آریا خلیق ۹۵۲۴۰۱۴ تمرین سری دوم درس سیستمهای عامل در بسیاری از سیستمها فرایندها دایما در حال تغییر هستند به همین دلیل یک روند ثابت برای اجرای این فرایندها وجود ندارد به همین دلیل فقط میتوان تقسیم زمانی CPU را مورد بحث قرار داد. به خاطر همین ما از تخمین استفاده میکنیم و یک تابع توزیع احتمال به دست میاوریم. این تابع توزیع نمایی میباشد.

یک کامپیوتر را میتوان شبکهای از سرورها دانست که هر سرور یک صف از فرایندهای در صف دارد. CPU هم یک سرور میباشد که به فرایندها سرویس میدهد.

فرض کنیم یک صف به طول متوسط n داریم و W هم متوسط زمان انتظار فرایند در صف باشد و L متوسط نرخ رسیدن و وارد شدن فرایندهای جدید به صف باشد. اگر صف در حالت پایدار باشد باید تعداد قرایندهایی که وارد میشوند برابر تعدادی باشد که خارج میشوند که بنابراین باید رابطه زیر برقرار باشد:

n = L * W

این فرمول، فرمول Little نام دارد و همچنین برای هر الگوریتم زمانبندی و رسیدگی به ورودی ها صادق است. با این فرایند میتوان هر یک را با دوتای دیگر حساب کرد و تخمین زد.

مشکل زمانبندی زمانی رخ میدهد که یک فرایند با اولویت بالاتر میخواهد اطلاعات کرنل هم را که الان فرایند با اولویت پایینتر به آن دسترسی دارد را بخواند یا بنویسد. اطلاعات کرنل هم توسط Lock در مقابل یک فرایند دیگر نگه داری میشوند به همین دلیل اگر فرایند با اولویت بالا بخواهد به آن دسترسی پیدا کند باید منتظر پایان فعالیت فرایند کماولویتتر باشد. این مشکل در سیستمهای با تعداد اولویت بیشتر از ۲ رخ میدهد.

فرض کنیم سه فرایند ۱، ۱ و ۳ را داریم. (۱ بیشترین اهمیت، ۲ متوسط و ۳ کمترین) در ابتدا فقط فرایند ۱ در حال اجرا است که وارد CS میشود، حال فرایند ۲ شروع شده و CPU را در دست میگیرد. حال فرایند ۱ شروع به کار میکند و CPU را از فرایند ۲ میگیرد. فرایند ۱ میخواهد وارد CS خود شود ولی چون فرایند ۳ داخل CS است بلاک میشود. CPU به فرایند ۲ داده میشود و تا زمانی که اجرای آن تمام نشده و CPU به فرایند ۳ داده نشده که از CS درآید در همین حالت میماند.

برای رفع مشکل برای اولویتهای بیشتر از ۲ تا از priority-inheritance protocol استفاده میشود که اولویت فرایند کمتر که داخل CS است به بالاترین سطح یا سطح اولویت بالاتر میرسد و اینگونه از شر اولویتهای وسط در امان میماند.

سیستمهای بلادرنگ سیستمهایی هستند که زمان در آنها نقش بسیار مهمی دارد. این سیستمها به دود دسته hard real time و soft real time تقسیم میشوند.

در hard real time باید سر زمان مشخص **حتماً** پاسخ و نتیجه را دریافت کنیم و این زمان به هیچ عنوان نباید رد شود ولی در حالت soft ممکن بسیار سعی میشود رد نشود ولی ممکن است کمی تفاوت داشته باشد.

وقایعی که سیستمهای بلادرنگ به آن باید پاسخ دهند به دو بخش periodic و periodic و تقسیم میشوند که در مواقعی سیستم باید به تعدادی وقایع periodic پاسخ دهد. ممکن است تعداد وقایع و زمان بندی آنها به گونهای باشد که پاسخگویی به همه آنها غیر ممکن باشد.

فرض کنیم واقعه ۱ دارای پریود P1 و زمان C1 برای پاسخ به آن نیاز باشد. باید جمع تمامی Ci/Pi ها کمتر از ۱ باشد که سیستم بتواند به آنها پاسخ دهد و اگر بیشتر باشد غیر ممکن overhead) پردازنده بسیار کوچک در نظر گرفته شده)

الگوریتمهای سیستمهای بلادرنگ میتواند static یا dynamic باشد. در حالت استاتیک تصمیمات باید قبل از شروع کار سیستم گرفته شوند و در حالت داینامیک سیستم در حال اجرا تصمیمات را میگیرد. حالت استاتیک زمانی ممکن است که اطلاعات خیلی خوبی در مـورد فرایندها قبل از شروع داشته باشیم.

۳ دسته بندی کلی داریم: Batch – Interactive – Real time

به طور کلی یک الگوریتم زمانبندی باید دارای شرایط زیر باشد:

- دادن عادلانه CPU به هر فرایند => باید به فرایندها بر اساس نیازشان، اهمیتشان و ... CPU اختصاص داد.
- قسمتها را به صورت بالانس فعال و مشغول نگه داشتن => مثلاً اینگونه نباشد که اول فعالیتهای در ارتباط با CPU اجرا و بعد فرایندهای مربوط به I/O Device اجرا شوند و این دو باید همزمان با هم اجرا شوند.
- باید اطمینان داشت که قوانین به درستی اجرا میشوند => مثلاً یک فرایند برای نباید برای مدت زیادی داخل CS باشد و دایم باید برای درست اجرا شدن قانون چک شود.

در Batch سیستمها:

- زمان Turnaround Time حداقل
- CPU باید همیشه مشغول باشد => CPU گران است و بهتر است همواره رد حال اجرای کاری باشد.
 - Throughput (تعداد کارهایی که در یک ثانیه کامل میشوند) باید حداکثر باشد نکته: لزوماً افزایش turnaround time نکته: لزوماً افزایش

در Interactive سیستمها:

- Response time سریع => برای کاربر بسیار مهم است.
- برآورده کردن انتظارات کاربران => مثلاً اتصال به مودم به زمان ۴۰ ثانیه نیاز دارد و این برای کاربر پذیرفته شده است ولی قطع کردن آن نه.

در Realtime سیستمها:

- اجرای هیچ فرایندی نباید ددلاین را رد شود => ددلاینها بسیار مهم هستند.
- جلوگیری از افت کیفیت در سیستمهای چند رسانهای => مثلاً در سیستمهایی که ارتباط با صدا دارند لگ زیاد پذیرفته شده نیست.

در سیستمعامل زمانی که یک فرایند fork میشود یعنی یک کپی از خود ایجاد میکند. در system call سیستمعامل UNIX و سیستمعاملهای مشابه (GNU/Linux) این کار از طریق tork و میشود. فرایند ایجاد شده را فرایند فرزند و فرایند اصلی که فرایند از آن fork شده را والد مینامند.

فرایند والد اجرای خود را ادامه میدهد در حالی که فرایند فرزند شروع به اجرا از همان جایی که fork شده میکند.

دو فرایند جدا یک برنامه را اجرا میکنند ولی stack، data و heap آنها متفاوت از هم است. در زمانی که تابع fork فراخوانی شده بود قسمتهای stack، data و الد و فرزند مشابه هم هستند ولی پس از آن متفاوت خواهند بود به همین دلیل هر فرایند که دادههای خود را تغییر دهد روی دیگر تأثیری نمیگذارد.

فراخوانی fork که در کتابخانه unistd.h تعریف شده است پس از فراخوانی سه مقدار باز میگرداند.

- If fork() returns negative value; creation of child process was unsuccessful.
- If fork() returns a zero; newly child process is created.
- If fork() returns positive; process _ID of child process to the parrent. It is type of pid_t defined as in sys/types.h

زمانبندی در لینوکس:

non-real-time به قبل لینوکس از الگوریتم زمانبندی real-time به همراه استفاده میکرد. از لینوکس ۲.۶ به بعد الگوریتم non-real-time در لینوکس بازنگری شد.

الگوریتم زمانبندی Real-Time:

سه کلاس زمانبندی در لینوکس عبارتند از:

- SCHED_FIFO => FIFO real-time thread
- SCHED_RR => Round-Robin real-time thread
- SCHED_NORMAL => Other, Non-Real_Time thread

داخل هر کلاس میتوان از انواع اولویتها استفاده کرد. هر یک از این مقادیر حالت پیشفرض دارند.

برای نخها FIFO قوانین زیر اعمال میشود:

- ۱) سیستم به این رشته اینتراپت نمیدهد جز در مواقع زیر:
 - یک نخ دیگر با اولویت بالاتر در حالت ready باشد.
- اجرای نخ بلاک شود برای دریافت یک event
 - خود نخ به صورت داوطلبانه CPU را پس بدهد.
- ۲) زمانی که در حین اجرای نخ اینتراپت داده میشود در جای مناسب خود با توجه به اولویتش قرار میگیرد.
 - ۳) اگر نخ در حالت ready قرار میگیرد و اولویت بالاتری از نخی که در حال اجرا است دارد CPU به نخ با اولویت بالاتر داده میشود.

کلاس SCHED_RR بسیار شبیه به کلاس SCHED_FIFO میباشد ولی time slice بیشتری در در هنگام انتظار به آن میرسد. زمانی که نخ SCHED_RR برای یک time slice اجرا میشود آن نخ suspend میشود و یک نخ با اولویت برابر یا بیشتر برای اجرا انتخاب میشود.

زمانبندی در ویندوز:

هدف ویندوز ارایه تجربه کاربری عالی به کاربران در هر دو حالت دسکتاپ و سرور میباشد. ویندوز یک زمانبند قابل پس گیری را با حالتها و درجههای مختلف قابل پس گیری بودن پیادهسازی میکند. در ویندوز نخها واحدهای زمانبندی هستند(به جای فرایندها)

اولویتها در ویندوز به دو دسته بلادرنگ و متغیر تقسیم میشوند و هر کدام هم دارای ۱۶ real- اولویت مختلف هستند. نخهایی که باید به صورت بلادرنگ رسیدگی شوند در دسته بندی -time قرار میگیرند و به صورت بلادرنگ اجرا میشوند. ویندوز از یک سیستم دارای اولویت استفاده میکند که نخهای در گروه بلادرنگ دارای اولویت بالاتر نسبت به بقیه هستند. در کلاس نخهای بلادرنگ اولویتها ثابت هستند و در این کلاس به صورت Round-Robin سرویس داده میشود.

این الگوریتم تقریباً سعی میکند به صورت یکسان به فرایندهایی که منتظر CPU هستند سرویس دهیم و زمانی که یک فرایند وارد شد با آن مثل زمانی رفتار میشود که انگار قبلاً وجود داشته. به عبارتی الگوریتم به گذشته نزدیک خود نگاه میکند و سعی میکند به فرایندها به گونهای سرویس دهد که به صورت مساوی CPU دریافت کنند.

در این حالت بین فرایندها با اولین بالاتر و عطش برای CPU و فرایندهای دیگر تفاوتی قایل نمیشویم و هر دو را به صورت برابر میبینیم چون ممکن است فرایند با عطش CPU در ابتدا انجام شود ولی بعد از آن همان زمان را به فرایندی با اولویت کمتر خواهیم داد. در این حالت عدالت رعایت نمیشود. برای فرایندها با عطش I/O هم به همین شکل است.

دو نوع سیستم Real-Time داریم، اولی که Soft Real-Time است سیستم گارانتی نمیکند که عملیات سر ددلاین انجام شوند در عوض فعالیتهای بحرانی را به فعالیتای غیر بحرانی ترجیح میدهد. در Hard Real-Time نیازمندیهایمان بیشتر است و هر فعالیت و وظیفه باید در ددلاین خود انجام شود و رد شدن ددلاین به هیچ وجه مورد قبول نیست.

مقداری تأخیر با توجه به نوع کار سیستم و آن فعالیت قابل تحمل است، دو نوع تاخیر داریم:

- تأخیر وقفه => فاصله زمانی که وقفه به وجود میآید تا زمانی که CPU شروع به پردازش وقفه میکند.
- تأخیر Dispatch => مدت زمانی که نیاز است تا dispatcher یک فایند را متوقف و فرایند دیگر را شروع کند.

انواع زمانبندیها:

زمانبندی Priority-Based:

مهمترین کاری که الگوریتمهای زمانبندی در سیستمهای Real-Time باید انجام دهند، دادن هر چه سریعتر CPU به فرایندی است که اولویت بالاتری دارد و CPU میخواهد. در این نوع زمانبندی ما اولویتهایی برای انواع مختلف فرایندها داریم (در ویندوز ۳۲ نوع اولویت مختلف که از ۱۶ تا ۳۱ برای فرایندهای real-time است). این نوع زمانبندی Soft Real-Time را تضمین میکند.

زمانبندی Rate-Monotonic:

و این کار زمان انتظار را کاهش میدهد.

این الگوریتم از الگوریتم زمانبندی کلی مبتنی بر دادن اولویت به کارها پیروی میکند. در ای نوع الگوریتم هر فرایندی که وارد میشود بر اساس پریود تکرار آن یک اولویت (هر چقدر پریود آن کوتاهتر باشد اولویت آن بیشتر است) علت این کار دادن اولویت بالاتر به task هایی هست که در بازه زمانی کوتاهتری تکرار میشوند

زمانبندی Earliest-Deadline_First:

در این نوع زمانبندی اولویتبندی به صورت داینامیک با توجه به ددلاین فرایندها صورت میگیرد به این صورت که هر چه ددلاین کاری زودتر باشد اولویت آن بالاتر میشود و برعکس. بر طبق قوانین این الگوریتم زمانی هر فرایند باید ددلاین خود را به سیستم اعلام کند. در این نوع زمانبندی بر خلاف Rate-Monotonic دیگر اولویتها ثابت نیستند.

زمانبندی Proportional:

در این زمان بندی کلاً T سهم بین تمامی اپلیکیشنها تقسیم میشود. هر اپلیکیشن میتواند n سهم داشته باشد و در کل صاحب n/T زمان CPU خواهد بود.

این نوع زمانبندی باید همراه قانون admission-control باشد که مطمین شویم هر اپلیکیشن زمان مناسب و مشخص خود را دریافت میکند.

زمان بندی POSIX Real-Time;

POSIX یک استاندارد برای برنامهریزی سیستمهای بلادرنگ است. این سیستم دو کلاس برای threadها مشخص میکند(قبلا بررسی شده):

- SCHED_FIFO
- SCHED_RR

کلاس اول thread را بر اساس first come firs serve و صف FIFO اولویت بندی میکند و فرایند با اولویت بیشتر که در ابتدای صف است CPU را گرفته و تا زمانی که تمام شود یا بلاک شود آن را نگه میدارد. در حالت SCHED_RR به صورت round robin به نخهایی با اولویت برابر رسیدگی میشود.

تمرین امتیازی ۱)

۱) چرا باید اجازه اجرای همزمان برنامهها را داد؟

برنامهها میتوانند به صورت همزمان اجرا شوند، چندین برنامه که همزمان اجرا میشوند میتوانند با هم همکاری داشته باشند. از فواید اجازه اجرای همزمان دادن: فرض کنیم چند کاربر یا برنامه میخواهند با همزمان با هم به یک منبع دسترسی داشته باشند. باید این اجازه را داد . در سیستمهایی با چند هسته پردازنده میتوان یک برنامه را برای افزایش سرعت به تکههایی تقسیم کرد و اجرای آن را به صورت همزمان جلو برد. کاربر ممکن است همزمان چندین برنامه را با هم اجرا کند.

۲) فواید استفاده از نخ به جای ایجاد چند فرایند همزمان را در سرور شرح دهید.

در گذشته این اتفاق میافتاد که زمانی که یک سرور یک request دریافت میکرد، یک فرایند جدا برای سرویس دادن به آن request ایجاد میکرد. اینکه برای هر request یک فرایند بسازیم زمان زیادی میبرد و منابع زیادی هم مصرف میکند به جای این کار بهتر است از یک فرایند request استفاده کنیم که خود چند نخ دارد. در این حالت سرور یک نخ درست میکند که به request ها گوش میکند. زمانی که یک request از client آمد به جای ایجاد فرایند یک نخ برای پاسخ به آن request ایجاد میشود.

۳) راه حلی عملی برای رفع مشکل starvation در سیستمهای دارای اولویت ارایه کنید. در اینگونه سیستمها ممکن است فرایندهایی که اولویت کمی دارند همواره منتظر CPU بمانند ولی هیچگاه CPU دریافت نکنند و تا همیشه در انتظار CPU باشند. برای رفع این مشکل از Aging استفاده میشود به گونهای که فرایندی که مدت زیادی منتظر CPU مانده به تدریج اولویتش زیادتر میشود، اینگونه اگر برای مدت زمان زیادی CPU به آن نرسد اولویتش به گونهای افزایش پیدا میکند که CPU دریافت کند.

تمرین امتیازی ۲)

۱) چرا نرمافزارها نباید به زمانبندی خود سیستم اتکا کنند؟

فرض کنید میخواهیم یک ویدیو اجرا کنیم که همراه با خود یک صوت دارد. اگر به زمانبندی خود سیستم اتکا کنیم و برای مثال یک loop خالی با ۱۰۰ بار اجرا بگذاریم ممکن CPU در وسط از صدا گرفته شود و زمانبندی صدا و ویدیو به هم بخورد.

۲) پیادهسازی نخها در فضای کاربر چه مزیتی نسبت به کرنل دارد؟

اولین مزیت این است که کرنل ممکن است اصلاً threading را ساپورت نکند و این مشکل را می میتوان به کمک کتابخانه ها و پیادهسازی در فضای کاربر رفع کرد. مورد بعدی این است که میتوان الگوریتم زمانبندی را به صورت دلخواه پیادهسازی کرد (مثلاً garbage collector در یک حالت نامناسب متوقف نشود). به scalability هم کمک میکند چون table space های کاربران و کرنل جداست وجود تعداد بسیار زیادی thread در کرنل باعث ایجاد مشکل میشود.

۳) اصلیترین و مهمترین اهداف الگوریتمهای زمانبندی چیست؟

باید با در نظر گرفتن اولویت فرایندها به آنها CPU داده شود. مثلاً اینکه همیشه یک فرایند CPU داشته باشد یا یک برنامه بسیار پر اهمیت به میزان کمی CPU دریافت کند پذیرفته نیست. برنامه زمانبند باید قوانینی را که برای فرایندها وجود دارد چک کند و از اجرا شدن آنها مطمین شود. در مورد سوم زمانبند باید از اجرای همیشه و لازم CPU مطمین باشد زیرا CPU یک قطعه گران است و زمانی که کار داشته باشیم باید مشغول انجام دادن کارها باشد.

پاسخ تمرینات تشریحی:

ارده دولتي ه ۲۷:

 $A \rightarrow 5' \rightarrow 5'$ $E \rightarrow 5' \rightarrow 5' \rightarrow 5' \rightarrow 1$ $C \rightarrow 5' \rightarrow 5' \rightarrow 1$ $D \rightarrow 7' \rightarrow 1$

(ontext switch , a RR o TS -> To

Gantt chart:

A V.	B	F. C. 200	Vo va	A 400	I C	184 140 160 1VA
) 9 / 9	C.B.T	20%	T.T	T.w	
A	0	4.	1 4.	116	٨٥	_
B	6	۲0	40	٤٥	10	
C	0	C) a	/νω	100	1 701	
D	O	۳.	140	140	١٣٠	

11 6111 1211 @

و زمان چوچی بای مانه عامای فراند مای آن نواند کاه

11, C 2000) 11A

F	Pr
CPYS	CP44
NEX 4	Disk3
CPV2	CEN 3
DIKK3	Net3
CPUB	cp43
DISK2	Net 3
CPy/2	CPU3

PI Pr	PI PI	Pr Pr Pr	PI Pr	
1// 8	Pr P	- 10.1	0.	

اس زمان مرابر ۱۲ و ۱۷۹۲ و مقدار هر رفتنی ۲ ات.

Q					ى دولتى ، دا
Pr V	P, P, F P	Pr 1	P _E	Pal Pu	1 F-1
Projet XT	Pal Pa	l P	1 Pal	1	
Pany XXX XXX		GBT	T.T	W.T	1
P4-) 1/2	P	4	4	0	
У	Pr	٧	11	4	
ATT- YWA	Pr	17	40	701	
AVg W.T= TUO AVg T. T= 01/ FA	PE	۲.	019	r9	
Avg T. 1= allin	Pa	TU	91	40	
	Py	40	91	ил	

AT 11:10 21 0

		?	-		
_		CBT	Arrival	T:T	WT
_	P _i	4	b	14	٨
_	Pr	٣	١	٣	0
	Pe	۲	٣	٣	1
	Pį	٣	V	۳	0

$$P_{1} \rightarrow Y_{0}$$

$$P_{7} \rightarrow Y_{0}$$

$$P_{7} \rightarrow Y_{0}$$

$$P_{4} \rightarrow Y_{0}$$

15671510 211 D

Pai T = V Pr: 11+v = IN

	BST	AT	T.T	w.
Po	4	0	17	٧
Pı	٤	۲	14	17
Pr	7	٣	٣	١
Pr	1	^	٢	1

PO MI PO MI

YN+ NX t (CS)

ازایز دهری کا ۱۶ = سیاسی زمان سروی

RP, 2 = looms

هر من آید ما خایش آیر که هر سال به ۲۲ زمان نیزراند ، هیدی دلی کدرمان کا ۱۰۱۰ در این کدرمان کا ۱۰۱۰ در کا درمان کا ۱۰۱۰ درمان کا درما

Y.5 = 0/F

ا مان بواران حداً فل زمان W.T وازالله ريم عرف استان يان.

دون الاستاز ۱۵ ات ب CBT آن ار ام فرایندیگر ستان ب اول واری این.

1	COT	TT	
PI	4	91	_
Pr.	V	94-15	
44	a	140	
PE	^	りょし	

T.T = EX + 0+15+50 = 150+50 = 15/10=191 = 15/10