

1. In fdisk, when using the p (print) command, there is a column labeled "sectors." What does "sector" refer to in this context?

سکتور یک بخش کوچک از فضای دیسک است که داده‌ها در آن ذخیره می‌شوند. معمولاً هر سکتور 512 بایت یا 4096 بایت فضا دارد. این ستون نشان می‌دهد که هر پارتیشن از کدام سکتور شروع می‌شود و تا کدام سکتور پایان می‌یابد. همچنین تعداد کل سکتورهای آن پارتیشن نیز ممکن است نمایش داده شود. با دانستن محل شروع و پایان سکتورها سیستم عامل می‌داند داده‌ها را از کجا بخواند یا بنویسد، آیا پارتیشن‌ها به درستی چیده شده‌اند یا تداخل دارند. برای بهینه‌سازی عملکرد دیسک، اطلاع از سکتورها ضروری است.

2. There are three types of servers mentioned. Please summarize your partitioning recommendations for each:

- Linux desktop systems

برای **/boot** حدود 250 الی 500 مگابایت ، برای / که حاوی داده‌های root می‌باشد حدود 30 الی 50 گیگابایت و برای **/home** بیشترین فضای باقی مانده دیسک مدنظر می‌باشد.

- Linux servers used for databases or web services with extensive logging

- Linux servers used in university labs or staging environments where multiple users require individual home directories and personal storage space

3. How does an operating system run multiple applications when the total available RAM is less than the combined memory requirements? Please summarize:

- How processes are scheduled to run on the CPU

سیستم‌عامل از Scheduler استفاده می‌کند تا پراسس‌ها را به نوبت روی CPU اجرا کند. هر پراسس برای مدت زمان مشخصی که به آن time slice گویند اجرا می‌شود. CPU بین پراسس‌ها سوئیچ می‌کند تا اجرای همزمان ایجاد شود. (چند وظیفگی multitasking)

- How process data is loaded into RAM

سیستم‌عامل فقط بخش‌هایی از برنامه‌ها را که در حال استفاده هستند به RAM می‌آورد، نه همه آن‌ها را. و این کار با کمک مکانیزمی به نام demand paging انجام می‌شود بطوریکه وقتی برنامه بخشی از حافظه را نیاز دارد، آن بخش از دیسک به RAM منتقل می‌شود.

- What happens when RAM is insufficient for a new process

اگر RAM پر باشد و یک پراسس جدید بخواهد اجرا شود، سیستم‌عامل باید فضا آزاد کند. پیج‌های کمتر استفاده شده از RAM به دیسک منتقل می‌شوند (به فایل swap). سپس فضای آزاد شده به پراسس جدید داده می‌شود.

- How and when the operating system uses disk space (paging and swapping)

Paging زمانی استفاده می‌شود که فقط بخشی از حافظه یک برنامه به دیسک منتقل می‌شود و در صورت نیاز دوباره به RAM بازگردانده می‌شود.

Swapping در شرایط خاص، کل پراسس از RAM خارج و روی دیسک قرار می‌گیرد. فایل یا پارتیشن swap محلی در دیسک است که نقش حافظه مجازی را بازی می‌کند.

• The memory management strategies involved (e.g., demand paging)

Demand Paging : فقط وقتی صفحه ای از حافظه نیاز باشد، به RAM بارگذاری می‌شود.

Least Recently Used (LRU): یکی از الگوریتم ها برای انتخاب اینکه کدام صفحه از RAM حذف شود.

• Whether all pages of a process must be loaded into RAM for execution

لازم نیست تمام صفحات حافظه یک پراسس هم زمان در RAM باشند. فقط صفحات فعال یا مورد نیاز فعلی بارگذاری می‌شوند، و سایر بخش ها در صورت نیاز از دیسک بازیابی می‌شوند.

4. What is the Translation Lookaside Buffer (TLB), and what role does it play in memory management?

یک حافظه کش (Cache) بسیار سریع است که برای سرعت بخشیدن به فرآیند ترجمه آدرس‌های حافظه مجازی به آدرس‌های حافظه فیزیکی استفاده می‌شود.

TLB در مدیریت حافظه وقتی یک پراسس به آدرسی دسترسی پیدا می‌کند، ابتدا CPU به TLB نگاه می‌کند تا ببیند آیا نگاشت آدرس قبلاً ذخیره شده یا نه. اگر یافت شود یعنی TLB hit و تبدیل آدرس سریع انجام می‌شود. اگر یافت نشود یعنی TLB miss که سیستم به جدول صفحات مراجعه می‌کند، نگاشت را پیدا می‌کند، و آن را در TLB برای استفاده‌های بعدی ذخیره می‌کند.

5. What are a page, a virtual page, and a context switch?

- صفحه یا page کوچک ترین واحد حافظه در سیستم های دارای مدیریت حافظه مجازی است. حافظه (RAM) و فضای آدرس دهی مجازی به صفحه‌های با اندازه ثابت (مثلاً ۴ کیلوبایت) تقسیم می‌شوند.
- صفحه مجازی بخشی از آدرس مجازی یک پراسس است. برنامه ها با آدرس های مجازی کار می کنند و این صفحات مجازی در صورت نیاز به صفحات فیزیکی RAM نگاشت می شوند. اگر صفحه ای هنوز در RAM نباشد، سیستم آن را از دیسک (Swap) بارگذاری می‌کند.
- Context switch وقتی اتفاق می‌افتد که CPU از اجرای یک پراسس به پراسس دیگر سوئیچ می‌کند. در این فرآیند، وضعیت (Context) پراسس فعلی ذخیره می‌شود و وضعیت پراسس جدید بارگذاری می‌گردد.

Additionally, how does increasing swap space affect context-switching performance?

افزایش فضای Swap بطوریکه سیستم مجبور شود مکرراً صفحات را بین RAM و دیسک جابه‌جا کند می‌تواند باعث کاهش عملکرد Context Switch شود. بخصوص اگر حافظه RAM کم باشد، Context Switch ممکن است باعث تبادل داده بیشتر با دیسک شود، که بسیار کندتر از RAM است.

How does page size influence these effects?

در صفحه های بزرگ تر تبادل داده بین RAM و دیسک در هر بار بیشتر است در نتیجه سرعت بارگذاری بهتر است اما احتمال internal fragmentation بیشتر است.

در صفحه های کوچک تر از حافظه استفاده بهینه تر می شود اما تعداد بیشتر صفحه ها سبب بار بیشتر روی TLB و جدول صفحات و در نتیجه احتمال بیشتر برای TLB miss و کندی می شود.

6. What is a huge page?

Please explain its purpose and when it is used.

Huge Pages یا صفحات بزرگ، یک تکنیک مدیریت حافظه در سیستم های کامپیوتری مدرن هستند که با استفاده از بلاک های حافظه بزرگتر از اندازه صفحه پیش فرض (معمولاً 4 کیلوبایت)، به بهبود عملکرد کمک می کنند. هدف اصلی از به کارگیری Huge Pages کاهش سربار مدیریت حافظه و بهبود کارایی دسترسی به حافظه، به خصوص در سیستم هایی با حجم زیاد حافظه رم است. استفاده از آنها برای برنامه هایی توصیه می شود که به مقدار زیادی حافظه نیاز دارند و دسترسی مکرر به حافظه دارند مانند پایگاه های داده بزرگ، ماشین های مجازی، برنامه های علمی و محاسباتی با حجم داده بالا، برخی از برنامه های کاربردی با نیازهای حافظه خاص.

7. What is memory fragmentation in RAM, and what problems can it cause?

زمانی اتفاق می افتد که بلوک های حافظه با اندازه های مختلف در طول زمان تخصیص و آزاد می شوند. این فرآیند می تواند منجر به وضعیتی شود که فضا های خالی حافظه کوچک و پراکنده ای در بین بلوک های تخصیص یافته ایجاد شوند. در نتیجه، حتی اگر کل حافظه آزاد کافی برای تخصیص یک درخواست جدید وجود داشته باشد، سیستم عامل نمی تواند یک بلوک پیوسته با اندازه مورد نیاز پیدا کند و درخواست تخصیص حافظه با شکست مواجه می شود.

سبب مشکلاتی مانند کاهش کارایی سیستم، افزایش زمان تخصیص حافظه، افزایش سربار مدیریت حافظه، عدم امکان تخصیص حافظه حتی با وجود حافظه آزاد کافی، افزایش احتمال استفاده از Swapping، مشکلات در سیستم های Real-time می شود.