

# معماری کامپیوتر

جلسه دهم: سلسله مراتب حافظه - TLB

# حافظه مجازی

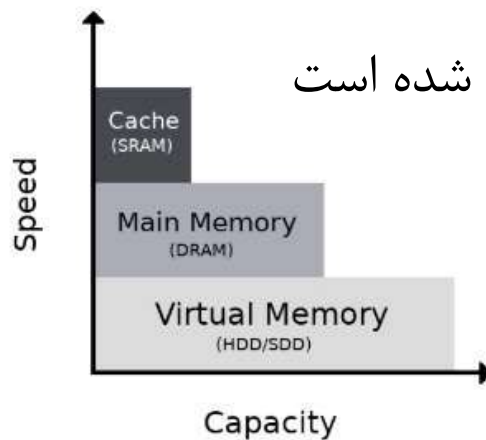


- هدف از حافظه نهان: تسريع ارتباط پردازنده با حافظه اصلي

- فرض ما در بحث‌های پيشين:

- دو سطح حافظه داريم و همه داده موردنياز در حافظه اصلي ذخيره شده است

- در واقعيت بيش از دو سطح داريم و بخشي از داده در ديسك ذخيره شده است



- دسترسي پردازنده به ديسك خيلي کند است

- حافظه مجازی: به‌عنوان cache برای ديسك

# حافظه مجازی



- فرض کنیم پردازنده ۳۲ بیتی و حافظه اصلی 16 MB داریم:
- قدرت پردازش:  $2^{32}$
- ظرفیت حافظه اصلی:  $2^{24}$
- برای افزایش کارایی و دسترسی به برخی آدرس‌ها لازم است به هارددیسک مراجعه کنیم:
- هارد دیسک نسبت به حافظه اصلی خیلی کند است
- تبدیل آن به حافظه مجازی (Virtual Memory) در کنار حافظه اصلی

# حافظه مجازی



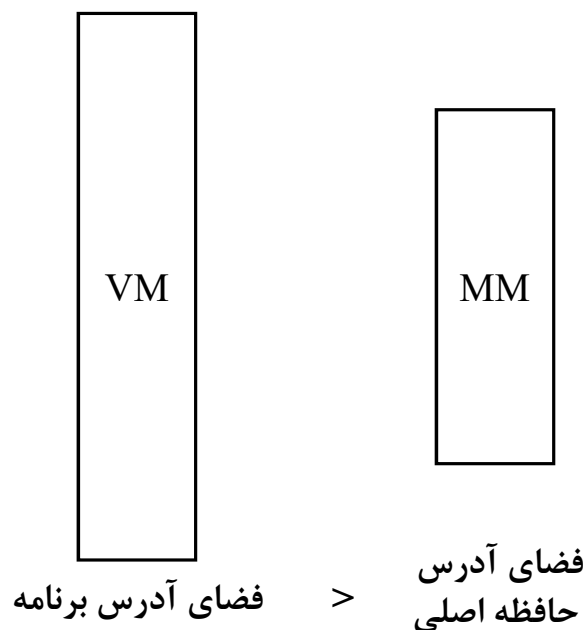
- کاربردهای دیگر حافظه مجازی

- تعریف فضای یکپارچه برای اجرای بخش‌های مختلف برنامه

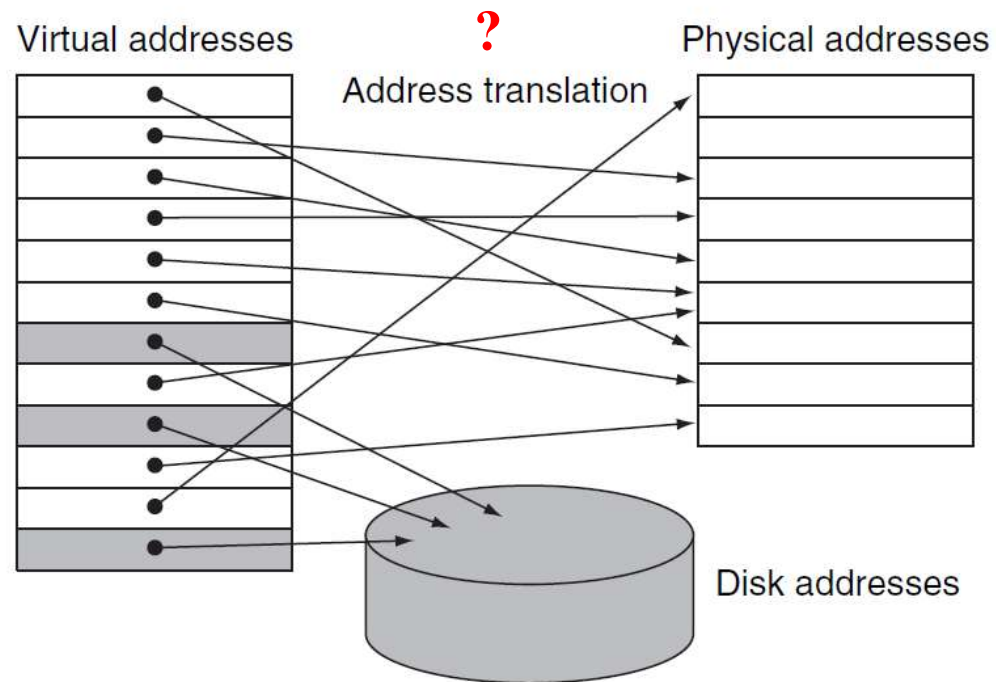
- تعریف فضای مستقل برای اجرای موازی چندین برنامه

- نحوه نگاشت فضای آدرس حافظه مجازی به حافظه اصلی؟

- ایده TLB



# حافظه مجازی



# حافظه مجازی



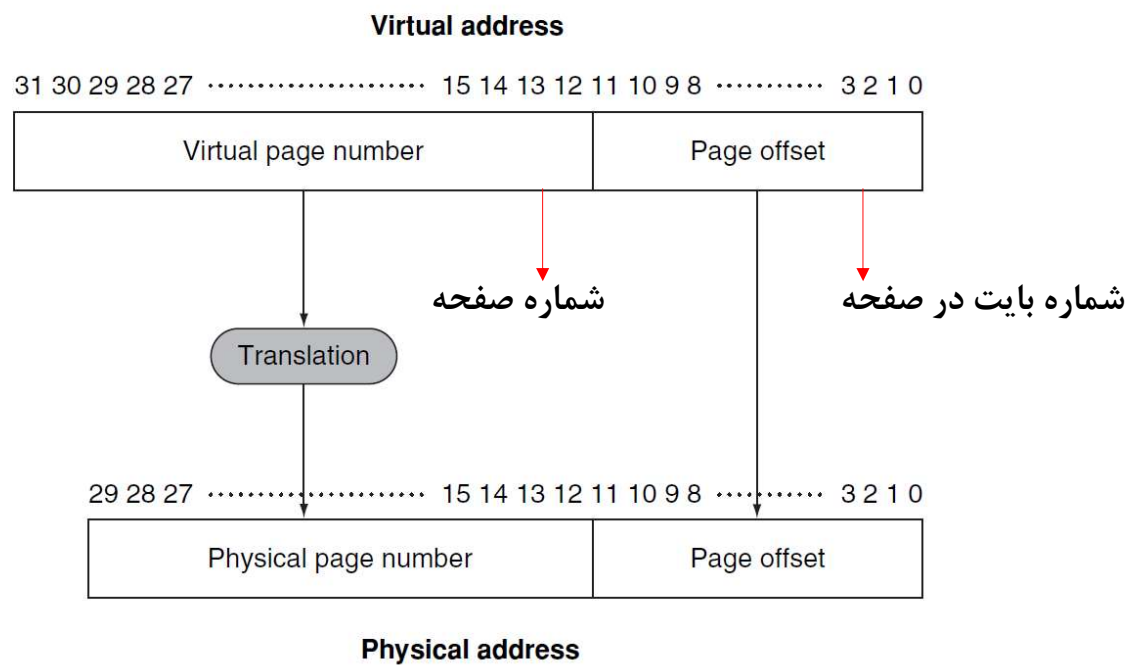
- در حافظه نهان، کل فضا به چندین set تقسیم می‌شد
- در حافظه مجازی، کل فضا به چندین page تقسیم می‌شود
- نگاشت از حافظه مجازی به حافظه اصلی با واحد page
- اندازه page وابسته به سرعت دیسک و ... است
- انتقال داده از هارد به حافظه اصلی در سطح page

$P'_0$
$P'_1$
$P'_2$

# حافظه مجازی



- نگاشت فضای آدرس حافظه مجازی و حافظه اصلی



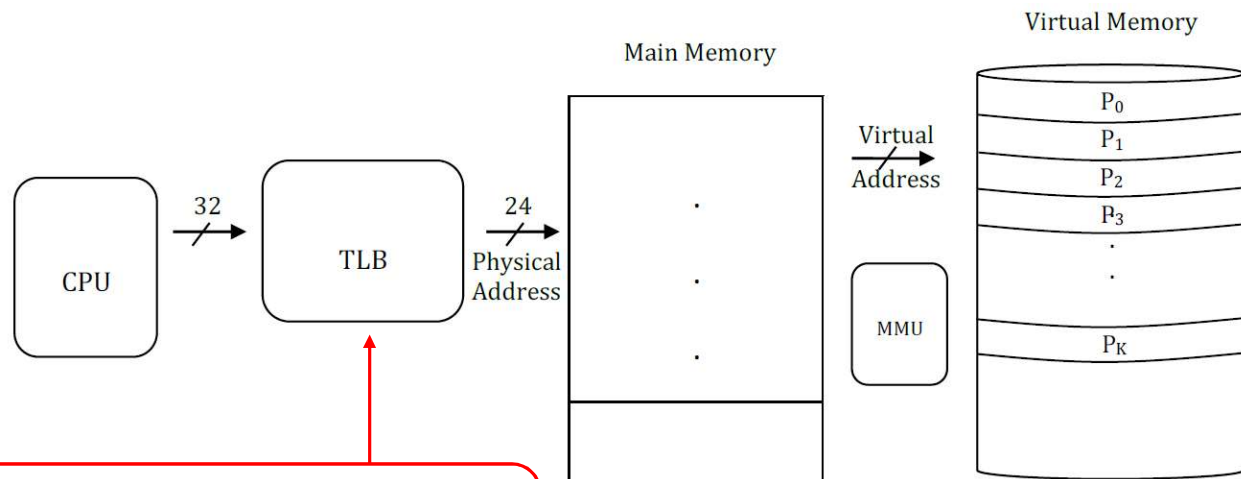
# TLB



- ترجمه آدرس مجازی به آدرس فیزیکی توسط TLB
- مخفف: Translation lookaside buffer
- TLB به صورت جدول ذخیره می شود

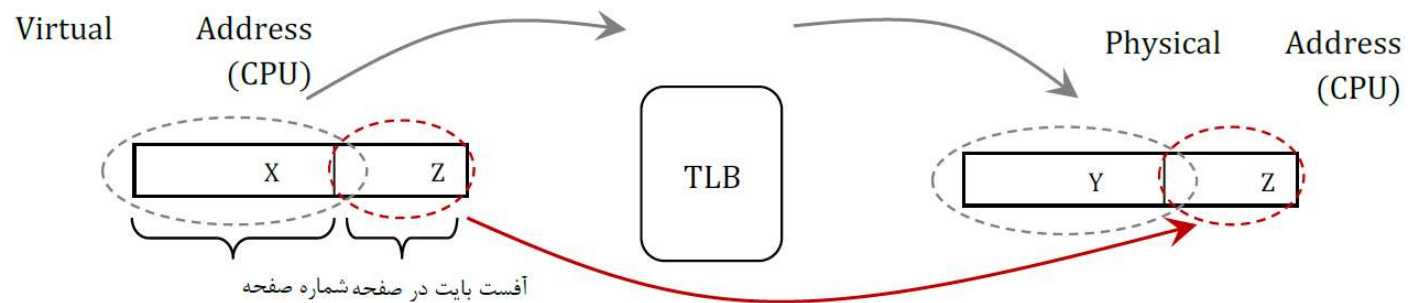


# TLB



- عملکرد به عنوان مترجم
- تبدیل آدرس مجازی پردازنده به آدرس فیزیکی حافظه اصلی  
(تبدیل شماره page اصلی به شماره آن در حافظه اصلی)

# TLB



# TLB



- ساختار TLB نمونه:

- سه page در حافظه اصلی وجود دارد

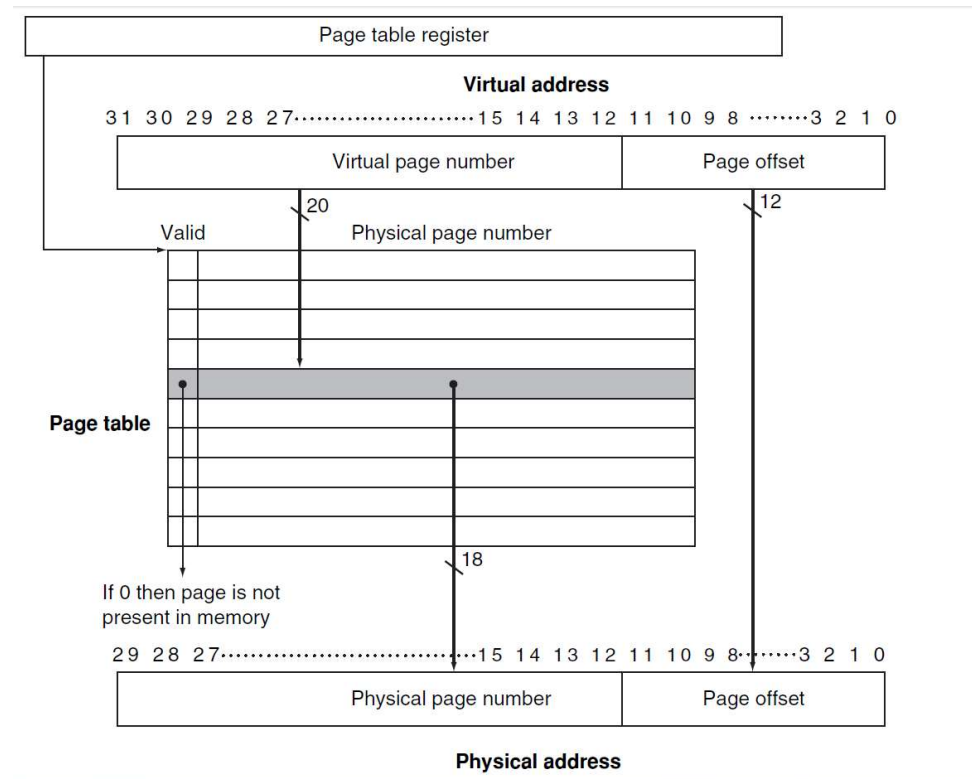
- بسیاری از سطرها خالی هستند

- سربار فضای زیاد

- راهکار: تغییر طراحی

آدرس مجازی	P	P'	آدرس فیزیکی
	0	2	
	1	0	
	2	X	
	3	X	
	.	.	
	.	.	
	.	.	
	232-	1	
	1		

# TLB



# TLB



- شیوه طراحی بهینه:

آدرس فیزیکی    آدرس مجازی

P	P'
1	0
99	1
0	2
2	3
.	.
.	.
.	.
6	K

- ثابت در نظر گرفتن آدرس فیزیکی و تخصیص آدرس مجازی بر حسب مکان HDD

- در نتیجه ساختار به صورت حافظه CAM طراحی می شود

- آدرس دریافت شده را با تمام سطرها (pageها) مقایسه کرده

- اگر موجود بود: hit و برگرداندن آدرس فیزیکی

- در غیر این صورت داده در حافظه اصلی نیست و باید از هارد دیسک آورده شود

# TLB



## • در صورت miss شدن آدرس در TLB (Page Fault)

- کل page به حافظه اصلی منتقل می شود

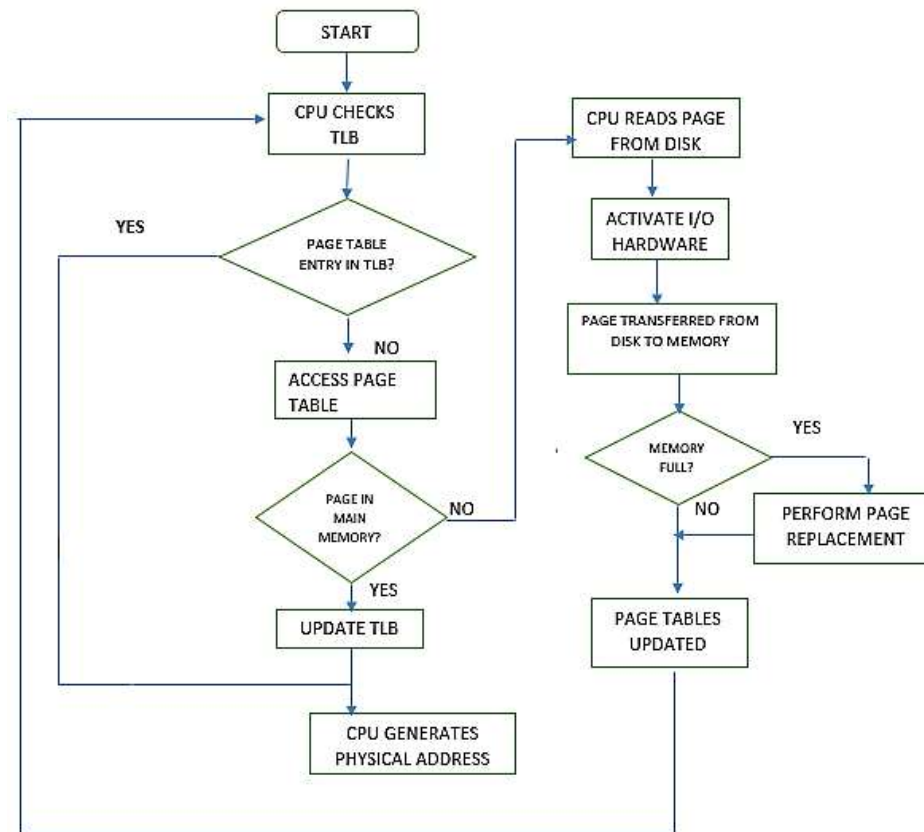
- بسیار زمان بر

- بیشتر توسط OS انجام می گیرد و نه سخت افزار

- اطلاعات TLB به روز می شود

- جایگزینی در حافظه اصلی: LRU

# TLB



# TLB



- مکان قرارگیری TLB:

- بعد از حافظه نهان: داده‌های HDD هم در حافظه نهان وارد می‌شوند
  - کارایی و هزینه بالا
- قبل از حافظه نهان: فقط داده‌های حافظه اصلی در حافظه نهان وارد می‌شوند
  - کارایی و هزینه کمتر
- معماری معمول: حالت دوم که حافظه نهان بعد از TLB است
  - معماری‌های مدرن: حالت موازی





برآورد کارایی

Performance Evaluation

# ارزیابی کارایی



- در طراحی اجزای مختلف سیستم‌های کامپیوتری، کارایی معیار بسیار مهمی است
- ارزیابی کارایی براساس بررسی روی برنامه‌های استاندارد انجام می‌گیرد.
- برنامه‌های استاندارد که اهداف معینی دارند: Benchmark
- گروه SPEC، برنامه‌های پرکاربرد را بررسی کرده و برآن اساس برنامه‌های سنجشی ارائه می‌دهند
- روش‌های طراحی و معماری‌های جدید توسط این محک‌ها تست می‌شوند
- برنامه‌ها با اهداف مشخص طراحی شده‌اند مثلاً حافظه‌ای، پردازشی و ....



# ارزیابی کارایی

- پارامتر کارایی در مقایسه روش‌های طراحی مختلف استفاده می‌شود
- مناسب بودن طراحی برای کاربرد خاص برحسب نیاز کاربر
- مقایسه دو طراحی و انتخاب بین آن‌ها

کاربر	کار مورد انتظار از رایانه	مفهوم کارا بودن رایانه برای کاربر	نگرانی کاربر پیرامون قیمت رایانه
سازمان هواشناسی	پردازش حجم زیادی از داده‌ها	قدرت پردازشی رایانه	نگران نیست
سازمان هوافضا	کنترل باله های موشک	سریع، دقیق و بی درنگ عمل کردن	نگران نیست
گرافیکست	کارهای گرافیکی سنگین	از پس تولید جلوه های بصری و تصاویر، در زمان کوتاهی برآید.	کمی نگران
کاربر خانگی	بازی، مرور اینترنت و برخی نیازهای روزمره	از هر نظر نسبتاً قابل قبول عمل کردن (سرعت، قدرت محاسباتی و ..)	کاملاً نگران

# ارزیابی کارایی



- تعریف کارایی

- Performance: مفهوم کلی است به معنای ایده‌آل کردن فرایندی در سیستم که این فرایند برحسب کاربرد می‌تواند زمان پاسخ، کلاک، سرعت، توان مصرفی و ... باشد.

- Throughput: مجموع کاری که سیستم در زمان معین می‌تواند انجام دهد (سرعت اجرای برنامه در پردازنده)

- نقطه مشترک تمامی کاربران: سرعت کارکرد سیستم‌های کامپیوتری

$$\text{Performance} \propto \frac{1}{\text{Execution Time}}$$

# ارزیابی کارایی



$$\text{Performance} \propto \frac{1}{\text{Execution Time}}$$

• منظور از زمان اجرا:

- مدت زمانی که طول می کشد تا یک سیستم کامپیوتری وظیفه مشخصی را به اتمام برساند
- شامل: زمان پردازش، دسترسی به حافظه، عملیات سیستم عامل و ...
- رابطه ذکر شده برای محاسبه کارایی استفاده می شود
- در ادامه هزینه هم در این رابطه وارد کرده و ارزیابی را جامع تر می کنیم

## ارزیابی کارایی



- **مثال:** اجرای برنامه‌ای روی کامپیوتر A در ۱۰ واحد زمانی و روی کامپیوتر B در ۱۵ واحد زمانی به اتمام می‌رسد، با فرض یکسان بودن همه مشخصه‌های این دو سیستم، کارایی این دو کامپیوتر را مقایسه کنید.

$$\frac{\text{Performance } A}{\text{Performance } B} = \frac{\text{Execution time } B}{\text{Execution time } A} = \frac{15}{10} = 1.5$$