



اهداف تمرین

- آشنایی با پروتکل BGP
- آشنایی با BGP finite state machine
- آشنایی با BGP message packets
- آشنایی با BGP routing policies
- آشنایی با BGP security

۱. مقدمه

پروتکل دروازه‌ای مرزی^۱ یکی از پروتکل‌های مسیریابی استاندارد است که ارتباط بین سامانه‌های مستقل^۲ را برقرار می‌کند. مسیریابی توسط این پروتکل بر اساس سیاست‌های تعیین شده برای سامانه انجام می‌گیرد. این پروتکل می‌تواند برای مسیریابی درون یک AS نیز استفاده شود، اما در این تمرین تاکید بر روی ارتباط خارجی بین ASها می‌باشد. برای آشنایی بیشتر با این پروتکل می‌توانید به اسلایدهای درس مراجعه نمایید.

۲. برقراری ارتباط BGP

برای برقراری ارتباط BGP، هر مسیریاب از یک ماشین حالت متناهی متشکل از ۶ حالت استفاده می‌کند. بسته به اینکه پروتکل در کدام یک از این حالات باشد، اعمال متفاوتی را انجام می‌دهد و پیام‌های متفاوتی را برای همتای

* با سپاس از امیرپاشا قابوسی و سولماز سلیمی

^۱BGP: Border Gateway Protocol

^۲AS: Autonomous Systems

خود ارسال می‌کند.

۱. حالت پایه Idle می‌باشد. در این حالت BGP هیچ ارتباط ورودی را نمی‌پذیرد و در صورت دریافت فرمان start یک زمان‌سنج به نام ConnectRetryTimer را شروع می‌کند. سپس درخواست برقراری یک ارتباط TCP را برای همتای دیگر ارسال می‌کند و به حالت Connect وارد می‌شود. اگر در مرحله‌ی Idle مشکلی ایجاد شود پس از به پایان رسیدن زمان‌سنج، مجدداً برای برقراری ارتباط تلاش می‌شود. زمان بین هر تلاش مجدد به صورت نمایی افزایش می‌یابد.

۲. حالت دوم Connect است. در این حالت BGP منتظر کامل شدن ارتباط TCP بوده و همزمان منتظر درخواست برقراری ارتباطی که ممکن است از طرف همتای دیگر برقرار شود نیز می‌باشد. در صورت برقرار ارتباط، ConnectRetryTimer ریست شده و متوقف می‌شود. سپس پیام OPEN ارسال شده، زمان‌سنج HoldTimer شروع شده و وضعیت به OpenState تغییر می‌کند. اگر ارتباط به دلیل retransmission timeout برقرار نشود زمان‌سنج ریست شده و وضعیت به ActiveState تغییر می‌کند. اگر به هر دلیل دیگری برقراری ارتباط میسر نشود وضعیت به Idle تغییر می‌کند.

۳. در حالت ActiveState انتظار می‌رود ارتباط از سمت همتا برقرار شود. اگر این ارتباط به طور کامل برقرار شد ConnectRetryTimer ریست شده و متوقف می‌شود. پیام OPEN به همتا ارسال شده، زمان‌سنج HoldTimer شروع شده و در نهایت وضعیت به OpenState تغییر می‌کند. اگر چنین ارتباطی برقرار نشد و تایمر به اتمام رسید، تایمر ریست شده و درخواست برقراری ارتباط ارسال می‌شود. در نهایت وضعیت مجدداً به Connect تغییر می‌کند. اگر به هر دلیل دیگری برقراری ارتباط میسر نشود وضعیت به Idle تغییر می‌کند.

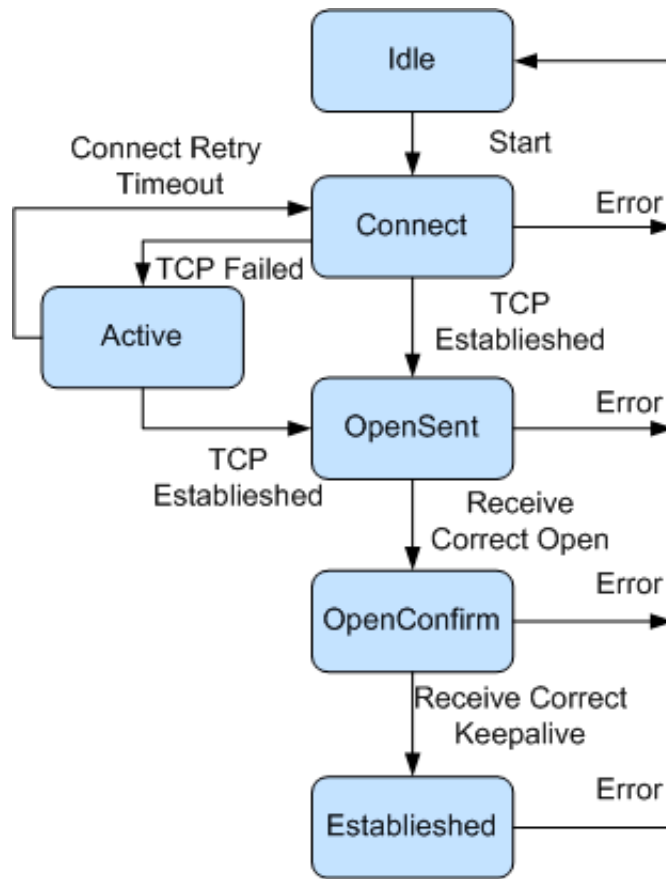
۴. در حالت OpenState انتظار می‌رود BGP منتظر دریافت پیام OPEN از طرف همتا باشد. پس از دریافت این پیام تمامی بخش‌های آن بررسی شده و در صورت وجود اشکال، پیامی از نوع NOTIFICATION به همتا ارسال می‌شود و وضعیت به Idle تغییر می‌کند. اگر هیچ اشکالی وجود نداشت پیام KEEPALIVE به همتا ارسال می‌شود، HoldTimer ریست شده و وضعیت به OpenConfirm تغییر می‌کند. اگر تایمر به پایان برسد و پیامی دریافت نشده باشد یک پیام NOTIFICATION با کد Hold Timer Expired به همتا ارسال شده و به حالت Idle می‌رویم. اگر به هر علت دیگری مشکلی در ارتباط به وجود آید، یک پیام NOTIFICATION با کد Finite State

Machine Error به همتا ارسال شده و وضعیت به Idle تغییر می‌کند.

۵. در حالت OpenConfirm بی‌جی‌پی منتظر پیام KEEPALIVE بوده و با دریافت آن به حالت Established تغییر می‌کند. اگر زمان‌سنج به اتمام برسد و KEEPALIVE دریافت نشده باشد، یک پیام NOTIFICATION با کد Hold Timer Expired به همتا ارسال شده و وضعیت به Idle تغییر می‌کند. اگر به هر علت دیگری مشکلی در ارتباط به وجود آید، یک پیام NOTIFICATION با کد Finite State Machine Error به همتا ارسال شده و وضعیت به Idle تغییر می‌کند.

۶. در حالت Established توانایی تبادل همه‌ی پیام‌های UPDATE ، NOTIFICATION و KEEPALIVE وجود دارد. اگر هر یک از این پیام‌ها دریافت شوند تایمر ریست می‌شود. در صورت دریافت پیام NOTIFICATION وضعیت به Idle تغییر می‌کند. اگر در پیام UPDATE دریافت شده خطایی وجود داشته باشد، ابتدا یک پیام NOTIFICATION به همتا ارسال می‌شود و سپس وضعیت به Idle تغییر می‌کند. اگر تایمر به اتمام رسید، یک پیام NOTIFICATION با کد Hold Timer Expired به همتا ارسال شده و وضعیت به Idle تغییر می‌کند. اگر به هر علت دیگری مشکلی در ارتباط به وجود آید، یک پیام NOTIFICATION با کد Finite State Machine Error به همتا ارسال شده و وضعیت به Idle تغییر می‌کند.

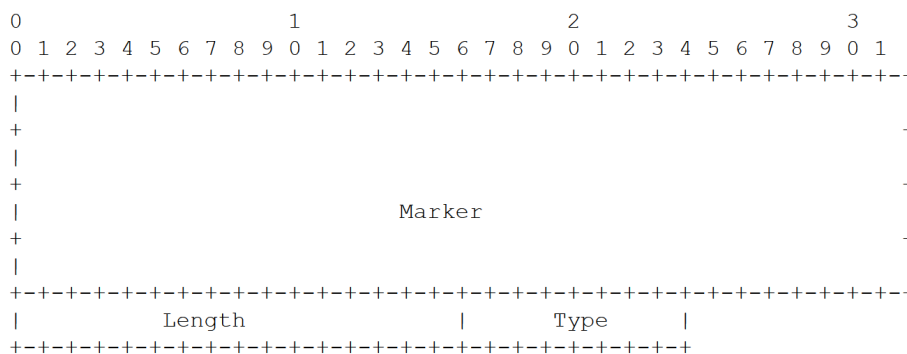
- اگر از هر حالتی به Idle تغییر وضعیتی اتفاق بیفتد، بلافاصله تقاضای برقراری ارتباط مجدداً ارسال شده و ConnectRetryTimer ریست می‌شود.
- مقدار اولیه‌ی ConnectRetryTimer معمولاً برابر ۶۰ ثانیه می‌باشد.
- مقدار HoldTimer معمولاً برابر ۲۴۰ ثانیه است.
- در تمامی حالات به جز حالت Idle فرمان start نادیده گرفته می‌شود.
- در حالت OpenConfirm یا Established اگر پس از گذشت زمانی معین پیام UPDATE یا KEEPALIVE فرستاده نشده باشد، یک پیام KEEPALIVE برای همتا ارسال می‌شود. این زمان معمولاً یک سوم HoldTimer است. در این تمرین از این مساله صرف نظر کنید.



شکل ۱: دیاگرام حالات در BGP، مرجع شکل

۳. انواع پیامها

تمامی بسته‌های BGP دارای سرآیندی در قالب زیر می‌باشند^۳:



شکل ۲: سرآیند بسته BGP

^۳مرجع تصاویر مربوط به توضیح سرآیند پیامها مستند (BGP-4) RFC 4271: A Border Gateway Protocol می‌باشد.

این سرآیند طبق قواعد زیر ساخته می‌شود:

- بخش Marker در سرآیند باید تماماً با بیت‌های ۱ پر می‌شود و اندازه‌ی آن ۱۲۸ بیت است.
- مقدار Type برای پیام OPEN ، NOTIFICATION ، KEEPALIVE و UPDATE به ترتیب برابر ۱، ۲، ۳، ۴ و ۱ است.
- مقدار Length برابر طول کل پیام به بایت است.

۱.۳ پیام OPEN

بسته‌های پیام OPEN به صورت زیر می‌باشند:

```
0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
|   Version   |
+-----+-----+-----+-----+
| My Autonomous System |
+-----+-----+-----+-----+
|           Hold Time           |
+-----+-----+-----+-----+
|                               BGP Identifier                               |
+-----+-----+-----+-----+
| Opt Parm Len |
+-----+-----+-----+-----+
|                               Optional Parameters (variable)                               |
+-----+-----+-----+-----+
```

شکل ۳: قالب پیام OPEN

این بسته طبق قوانین زیر ساخته می‌شود.

- بخش Version نشان دهنده‌ی نسخه‌ی مورد استفاده‌ی BGP بوده و در این تمرین برابر ۴ است.
- My Autonomous System شماره‌ی AS ارسال کننده‌ی بسته است.
- Hold Time مقدار اولیه‌ی HoldTimer ارسال کننده‌ی بسته می‌باشد. در این تمرین این بخش را همواره با ۰ پر کنید.
- BGP Identifier شماره‌ی IP روتر ارسال کننده‌ی بسته است.
- Opt Parm Len طول بخش اختیاری بسته به بایت است که در این تمرین ۰ در نظر گرفته می‌شود.

۲.۳ . NOTIFICATION پیام

این پیام تنها زمانی ارسال می‌شود که خطایی در سامانه رخ داده باشد و ارتباط BGP بلافاصله پس از ارسال این پیام بسته می‌شود.

0					1					2					3																										
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Error code										Error subcode										Data (variable)																					

شکل ۴: قالب پیام NOTIFICATION

این بسته طبق قوانین زیر ساخته می‌شود:

- قسمت Error code برای Error Message Header, OPEN Message Error, UPDATE Message Er- در این بخش به ترتیب برابر ۱، ۲، ۳، ۴، ۵ و ۶ مقدار می‌گیرد.
- بخش Error subcode و Data (variable) برای دادن اطلاعات بیشتر در مورد دلیل خطا می‌باشند که در این تمرین از آنها صرف نظر شده و لازم نیست بخش مربوط به آنها را در بسته‌ها لحاظ کنید.

۳.۳ . KEEPALIVE پیام

BGP از مکانیسم طراحی شده در TCP برای KEEPALIVE استفاده نمی‌کند و بسته‌ی مخصوص به خود را دارد. این بسته تنها شامل هدر BGP می‌باشد.

۴.۳ . UPDATE پیام

از این پیام برای تبلیغ یک مسیر قابل دسترس یا اعلام غیرقابل دسترس بودن یک مسیر استفاده می‌شود.

+-----+	
	Withdrawn Routes Length (2 octets)
+-----+	
	Withdrawn Routes (variable)
+-----+	
	Total Path Attribute Length (2 octets)
+-----+	
	Path Attributes (variable)
+-----+	
	Network Layer Reachability Information (variable)
+-----+	

شکل ۵: قالب پیام UPDATE

این بسته طبق قوانین زیر ساخته می‌شود (در این تمرین این بخش ساده شده است):

- Withdrawn Routes Length تعداد پریفیکس‌های موجود در Withdrawn Routes را مشخص می‌کند. طول این بخش همواره مضربی از ۵ است و ابتدا ۴ بایت آن به آیبی و ۱ بایت بعدی به سابنت ماسک یک پریفیکس اختصاص می‌یابد (سابنت ماسک عددی بین ۰ تا ۳۲ است). پیش‌وندهای مختلفی که در این بخش قرار می‌گیرند پشت سر قرار می‌گیرند.
- قسمت Total Path Attribute Length برابر تعداد مسیرهای موجود در Path Attributes می‌باشد.
- هر Path Attribute یک متغیر دو بخشی به صورت <attribute length, attribute value> است.

— attribute length شامل دو بایت بوده و نشان دهنده‌ی تعداد AS‌های موجود در attribute value است.

— attribute value برای تبلیغ یک AS_PATH استفاده می‌شود. بدین صورت که شماره‌ی AS‌هایی که در این مسیر وجود دارند به ترتیب در این بخش قرار می‌گیرند. شماره هر AS در دو بایت نوشته می‌شود. AS صاحب پیش‌وند به عنوان آخرین (کم ارزش‌ترین) دو بایت آخر قرار می‌گیرد.

- Network Layer Reachability Information طولی به مضرب ۵ دارد و مانند بخش Withdrawn Routes شامل پیش‌وند مسیرهای تبلیغ شده (به ترتیب قرارگیری در بخش قبل) می‌باشد.

پیام‌های یاد شده در بالا بسیار ساده شده‌اند و درباره‌ی بسیاری از حالات آن‌ها صحبتی نشده است. در صورت تمایل به مطالعه‌ی بیشتری می‌توانید به RFC 4271 مراجعه نمایید.

۴. قواعد انتخاب مسیر

رابطه‌ی بین دو AS می‌تواند یکی از دو حالت مشتری-تامین کننده یا همتا-همتا باشد. البته روابط دیگری نیز بین دو AS می‌تواند وجود داشته باشد، ولی در اینجا به بررسی همین دو حالت بسنده می‌کنیم. در صورتی که تمایل به ترانزیت نداشته باشیم، مسیرهایی که از یک مسیر همتا-همتا به ما تبلیغ شده باشند را به سایر مسیرهای همتا-همتا تبلیغ نمی‌کنیم ولی در روابط مشتری-تامین کننده همواره تمامی مسیرها را به مشتری خود تبلیغ می‌کنیم. همچنین یک مشتری تمامی مسیرها را به تامین کننده‌ی خود تبلیغ می‌کند به جز مسیرهایی که توسط سایر تامین کننده‌هایش به او تبلیغ شده باشند. در این تمرین ترانزیت غیرفعال است. برای انتخاب از بین مسیرهای مختلف منتهی به یک پیش‌وند به این ترتیب عمل می‌کنیم. ابتدا مسیری را انتخاب می‌کنیم که دارای بیشترین اولویت باشد. اولویت یک عدد صحیح مثبت است که برای هر مسیر به صورت دلخواه تعیین می‌شود. در صورت یکسان بودن اولویت چند مسیر مختلف، کوتاه‌ترین مسیر انتخاب

می‌شود. در صورت یکسان بودن طول مسیرها، مسیری انتخاب می‌شود که به واسطه با شماره‌ی کمتر متصل است. اگر این شماره نیز برای دو مسیر یکسان بود به شماره‌ی AS بعد از این همسایه نگاه می‌کنیم.

۵. مسائل امنیتی

در BGP ممکن است برخی مشکلات امنیتی رخ دهد. مثلاً ممکن است یک AS پیش‌وندی را تبلیغ کند که متعلق به خودش نیست. همچنین ممکن است یک AS اقدام به کوتاه‌تر، یا بلندتر نشان دادن یک مسیر نماید یا اینکه یک AS خاص را از مسیری حذف یا به آن اضافه کند. در این تمرین از شما خواسته می‌شود تا در صورت بروز برخی از این مشکلات آن‌ها را شناسایی کنید.

۶. پیاده‌سازی تمرین

شروع برقرار ارتباط با دستور `<I> start connection on interface` آغاز می‌گردد.

در تمامی مراحل برقراری ارتباط اگر از یک حالت به حالت دیگر، گذاری انجام شود باید خروجی زیر چاپ شود:

```
state changed from <X> to <Y> on interface <I>
```

در صورت برخورد به خطا باید پیام زیر چاپ شود:

```
an error occurred. state changed from <X> to <Y> on interface <I>
```

تنها بخش‌های خواسته شده‌ی سرآیند TCP را پر کنید. توجه کنید که از TCP تنها برای برقرار ارتباط اولیه استفاده می‌شود و پس از برقراری این ارتباط فقط از سرآیند BGP استفاده خواهد شد.

تنها بخش‌های Data offset, Flags را مقداردهی کنید و سایر بخش‌ها را ۰ قرار دهید.

برای سرآیند IP تنها بخش‌های TTL, Src IP, Dst IP, Length, IHL, Verion را مقداردهی کنید و سایر بخش‌ها را خالی بگذارید. در سرآیند Ethernet نیز EtherType را برابر 0x0800 و پورت‌ها را Broadcast قرار دهید.

TCP Header																																	
Offsets	Octet	0				1				2				3																			
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Source port								Destination port																							
4	32	Sequence number																															
8	64	Acknowledgment number (if ACK set)																															
12	96	Data offset	Reserved 0 0 0	N S	C R	E G	U K	A H	P T	R N	S N	F N	Window Size																				
16	128	Checksum								Urgent pointer (if URG set)																							
20	160	Options (if data offset > 5. Padded at the end with "0" bytes if necessary.)																															
...																															

شکل ۶: سرآیند TCP، مرجع شکل

IPv4 Header Format																																	
Offsets	Octet	0				1				2				3																			
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version		IHL		DSCP		ECN		Total Length																							
4	32	Identification								Flags		Fragment Offset																					
8	64	Time To Live				Protocol				Header Checksum																							
12	96	Source IP Address																															
16	128	Destination IP Address																															
20	160	Options (if IHL > 5)																															
24	192																																
28	224																																
32	256																																

شکل ۷: سرآیند IP، مرجع شکل

در این تمرین لازم نیست زمان بین برقراری مجدد ارتباط را به صورت نمایی اضافه کنید. برای ارسال KEEPALIVE در صورتی که هیچ بسته‌ای به مدت ۶۰ ثانیه ارسال نشده بود اقدام کنید. همچنین همواره باید تمامی بسته‌های دریافتی را بررسی کنید و در صورتی که مشکلی در آن‌ها بود با توجه به جدول آخر تمرین پیام NOTIFICATION مناسب را ارسال کنید.

پس از برقراری ارتباط در صورتی که دستور advertise all دریافت کردید باید تمامی مسیرهایی که تاکنون یاد گرفته‌اید به همراه تمامی پیشوندهای خود را به همسایه‌های خود در صورت برقراری ارتباط، تبلیغ کنید. ترتیب قرارگیری مسیرها در بسته مهم نیست. در این بخش باید تمامی قوانین انتخاب مسیر (مانند ترانزیت و...) را رعایت کنید.

مسیر یاد گرفته شده از یک همسایه، در صورتی که شماره‌ی AS خودتان در آن باشد دیگر تبلیغ نمی‌شود.

در صورت دریافت دستور <PREFIX> print routes to تمامی مسیره‌ای یاد گرفته شده به این پیش‌وند را به ترتیب اولویت انتخاب چاپ کنید. در صورتی که هیچ مسیری به این پیش‌وند وجود نداشته باشد خروجی زیر چاپ می‌شود:

no routes found for <PREFIX>

در صورت دریافت دستور <I> is <X> priority of اولویت مسیرهایی که شروع آن‌ها از این واسط است را تغییر دهید. به صورت پیش فرض اولویت مسیرهایی که از مشتری‌های خود یاد گرفته‌اید برابر ۱۰۰، مسیرهایی که به صورت همتا-همتا هستند برابر ۹۰ و مسیرهایی که از تامین‌کننده‌ی خود یاد گرفته‌اید برابر ۸۰ است.

در صورت دریافت دستور <PREFIX> hijack شروع به تبلیغ این پیش‌وند به تمامی همسایه‌های خود کنید. هنگامی که یک AS متوجه این اتفاق می‌شود این مسیر را تبلیغ نمی‌کند و باید خروجی زیر را چاپ کند:

<PREFIX> is hijacked!

شناسایی این اتفاق بدین شکل صورت می‌گیرد که اگر از قبل یک AS را به عنوان صاحب این پیش‌وند بدانیم و ناگهان یک AS دیگر ادعای مالکیت آن را کند، این یک hijacking محسوب می‌شود.

در صورت دریافت دستور <PREFIX> withdraw این پریفیکس را پاک کرده و پیام UPDATE مربوط به این مساله را به تمامی همسایه‌ها ارسال کنید. ابتدا از اینترفیس شماره صفر خود شروع کنید. همسایه‌ها نیز به محض دریافت این پیام، مسیری را که از ما به این پریفیکس می‌رسیده را پاک کرده و این مساله را به همسایه‌های دیگر خود اعلام می‌کنند.

برای کاهش زمان داوری تمارین، مدت زمان اولیه‌ی ConnectRetryTimer را ۳۰ ثانیه در نظر بگیرید. تمامی اطلاعات لازم به عنوان ورودی در AS_information به شما داده می‌شوند.

Value	Name
0	Reserved
1	Message Header Error
2	OPEN Message Error
3	UPDATE Message Error
4	Hold Timer Expired
5	Finite State Machine Error
6	Cease
7	ROUTE-REFRESH Message Error
8-255	Unassigned

شکل ۸: کدهای خطا در Notification، مرجع شکل

۷. نقشه‌ی تمرین و آزمون‌های نمره‌دهی

در تصویر صفحه‌ی بعد می‌توانید نقشه‌ای که در اختیار شما قرار داده شده تا کدهای خود را تست کنید را مشاهده فرمایید.

همچنین سناریوهای نمره‌دهی و تست این تمرین شامل سه سناریوی زیر می‌باشد:

- برقراری ارتباط: در این بخش تنها برقراری ارتباط بین AS ها و درستی آنها سنجیده شده و سایر قابلیت‌های کد شما تاثیری در نمره‌ی این بخش ندارد. نمره‌ی این سناریو ۴۰ درصد نمره‌ی تمرین است.
- تبلیغ مسیرهای جدید: در این بخش تنها درستی ارسال مسیرهای جدید به سایر AS ها مورد ارزیاب قرار گرفته و سایر توانایی‌های کد شما تاثیری در نمره‌ی این بخش ندارد. نمره‌ی این سناریو ۲۰ درصد نمره‌ی تمرین است.
- حذف مسیرهای withdraw شده: این بخش از تست‌ها فقط به بررسی درست عمل کردن دستور withdraw می‌پردازد. برای گرفتن نمره‌ی این بخش باید نمره‌ی قسمت قبل را کامل گرفته باشید. نمره‌ی این سناریو ۲۰ درصد نمره‌ی تمرین است.
- یافتن مسیرهای جعلی: این بخش تنها به بررسی کارکرد درست دستور hijack پرداخته و برای گرفتن نمره‌ی آن باید بخش ارسال مسیرهای جدید به درستی کار کند. نمره‌ی این سناریو ۲۰ درصد نمره‌ی تمرین است.



شکل ۹: نقشه

نکات ضروری

- به علت اینکه نمره‌ی تمرین به صورت خودکار داده می‌شود، ساختار پیام‌های گفته شده باید دقیقاً به صورت گفته شده باشد.
- نقشه‌ای که برای ارزیابی استفاده می‌شود با نقشه تست که در اختیار شما قرار گرفته فرق می‌کند.
- داوری خودکار بصورت برخط پس از اتمام مهلت ارسال "مستند طراحی" فعال می‌شود.
- به دلیل مشکلات اینترنتی بهتر است داوری را هنگامی که به اینترنت دانشگاه متصل هستید انجام دهید.
- در صورتی که هر مشکل یا پرسشی داشتید که فکر می‌کنید پاسخ آن برای همه مفید خواهد بود، آن را به گروه اینترنتی درس ارسال کنید.
- از فرستادن جواب تمرین به گروه اینترنتی درس خودداری کنید.
- تمام برنامه‌ی شما باید توسط خود شما نوشته شده باشد. فرستادن کل یا قسمتی از برنامه‌تان برای افراد دیگر، یا استفاده از کل یا قسمتی از برنامه‌ی فرد دیگری، حتی با ذکر منبع، تقلب محسوب می‌شود.
- پس از اتمام کارتان لازم است با اجرای دستور `make archive` فایل زیبایی شامل تمام فایل‌هایی که برای اجرا شدن کد شما نیاز است بسازید. در صورتی که از کلاس‌ها و فایل‌های اضافه شده خودتان استفاده می‌کنید، سعی کنید در پوشه گفته شده باشد. در هر صورت فایل آرشیو شما باید قابلیت کامپایل/اجرا شدن را به روش سیستمی داشته باشد، در غیر اینصورت نمره شما صفر خواهد شد.
- در این تمرین Judge فقط ابزار کمکی است و هر گونه خرابی در این سیستم تاثیری بر روی زمان تحویل ندارد.
- نسخه نهایی تمرین را در مخزن خود در [وب سایت ترشت](#) بارگذاری نمایید.