



Homework 1- Solution

Lectures 1, 2

Operating Systems

Dr. Javadi

Spring 2023

۱- با توجه به Direct Memory Access (DMA) به سوالات زیر پاسخ دهید.

الف) مفهوم و کاربرد DMA را به صورت خلاصه شرح دهید.

تکنولوژی DMA یا Direct Memory Access روشی برای انتقال داده‌ها از حافظه RAM به اجزاء دیگر کامپیوتر بدون پردازش مستقیم داده‌ها توسط CPU می‌باشد. با اینکه داده‌های ورودی و خروجی از کامپیوتر توسط CPU پردازش می‌شوند، اما برخی از داده‌ها نیاز به پردازش ندارند و یا می‌توانند توسط دستگاه دیگری در کامپیوتر مورد پردازش قرار گیرند. همین امر فلسفه‌ی وجودی DMA را فراهم می‌کند.

ب) چگونه استفاده از این فناوری در کامپیوترهای جدید باعث کاهش حجم کار پردازنده در یک عملیات I/O می‌شود؟

همانطور که می‌دانید CPU یا پردازنده سیستم در فرآیند فراخوانی اطلاعات از حافظه (RAM) دخیل می‌باشد و برای انجام پردازش روی داده‌ها بایستی اطلاعات را از حافظه RAM فرخوانی کند. همین امر موجب می‌شود که زمان زیادی برای انتقال داده‌ها از حافظه RAM به سایر دستگاه‌ها در کامپیوتر صرف شود. اینجاست که تکنولوژی DMA پا به میان می‌گذارد تا انجام این فرآیند با سرعت بیشتری صورت پذیرد و نیاز به پردازش اطلاعات توسط CPU را از میان برمی‌دارد. به همین صورت DMA از بار پردازشی CPU می‌کاهد و راهی بهینه برای انتقال داده‌ها از حافظه RAM به سایر دستگاه‌های موجود در کامپیوتر ایجاد می‌کند. همینطور، در حین انجام عملیات I/O با استفاده از DMA، پردازنده قادر است به صورت موازی برنامه‌های دیگری را اجرا کند.

پ) در حین انجام عملیات I/O به وسیله DMA، پردازنده می‌تواند برنامه دیگری را اجرا کند. آیا این کار می‌تواند باعث ایجاد اختلال بین عملیات I/O و برنامه در حال اجرا شود؟ توضیح دهید و در صورتی که پاسخ شما مثبت است، روش جلوگیری از این مشکل را شرح دهید.

CPU و DMA ممکن است برای دسترسی به BUS سیستم، که مسیر ارتباطی مشترک بین CPU، DMA و سایر اجزای سیستم است، رقابت کنند. اگر BUS مشغول سرویس دهی به یک دستگاه I/O باشد، دستگاه دیگر ممکن است مجبور باشد در انتظار بماند که همین امر موجب کندی در پردازش می‌شود. برای جلوگیری از این مشکل می‌توان از روش‌های اولویت‌بندی دستگاه‌ها استفاده کرد و برای مثال به CPU اولویت بالاتری برای دسترسی به BUS اختصاص داد. علاوه بر این، اگر CPU و DMA همزمان به یک ناحیه از حافظه دسترسی داشته باشند، این امر ممکن است باعث بروز اختلال و خطا در داده‌ها شود. برای جلوگیری از این امر، DMA ممکن است طوری پیکربندی شود که به طور موقت عملیات خود را متوقف کند، در حالی که CPU به همان مکان حافظه دسترسی دارد.

د) در آینده با مفاهیم Security و اهمیت آن در سیستم‌های کامپیوتری آشنا می‌شوید. یکی از انواع حمله‌هایی که به سیستم‌های کامپیوتری وارد می‌شود DMA Attack می‌باشد. نحوه انجام این حمله را به صورت مختصر توضیح دهید.

همانطور که گفته شد، از آن جایی که در DMA دسترسی مستقیم به حافظه فراهم است، مهاجمین می‌توانند با استفاده از DMA Attack به RAM دسترسی داشته باشند و اطلاعات دلخواه را بنویسند یا بخوانند. این روش می‌تواند به صورت فیزیکی توسط یک دستگاه متصل به کامپیوتر که قابلیت‌های DMA برای آن فعال است صورت گیرد یا از راه دور بر شبکه‌ای از دستگاه‌ها که به یکدیگر متصل هستند اعمال شود.

۲- قطعه کد زیر که به زبان اسمبلی نوشته شده است و در یک ماشین با سیستم عامل Linux با معماری ۶۴ بیتی اجرا می‌شود را در نظر بگیرید و به سوالات پاسخ دهید.

```

1 segment .data:
2     message db "AUT is Awesome!"
3     len equ $-message
4
5 segment .bss:
6
7 segment .text:
8     global _start
9     _start:
10        mov eax, 0x4
11        mov ebx, 1
12        mov ecx, message
13        mov edx, len
14        int 0x80
15
16        mov eax, 0x1
17        mov ebx, 0
18        int 0x80
19
```

الف) با توجه به عددی که استفاده شده، مشخص کنید که وقفه‌هایی که در خطوط ۱۴ و ۱۸ رخ می‌دهد، باعث ایجاد چه عملیاتی می‌شود؟

Intx080 یک دستورالعمل وقفه نرم افزاری است که در زبان اسمبلی x86 استفاده می‌شود. این وقفه برای ایجاد درخواست فراخوانی سیستمی به Kernel در لینوکس و سایر سیستم‌های عامل مشابه Unix استفاده می‌شود. هنگامی که دستور Intx080 اجرا می‌شود، پردازنده به حالت Kernel تغییر حالت می‌دهد و کنترل پردازنده به یک ISR خاص در کرنل واگذار می‌شود. اینکه چه فراخوانی سیستمی‌ای اجرا شود به مقدار موجود در ثبات EAX بستگی دارد.

ب) تحقیق کنید که اعداد 0x4 و 0x1 که در رجیستر eax قرار می‌گیرند، باعث اجرای کدام فراخوانی‌ها می‌شوند؟ در انتها نیز مشخص کنید که این قطعه کد چه کاری را انجام می‌دهد؟

بسته به نوع سیستم عامل، اینکه با مقادیر 0x4 و 0x1 چه فراخوانی سیستمی اجرا شود متفاوت است. در این سوال، مقدار 0x4 معادل فراخوانی دستور **write** است و 0x1 معادل دستور **exit** می‌باشد.

(پ) در مورد نحوه انتقال پارامترها در این قطعه کد تحقیق کنید و توضیح مختصر بنویسید.

در این قطعه کد انتقال پارامترها با نوشتن آن‌ها در ثبات‌ها انجام می‌شود. برای مثال در فراخوانی سیستمی **write** مقادیر ورودی که به ترتیب آدرس فایل، آدرس بافر حاوی داده و طول داده می‌باشد در رجیسترهای **ebx** و **ecx** و **edx** قرار داده شده‌اند. فراخوانی سیستمی **exit** ورودی‌ای نمی‌پذیرد و مقداری که در رجیستر **ebx** ریخته می‌شود توسط پردازنده برای محاسبه مقدار خروجی **exit** استفاده می‌شود.

(د) اگر این ماشین به صورت **dual mode** باشد، برای اینکه این کد به صورت موفقیت آمیز انجام شود، باید سیستم در چه حالت (mode) عملیاتی باشد؟

اجرای فراخوانی‌های سیستمی باید فقط در حالت **Kernel** انجام می‌شود ولی درخواست برای اجرای آن‌ها می‌تواند در حالت **User** ایجاد شود.

۳- در برخی از پردازنده‌ها، بیشتر از دو حالت عملیاتی (modes of operation) تعبیه شده است. حداقل دو پردازنده با این ویژگی را نام ببرید و دو استفاده ممکن از این حالت‌های چندگانه را شرح دهید.

در برخی از پردازنده‌ها بیشتر از ۲ مود اجرایی **kernel** و **user** تعبیه شده است که از آن‌ها برای مصارف خاصی مانند مجازی‌سازی، دیباگ و کنترل توان استفاده می‌شود.

برای مثال پردازنده‌های **Intel** با قابلیت **Intel Virtualization Technology** چندین مود کاری برای مجازی‌سازی دارند از جمله مودهای **root**, **guest**, **non-root**. به واسطه‌ی این مودهای کاری پردازنده قابلیت کنترل چندین ماشین مجازی با سطوح دسترسی مختلف را داراست. مثالی دیگر پردازنده‌های سری **ARM-v8** که دارای مود دیباگ هستند که به دولوپرها اجازه می‌دهد تا عملیات دیباگ را در سطح **low-level** به روی پردازنده و سیستم انجام دهند. یا مثال‌های دیگر...

۴- الف) نحوه بوت شدن سیستم عامل توسط برنامه bootstrap به صورت خلاصه توضیح دهید و تعدادی از وظایف برنامه bootstrap را نیز بنویسید.

با روشن شدن سیستم یا در هنگام reboot، برنامه Bootstrap اولین برنامه‌ای است که لود و اجرا می‌شود. Bootstrap یک firmware می‌باشد که در ROM یا EPROM ذخیره شده است و تمامی ویژگی‌ها و تنظیمات سیستم را مقداردهی اولیه می‌کند. از وظایف bootstrap می‌توان به بارگذاری و اجرای kernel، بررسی اولیه سخت‌افزار برای یافتن اشکال یا خطا، بارگذاری اولین بخش/sector از دیسک به روی RAM و اجرای آن اشاره کرد. در اولین بار روشن کردن سیستم، این بخش همان برنامه سیستم عامل است و از آن پس، سیستم با مراجعه به RAM، سیستم عامل نصب شده را اجرا می‌کند. پس از آن سیستم عامل برنامه‌های نرم‌افزاری اش را اجرا می‌کند و کنترل سیستم را به دست می‌گیرد.

ب) چگونه می‌توان سیستمی طراحی کرد که اجازه‌ی انتخاب یک سیستم عامل از چند سیستم عامل را هنگام بوت شدن به کاربر بدهد؟ برنامه‌ی bootstrap برای این منظور چه کاری باید انجام بدهد؟

زمانی که می‌خواهیم چند سیستم عامل داشته باشیم، در واقع سیستم به چند partition (هرکدام برای یک سیستم عامل) تقسیم می‌شود. برنامه bootstrap با روشن شدن سیستم، متوجه می‌شود که چند سیستم عامل روی سیستم نصب شده است و از کاربر می‌خواهد که سیستم عامل مورد نظر را انتخاب کند. سپس مانند توضیح داده شده در بخش الف عمل می‌کند، یعنی برنامه سیستم عامل مورد نظر را بر روی RAM می‌آورد و سپس اجرا می‌کند. برای هر سیستم عامل یک bootloader داریم.

۵- در مورد multiprocessing و multitasking به سوالات زیر پاسخ دهید.

الف) تفاوت‌های multiprocessing و multitasking را بیان کنید.

در Multiprogramming اساس کار بر مبنای context-switch است اما Multitasking بر مبنای time-sharing است. تمرکز اصلی Multiprogramming استفاده‌ی حداکثری از cpu است. اینکه cpu هرگز بیکار نباشد. اما تمرکز اصلی Multitasking کم کردن زمان پاسخگویی cpu است (کمتر از ثانیه به طوری که حس اجرای موازی را القا کند). در Multiprogramming بیشترین زمان ممکن صرف اجرای یک پردازش می‌شود اما در Multitasking کمترین زمان ممکن صرف اجرای یک پردازش می‌شود. Multiprogramming با وجود حجم کم حافظه هم امکان‌پذیر است ولی Multitasking نیازمند حجم زیادی از حافظه است چرا که تمام پردازش‌هایی که با روش Multitasking در حال اجرا هستند باید در حافظه بارگذاری شده باشند.

ب) اگر یک سیستم تک پردازنده و یک هسته‌ای داشته باشیم که از multiprogramming پشتیبانی کند، در هر لحظه چند پردازنده می‌تواند در وضعیت running باشد؟

در هر لحظه تنها یک پردازنده می‌تواند در حالت running باشد و با رفتن این پردازنده به حالت wait یا terminated، پردازنده بعدی انتخاب و اجرا می‌شود.

ج) فرض کنید سه پردازنده A, B, C داریم که فرآیند کاری آنها به صورت زیر است:

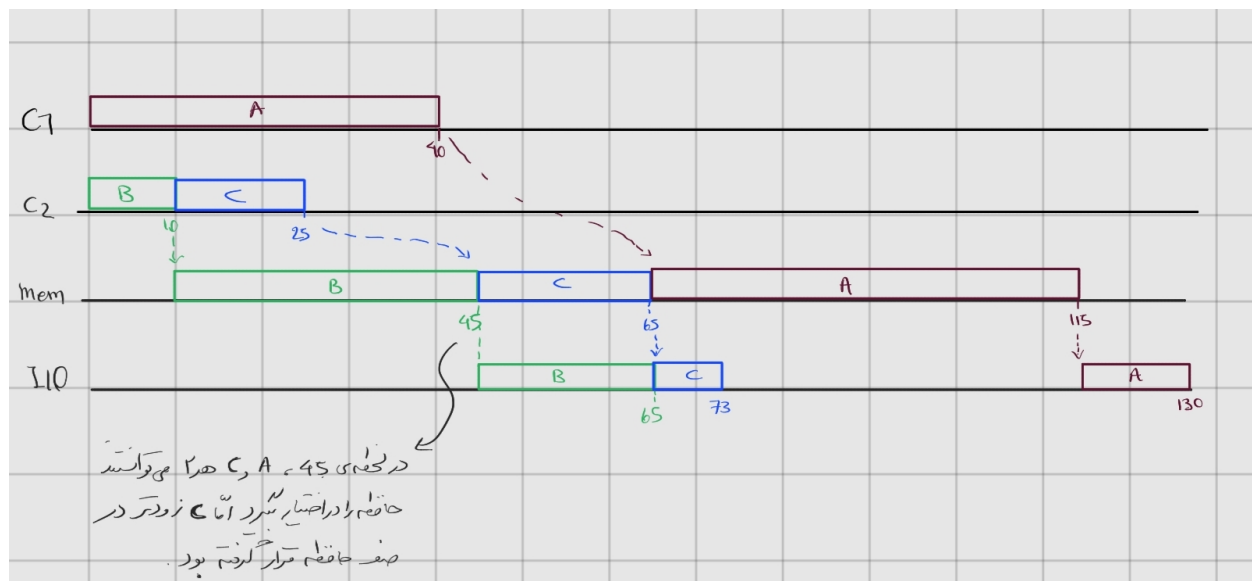
A: 50ns Memory, 40ns CPU, 15ns I/O

B: 35ns Memory, 10ns CPU, 20ns I/O

C: 20ns Memory, 15ns CPU, 8ns I/O

مقدار CPU Utilization را در صورتی که یک سیستم چند پردازنده و دو هسته‌ای داشته باشیم که از multiprogramming پشتیبانی می‌کند حساب کنید.

توجه: فرض کنید دستگاه‌های I/O دسترسی به حافظه به صورت موازی عملیات‌های پردازنده‌ها را انجام می‌دهند.



کل زمان: $10 + 35 + 20 + 50 + 15 = 130$ نانو ثانیه

بهره‌وری هسته اول: 0.307 یا 30.7 درصد | بهره‌وری هسته دوم: 0.269 یا 26.9 درصد

