на *физичке странице* (или *оквире за странице*)
  - Битови физичке адресе се деле на *број физичке странице* и *адресу у страници*

Пресликавање адреса

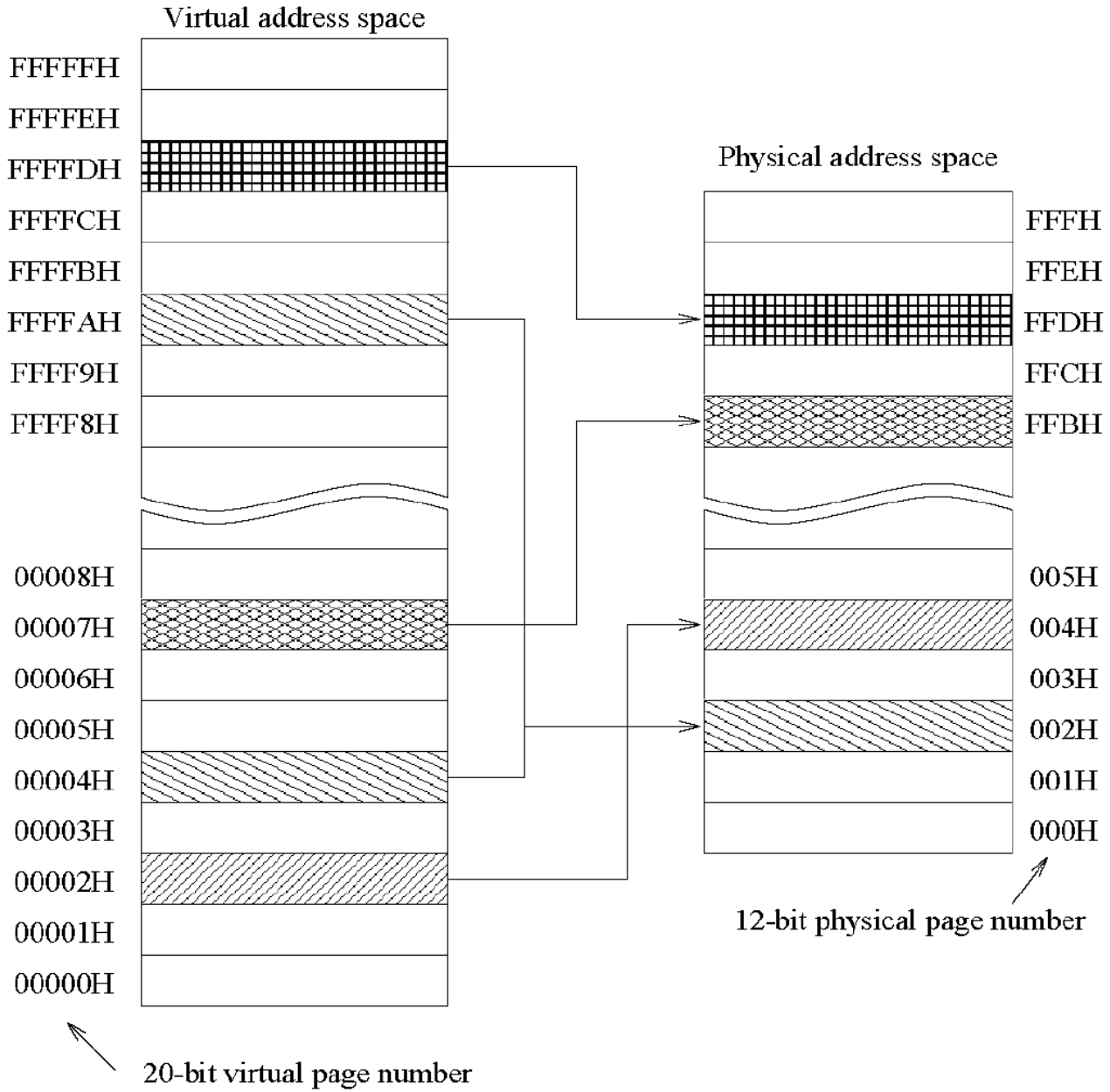
- При пресликавању виртуелних адреса у физичке
  - претпоставља се да су физичке и виртуелне странице исте величине
  - адреса у оквиру странице се не мења пресликавањем
  - број виртуелне странице се пресликава у број одговарајуће физичке странице
  - уобичајена величина странице је 4*KiB*
- За пресликавање је задужена јединица за управљање меморијом (енгл. *memory management unit*)

Пример

- Пресликавање 32-битне виртуелне у 24-битну физичку адресу, са страницом од 4*KiB*



Пример (2)


Virtual address space
FFFFFH
FFFFEH
FFFFDH
FFFFCH
FFFFBH
FFFFAH
FFFF9H
FFFF8H
00008H
00007H
00006H
00005H
00004H
00003H
00002H
00001H
00000H
20-bit virtual page number
Physical address space
FFFH
FFEH
FFDH
FFCH
FFBH
005H
004H
003H
002H
001H
000H
12-bit physical page number

Употреба диска

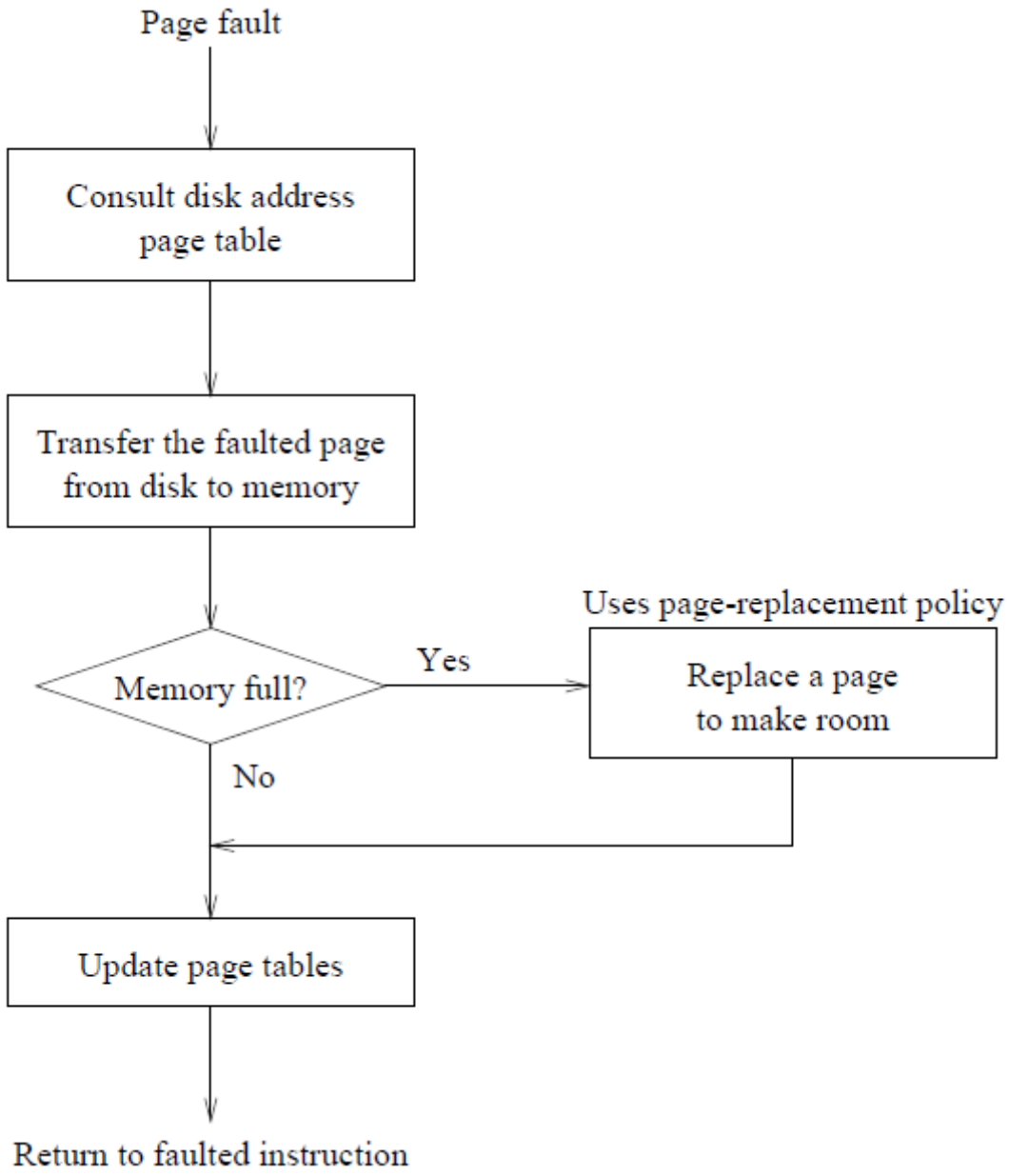
- Из угла ОС-а свака страница ВМ је
  - у главној меморији или
  - на диску
- Штавише, и странице које су у главној меморији имају своју слику на диску

Страничење по захтеву

- Ако се покуша приступити садржају странице која није у главној меморији, тада настаје *грешка страничења* (*странична грешка*, енгл. *page fault*)
  - има сличности са промашајем кеша
  - у том случају ОС је задужен да одреагује и учита тражену страницу у главну меморију и ажурира податке који се односе на пресликавање страница
  - ово се назива *страничење по захтеву*, зато што се обавља када се захтева дата страница

- Ако је главна меморија пуна, примењује се *политика замене странице*

Страничење по захтеву



Имплицитно страничење

- ОС може да предузима страничење и имплицитно, тј. без експлицитно исказане потребе за страницом, уколико процени
  - да се страничењем неће значајно ослабити перформансе активних процеса
    - нпр. да се диск тренутно не употребљава
  - да се са релативно високом вероватноћом може претпоставити да ће у скорој будућности бити потребне неке странице које трнутно нису у физичкој меморији

Политике замењивања страница

- При разматрању политика замењивања узимају се у обзир разлике између ВМ и кеша
  - цена промашаја у случају ВМ је вишеструко већа него у случају кеша, па је већи значај добре политике
  - политике замењивања кеша се имплементирају у хардверу, а у случају ВМ у софтверу, што пружа већу слободу
- Због тога се у случају ВМ тежи постизању што квалитетнијег одабира страница, а по цену сложености алгоритма

Политике замењивања страница (2)

- Неке од политика замењивања су:
  - *FIFO*
  - политика друге шансе
  - политика ретко употребљаване странице
  - политика најдуже неупотребљаване странице

Политика *FIFO*

- Замењује се она страница која је најдуже у главној меморији
  - једноставна имплементација
  - мана је што не узима у обзир употребу странице
  - релативно слабе перформансе
  - ретко се примењује

Политика *друге шансе*

- Унапређена политика *FIFO*
- Замењује се она страница која је најдуже у главној меморији, а да није ниједанпут реферисана
  - релативно једноставна имплементација
    - додаје се по један додатни бит за сваку страницу, који означава да ли је поново употребљавана након иницијалног учитавања
  - умерено значајан допринос перформансама

Политика *ретко употребљаване*

- Замењује се страница која је најмање пута реферисана
  - релативно једноставна имплементација
    - свакој страници се додаје бројач реферисања
    - оперативни систем повремено поништава бројаче
  - умерено значајан допринос перформансама

Политика *најдуже неупотребљаване*

- Замењује се страница чији садржај најдуже није реферисан
  - иако се имплементира у софтверу а не у хардверу, ипак је непрактична пуна имплементација
  - најчешће се апроксимира
    - начелно слично као у случају кеша
    - али ипак се тежи бољој апроксимацији
  - ово је најчешће примењивана политика

Политике писања

- У случаја кеша разматрали смо писање са пропуштањем и писање са преписивањем

- У случају ВМ писање са пропуштањем није опција, због велике разлике у брзини диска и главне меморије
- Додатно, у случају ВМ постоји заштита на нивоу страница и процеса (попут закључавања) па је писање са преписивањем безбедније него у случају кеша
  - не постоји опасност да неко употребљава неажуриране податке из странице на диску

Виртуелна меморија

Величина странице, таблице страница

Улога величине странице

- Мале странице су добре због:
  - интерне фрагментације
    - величина података, програма и стека није цео број страница
    - просечно је пола странице неупотребљено
    - што је страница мања, искоришћеност је већа
  - бољег погађања
    - што је већа страница, то ће при њеном учитавању у главну меморију заступљеност непотребних садржаја бити већа

Улога величине странице (2)

- Велике странице су добре због:
  - величине таблице страница
    - што су веће странице, биће их мање, па су и таблице мање
  - времена приступа диску
    - време приступа диску је много веће него време читања
      - реда 10ms по приступу, 100*MiB*/s читање
      - читање 100 страница од 4*KiB* траје
        - 100 x (10ms + 4/100ms) = 10.04s
      - читање 25 страница од 16*KiB* траје
        - 25 x (10ms + 16/100ms) = 2.54s
    - веће странице редукују број приступа диску и убрзавају рад

Примери величине странице

- *Intel x86*
  - странице су уобичајено 4*KiB*
  - подржане су и странице величине 2/4*MiB*
    - почев од процесора *Pentium*
    - документовано тек од процесора *PentiumPro*
    - у основи 4*MiB*, али 2*MiB* ако се користи *PAE*
- *Power PC*
  - странице су 4*KiB*
- *MIPS R4000*
  - подржава 7 величина страница, од 4*KiB* до 16*MiB*

Начин пресликавања

- Због високе цене промашаја потребно је да буде што мањи степен промашаја
- Због тога се у случају ВМ уобичајено примењује пуно асоцијативно пресликавање страница
- Имплементира се применом *таблица транслација*
  - краће се називају *таблице страница*

Таблице страница

- Примитивна имплементација
  - у основном облику таблица садржи скуп парова
    - (број виртуелне странице, број физичке странице) или
    - (број виртуелне странице, локација на диску)
    - ту су и додатни подаци о стању страница
  - по правилу се приступа по броју виртуелне странице
  - оваква имплементација захтева сувишна уређивања

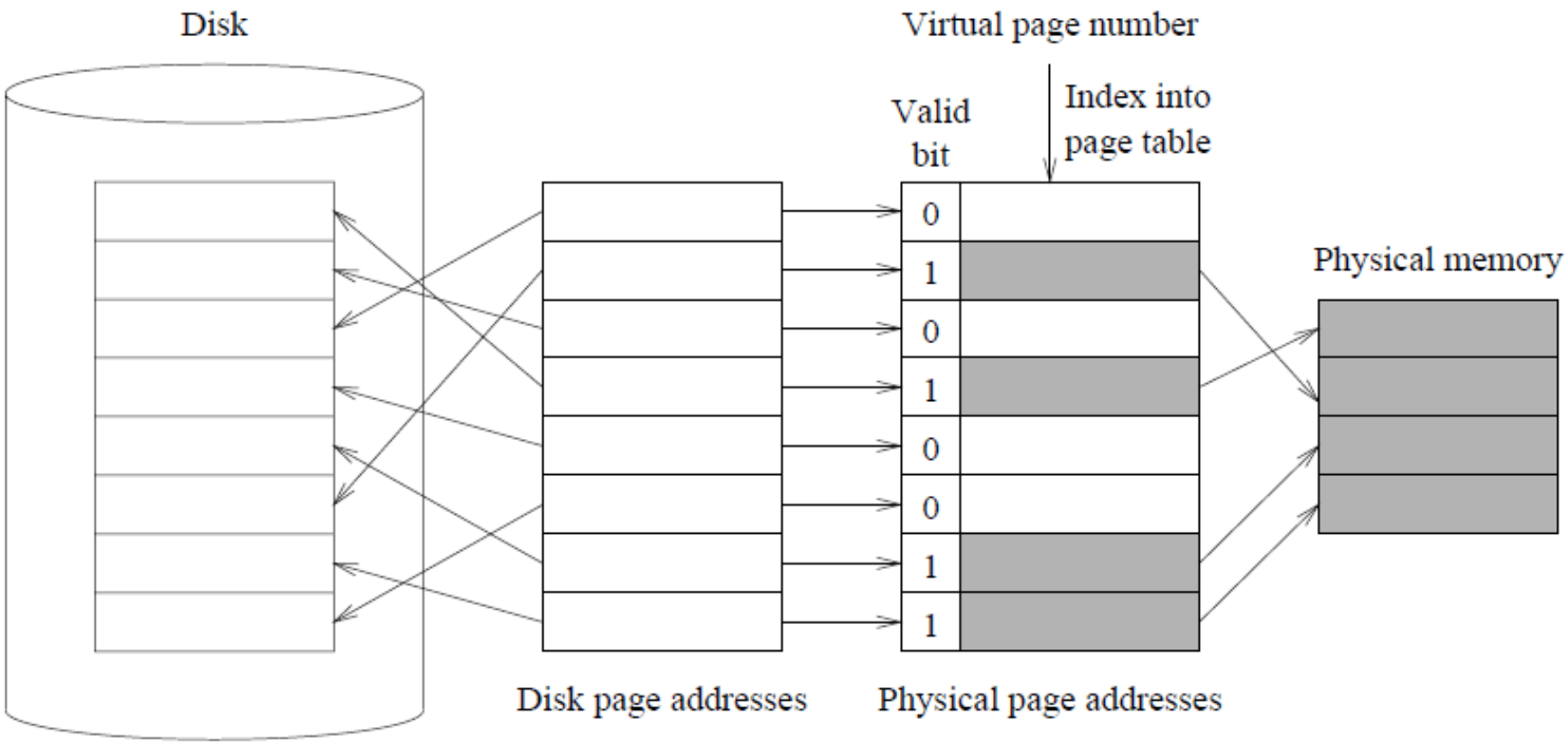
- Поједностављена имплементација
  - таблица се имплементира као низ индексиран бројевима виртуелних страница

Таблице страница (2)

- Може се моделирати са две таблице
  - таблица физичких страница
    - низ индексиран бројевима виртуелних страница
      - (енгл. *virtual page number - VPN*)
    - садржи бројеве физичких страница
    - садржи и додатне контролне битове
  - таблица адреса на диску
    - низ индексиран бројевима виртуелних страница
    - садржи локације страница на диску
- Обично се имплементира као једна таблица која садржи обе врсте података

Илустрација таблице страница

- за физичку меморију од 4,
  и ВМ од 8 страница



Ставке табеле страница

- Свака *ставка табеле страница* (енгл. *page table entry - PTE*) садржи:
  - број физичке странице (енгл. *physical page number – PPN*)
    - локација дате странице у главној меморији
    - води се за странице које постоје у главној меморији
  - адреса странице на диску (енгл. *disk page address*)
    - локација дате странице на диску
    - води се за све странице
  - бит исправности (енгл. *valid bit*)
    - означава да ли је страница у меморији или не
  - ...

Ставке табеле страница (2)

- ...
- бит измењености (енгл. *dirty bit*)
  - означава да ли је садржај странице мењан
  - ако је мењана, страница се при уклањању из главне меморије записује на диску
- бит реферисања (енгл. *referenced bit*)
  - користи се за имеплементацију алгоритма псеудо-*LRU*
  - ОС повремено поставља све битове реферисања на 0
  - при приступању страници бит се поставља на 1
- информација о власнику
  - ОС мора да зна ком процесу припада страница
- битови заштите
  - означавају тип приступа који власник остварује (само читање, читање и писање, извршавање,...)

Имплементација табеле страница

- Имплементација у софтверу је неефикасна
  - сваки приступ меморији би морао да се имплементира као бар два приступа меморији
    - приступање таблици пресликавања
    - приступање физичкој адреси податка
  - уобичајена употреба кеша може да омогући одређено убрзавање али то није довољно

- Имплементација у хардверу је скупа
  - савремени процесори имају велики адресни простор
  - таблице пресликавања би биле огромне и њихова имплементација у хардверу веома скупа

- Комбинована имплементација
  - у процесору се обезбеђује специјализован кеш за таблицу страница, тзв. *бафер таблице страница* (енгл. *translation lookaside buffer - TLB*)
  - имплементација свих алгоритама и политика је у софтверу – користи се само у случају промашаја

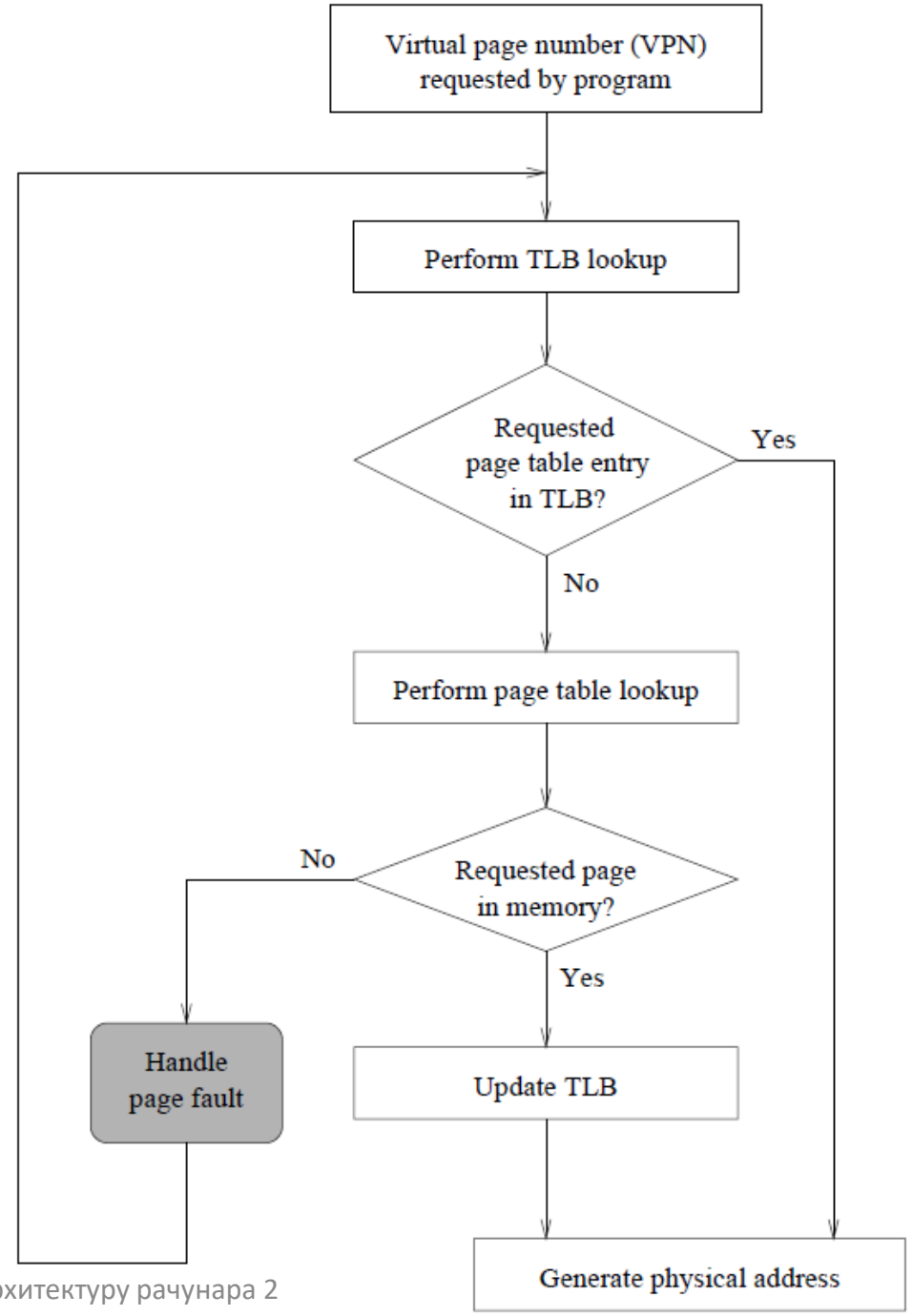
Бафер таблице страница

- *TLB* се имплементира у процесору
  - садржи податке о последњим коришћеним *PTE*
  - свака ставка *TLB* садржи:
    - број виртуалне странице (*VPN*)
    - одговарајући број физичке странице (*PPN*)
    - контролне битове
  - све ставке у *TLB* се налазе и у главној меморији у оквиру пуне таблице страница
    - допуњене додатним подацима, као што је локација на диску,...
  - као што кеш може бити подељен, тако и *TLB* може да се подели према намени – за инструкције и за податке
    - уобичајено је да уз подељен кеш иде подељен *TLB*
  - по правилу се имплементира са пуним асоцијативним пресликавањем

Употреба бафера таблице страница

- Алгоритам употребе:
  - најпре се подаци о страници траже у *TLB*
  - ако се пронађу, помоћу њих се рачуна физичка адреса
    - обично је правило да се страница реферисана из *TLB* не замењује, па је она већ у меморији
  - иначе се подаци траже у таблици страница
  - ако страница није у меморији, мора да се учита
  - рачуна се физичка адреса
  - приступа се садржају физичке меморије
- Политика замењивања ставки *TLB* је обично једноставна
  - политика случајног избора или
  - политика псеудо најдуже некоришћене

Транслација адреса помоћу бафера таблице страница



Локација таблице страница

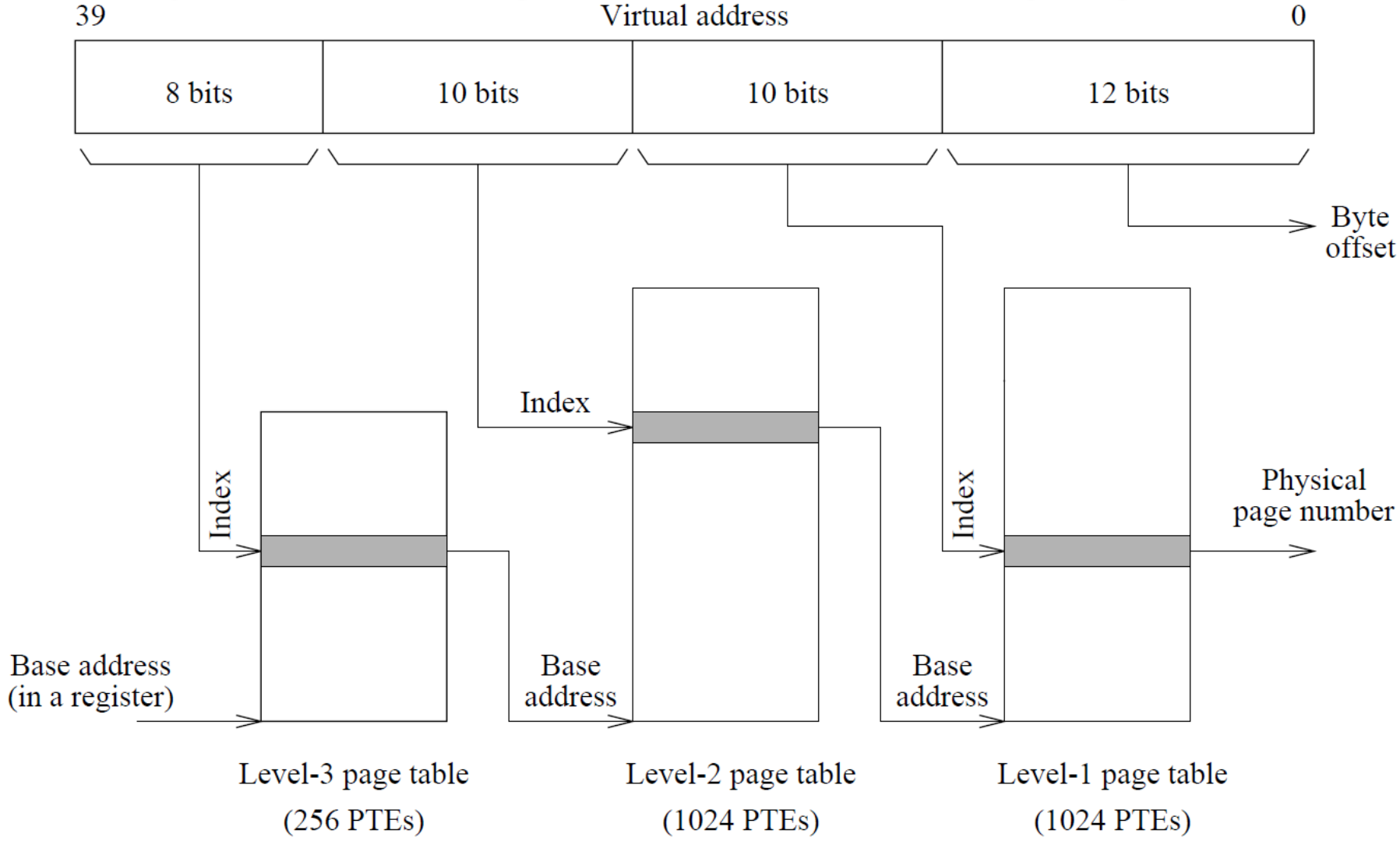
- Таблица страница може бити врло велика
  - пример: нека су виртуалне адресе 40-битне, величина странице 4*KiB*, а величина ставке 4 бајта
    - тада постоји 2^{28} страница
    - величина таблице је 1*GiB*

- Зато није практично (а понекад ни остварив) да се лоцира у физичкој меморији, већ се записује у виртуалној меморији
  - и таблица се дели на странице и само се потребне странице чувају у физичкој меморији
  - за ове странице се прави додатна таблица *другог нивоа*
  - по потреби се може правити више нивоа, све док се не дође до таблице чија је величина прихватљива за стално лоцирање у физичкој меморији

Пример

- Нека су виртуелне адресе 40-битне, величина странице 4*KiB*, а величина ставке 4 бајта
  - тада постоји 2^{28} страница
    - величина таблице првог нивоа је 2^{30}B = 1*GiB*
    - таблица првог нивоа заузима $2^{30}/ 2^{12} = 2^{18}$ страница
  - други ниво:
    - величина таблице другог нивоа је $2^{18} * 4$B = 1*MiB*
    - таблица другог нивоа заузима $2^{20}/ 2^{12} = 2^{8}$ страница
  - трећи ниво:
    - величина таблице трећег нивоа је $2^{8} * 4$B = 1*KiB*
    - таблица трећег нивоа стаје у једну страницу

Таблица страница организована у три нивоа



Инвертована организација таблице

- Проблем са описаном организацијом таблице страница је у величини таблице
  - за процесоре са 64-битним адресама, величина таблице може бити и 2^{54}B (!!!)
  - узрок проблема је употреба *VPN* као индекса таблице

- Алтернативни приступ је тзв. инвертована организација таблице
  - физичка адреса се користи као индекс
  - величина таблице је пропорционална физичкој меморији, а не виртуалној
  - штавише, постоји само једна таблица на нивоу система
    - у нормалној организацији постоји по једна за сваки процес
  - сложенија употреба

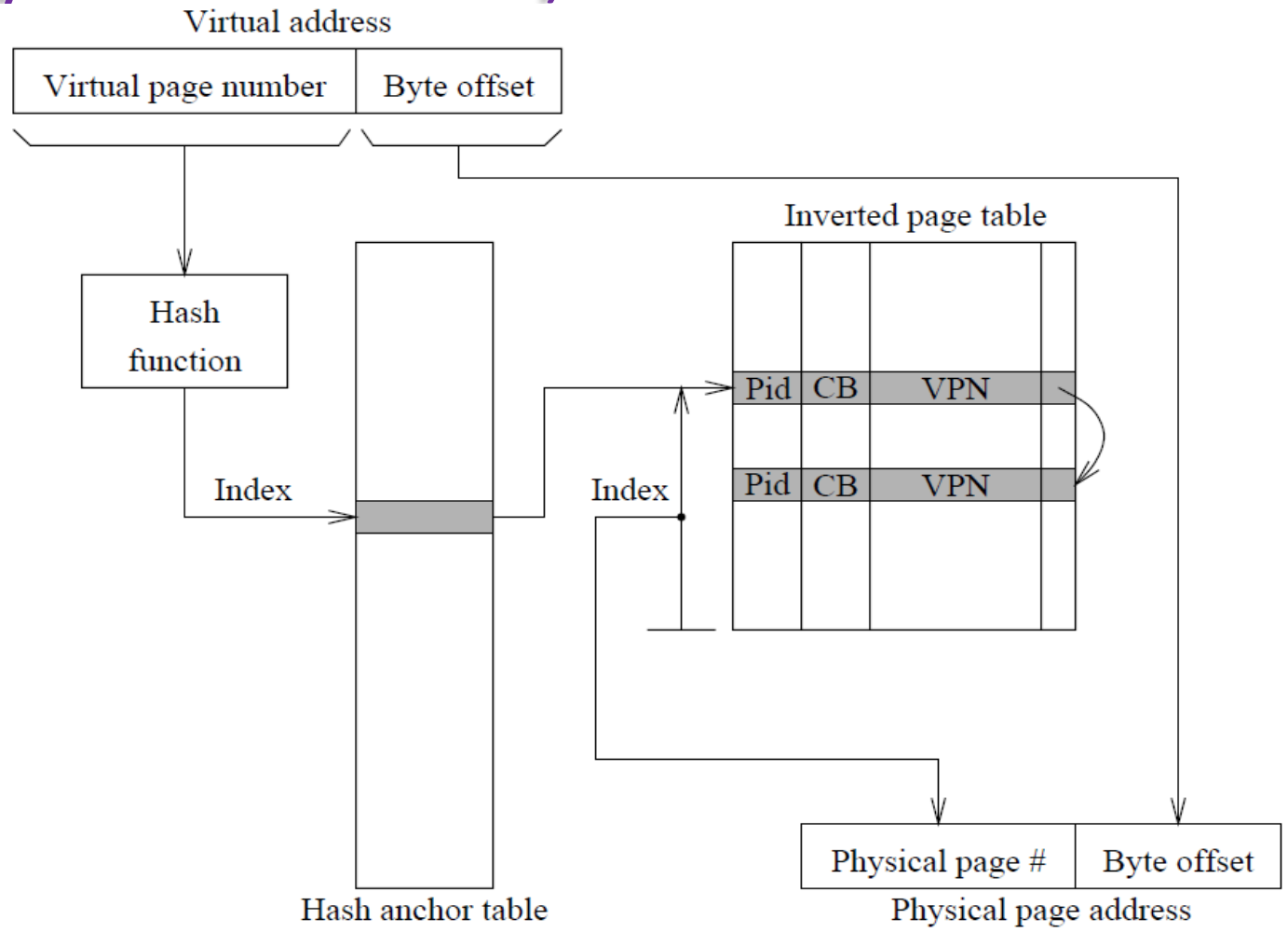
Инвертована организација таблице (2)

- Нама је потребно пресликавање *VPN* у *PPN*
  - ако су индекси *PPN*, ми заправо не можемо да их користимо
  - због тога се неком функцијом сецкања (енгл. *hash*) *VPN* пресликава у одговарајући индекс таблице
  - проблем са оваквим приступом је као и у случају функције непосредног пресликавања код кеша – више ведности се слика у исти индекс, тј. долази до *судара*

Обрада судара

- Постоје две опште технике за обраду судара при употреби функција сецкања:
  - Отворено уланчавање:
    - при писању
      - у случају судара се пише у првој наредној слободној локацији
      - све ставке са истим индексом се уланчавају
    - при читању
      - добијен индекс се користи као почетна тачка за тражење одговарајуће ставке
      - тражење се наставља низ ланац све до проналажења одговарајуће ставке
  - Секундарно сецкање:
    - у случају судара се примењује додатна функција сецкања за разрешавање судара

Инвертована организација таблице страница (примена уланчавања)



Сегментација

- Виртуализација представља линеарно пресликавање адресног простора
- Сегментација уводи додатну димензију пресликавања
- Основа концепта је да сваки процес добија *неколико независних* виртуалних адресних простора – *сегмената*
- Пуна адреса се састоји од две компоненте
  - броја (ознаке) сегмента и
  - адресе у оквиру сегмента

Сегментација (2)

- Сегментација је логичка организација која је видљива програмеру
- Пример: *Intel x86* процесори имају одвојене сегменте инструкција, података и стека
- Сегментација доноси
  - виши ниво заштите меморије
  - више адресних простора
  - могућност дељења сегмената међу процесима

Пример: *Intel Pentium*

- Подржава и виртуелну меморију и сегменте
  - свака од опција може да се искључи
  - *UNIX, Linux* не користе сегменте
- Сегментација пресликава 48-битне логичке адресе у 32-битне линеарне адресе, које се даље страничењем пресликавају у физичке адресе

Пример:
Intel Pentium (2)