به نام خداوند بخشنده مهربان

دانشگاه صنعتی شریف
مدرس: مهدی خرازی

تمرین برنامه‌نویسی سوم*
شبکه‌های رایانه‌ای

دانشکده‌ی مهندسی کامپیوتر
بهار ۱۳۹۸

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[1] یکی از پروتکل‌های مسیریابی استاندارد است که ارتباط بین سامانه‌های مستقل[2] 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. این تمرین حالت تغییر یافته‌ی پروتکل BGP است و بسیاری از پیچیدگی‌های آن برای شما ساده شده‌است.

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، سولماز سلیمی و کیانوش عباسی

[1] BGP: Border Gateway Protocol
[2] AS: Autonomous Systems

خود ارسال می‌کند. دقت کنید که هر ماشین برای هر کدام از واسط‌های خود یک ماشین حالت مستقل دارد. ماشین حالت متناظر با یک واسط دلخواه از حالات زیر تشکیل شده‌است:

۱. حالت پایه Idle می‌باشد. در این حالت BGP هیچ ارتباط ورودی را بر روی این واسط نمی‌پذیرد و در صورت دریافت فرمان start یک زمان‌سنج به نام ConnectRetryTimer را از ۳۱ ثانیه شروع می‌کند. سپس درخواست برقراری یک ارتباط TCP (یعنی یک پیام Syn را) برای همتای دیگر (آن سمت واسط) ارسال می‌کند و به حالت Connect وارد می‌شود.

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؛ به بیان دیگر در این‌جا سه پیام مختلف می‌تواند دریافت کند:

- اگر پیام SynAck دریافت کند یعنی پاسخ پیام Syn که قبلا فرستاده دریافت شده‌است و ارتباط برقرار شده‌است و در جواب یک پیام Ack می‌فرستد.
- اگر پیام Syn دریافت کند یعنی همتایش درخواست برقراری ارتباط داده و در پاسخ یک پیام SynAck به این همتا می‌فرستد.
- اگر پیام Ack دریافت کند یعنی پاسخ پیام SynAck که قبلاً فرستاده را دریافت کرده و و ارتباط برقرار شده‌است. (دقت کنید که ممکن است قبلاً پیام SynAck نفرستاده باشد و آنگاه شما باید این پیام دریافتی را نادیده بگیرید)

در صورتِ برقراری ارتباط، ConnectRetryTimer ریست شده و متوقف می‌شود. سپس پیام OPEN ارسال شده، زمان‌سنج دیگری به نام HoldTimer از ۲۴۰ ثانیه شروع شده و وضعیت به OpenState تغییر می‌کند.

اگر ارتباط به دلیل retransmission timeout پروتکل TCP برقرار نشود (یعنی در مدت معین پیام‌های SynAck و Ack لازم رد و بدل نشوند) زمان‌سنج ConnectRetryTimer ریست شده و وضعیت به ActiveStateتغییر می‌کند. مقدار retransmission timeout را برابر با ۳۰ ثانیه در نظر بگیرید.
اگر برقراری ارتباط به دلیل پایان یافتن ConnectRetryTimer میسر نشود، وضعیت به Idle تغییر می‌کند.

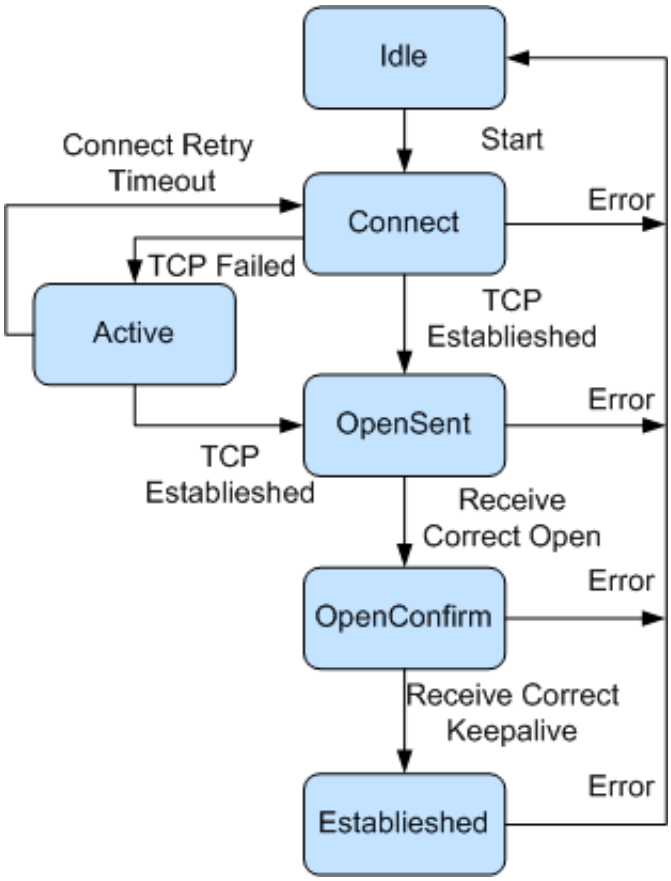
۳. در حالت ActiveState انتظار می‌رود ارتباط از سمت همتا برقرار شود؛ در واقع این حالت مانند یک وقت اضافه است که به همتا داده می‌شود تا ارتباطی که در حالت Connect تکمیل نشده را تکمیل کند. بنابراین پیام‌های دریافتی و پاسخ‌های متناظر مطابق بخش قبل است. اگر این ارتباط به طور کامل برقرار شد ConnectRetryTimer ریست شده و متوقف می‌شود، پیام OPEN به همتا ارسال شده، زمان‌سنج HoldTimerشروع شده و در نهایت وضعیت به OpenState تغییر می‌کند.
اگر چنین ارتباطی برقرار نشد و ConnectRetryTimer به اتمام رسید، زمان‌سنج ریست شده و مجددا درخواست برقراری ارتباط(پیام Syn) ارسال می‌شود. در نهایت وضعیت مجدداً به Connect تغییر

می‌کند. دقت کنید که در این حالت دیگر retransmission timeout مطرح نیست و صرفاً باید ConnectRetryTimerرا در نظر بگیرید.

۴. در حالت OpenState انتظار می‌رود BGP منتظر دریافت پیام OPEN از طرف همتا باشد. پس از دریافت این پیام، پیام KEEPALIVE به همتا ارسال می‌شود، HoldTimer ریست شده و وضعیت به OpenConfirmتغییر می‌کند.
اگر HoldTimer به پایان برسد و پیامی دریافت نشده‌باشد به حالت Idle می‌رویم.

۵. در حالت OpenConfirm بی‌جی‌پی منتظر پیام KEEPALIVE بوده و با دریافت آن ضمن ریست کردن HoldTimer، به حالت Established تغییر می‌کند. اگر زمان‌سنج به اتمام برسد و KEEPALIVE دریافت نشده باشد،وضعیت به Idle تغییر می‌کند.

۶. در حالت Established ارتباط بین دو همتا برقرار شده است، قابلیت دریافت پیام‌های KEEPALIVE و UPDATEوجود دارد. اگر HoldTimer به اتمام رسید و پیامی دریافت نشده‌بود، وضعیت به Idle تغییر می‌کند. با دریافت هرکدام از این پیام‌ها این زمان‌سنج باید ریست شود.



شکل ۱: دیاگرام حالات در BGP، مرجع شکل

- در تمامی حالات به جز حالت Idle فرمان start نادیده گرفته می‌شود.
- شرط لازم برای برقراری ارتباط بین دو همتا این است که هر دو فرمان start را بر روی واسطهای متناظر دریافت کنند. بنابراین واسطی که زودتر این پیام را دریافت می‌کند، یک پیام Syn به همتای خود می‌فرستد که با توجه به این‌که همتایش احتمالاً در حالت Idle است، این پیام نادیده گرفته می‌شود.
- اگر از هر حالتی به Idle تغییر وضعیتی اتفاق بیفتد، بلافاصله تقاضای برقراری ارتباط مجددا ارسال شده و ConnectRetryTimerریست می‌شود.
- در حالت OpenConfirm یا Established اگر پس از گذشت زمانی معین پیام UPDATE یا KEEPALIVEفرستاده نشده باشد، یک پیام KEEPALIVE برای همتا ارسال می‌شود. این زمان معمولا یک سوم HoldTimer است. در این تمرین از این مساله صرف نظر کنید.

۳. انواع پیام‌ها

در سرآیند TCP تنها قسمت Flags را مقداردهی کنید و بقیه‌ی بیت‌ها را برابر صفر بگذارید. برای سرآیند IP تنها بخش‌های Version, IHL, Length, Src IP, Dst IP, TTL را مقداردهی کنید و سایر بخش‌ها را صفر 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				NS	CWR	ECE	URG	ACK	PSH	RST	SYN	FIN	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.) ...																															

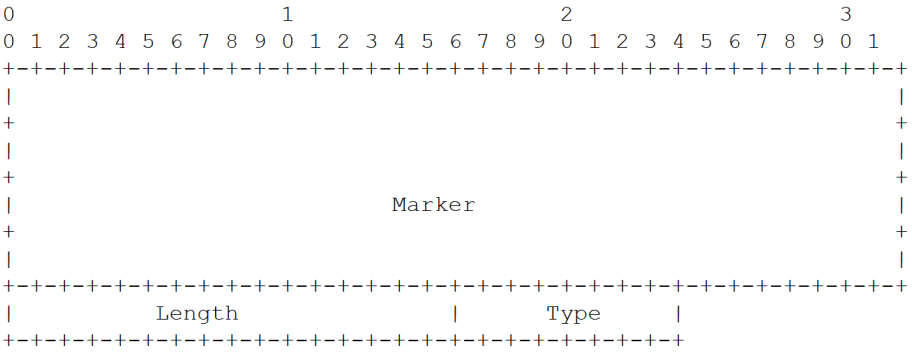
شکل ۲: سرآیند TCP، مرجع شکل

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 ۳: سرآیند IP، مرجع شکل

تمامی بسته‌های ارسالی به جز بسته‌های Syn, Ack, SynAck (که TCP هستند) از نوع BGP هستند. بسته‌های BGP مانند بسته‌های TCP از زیرمجموعه‌های بسته‌های IP هستند و دارای سرآیندی در قالب زیر می‌باشند[۳]



شکل ۴: سرآیند بسته BGP مرجع شکل

این سرآیند طبق قواعد زیر ساخته می‌شود:

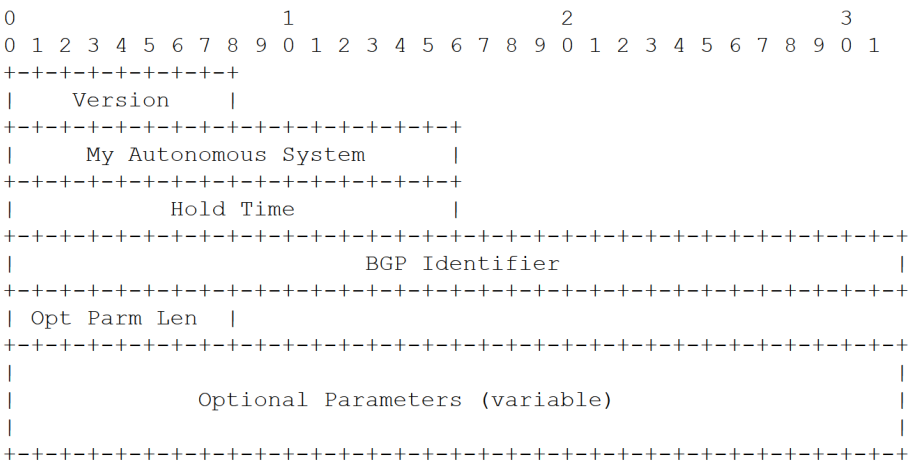
- بخش Marker در سرآیند باید تماماً با بیت‌های ۱ پر می‌شود و اندازه‌ی آن ۱۲۸ بیت است.
- مقدار Type برای پیام KEEPALIVE ، UPDATE و OPEN به ترتیب برابر ۴، ۲ و ۱ است.
- مقدار Length برابر طول کل پیام به بایت است.

۱.۳. پیام OPEN

بسته‌های پیام OPEN طبق قوانین زیر ساخته می‌شوند:

- 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.

[۳] برای اطلاع از استاندارد جهانی BGP می‌توانید به RFC 4271: A Border Gateway Protocol 4 (BGP-4) مراجعه کنید.



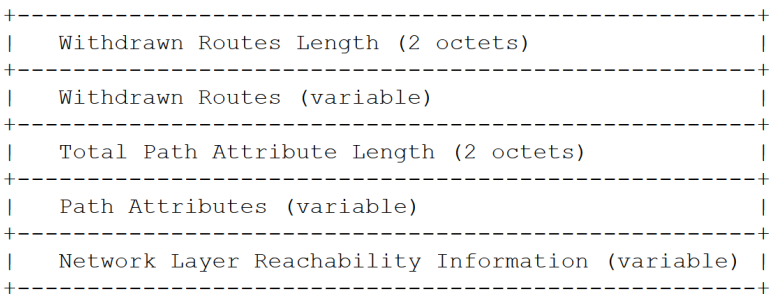
شکل ۵: قالب پیام OPEN مرجع شکل

۲.۳. پیام KEEPALIVE

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.

۳.۳. پیام UPDATE

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



شکل ۶: قالب پیام UPDATE مرجع شکل

به بیان دیگر این بسته برای اطلاع‌رسانی مسیرها بین AS ها استفاده می‌شود؛ به این صورت که یک AS مسیرهای جدیدی که به پیشوندهای مختلف آدرس (یعنی IP به همراه mask subnet آن) می‌شناسد را به همتای خود می‌گوید؛ و همچنین مسیرهای قدیمی‌ای که دیگر معتبر نیستند را نیز اطلاع دهد. این بسته طبق قوانین زیر ساخته می‌شود و شامل مسیر(Path)های غیر قابل دسترس (withdrawn) و مسیرهای جدید قابل دسترس است:

- Withdrawn Routes Length تعداد پیش‌وندهای موجود در Withdrawn Routes را مشخص می‌کند. طول این بخش همواره مضربی از ۵ است و ابتدا ۴ بایت آن به IP و ۱ بایت بعدی به mask subnet یک

پیشوند اختصاص می‌یابد (که عددی بین ۰ تا ۳۱ است). پیش‌وندهای مختلفی که در این بخش قرار می‌گیرند پشت سر هم 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 بیان یک مسیر 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.
یک AS به دلایل اقتصادی مسیرهایی که از یک مسیر همتا_همتا به وی تبلیغ شده باشند را به سایر مسیرهای همتا_همتا تبلیغ نمی‌کند ولی در روابط مشتری_تامین کننده همواره تمامی مسیرها را به مشتری خود تبلیغ می‌کند. 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.
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 واسط 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

بدنه اصلی کد شما در این تمرین در کلاس SimulatedMachine خواهد بود. در ابتدای کار برای دریافت اطلاعاتی از قبیل شماره‌ی AS و همسایه‌هایش، نوع روابط و آی‌پی واسط‌ها و نیز پیشوندهایی که در اختیار این AS است از تابع getCustomInformation استفاده کنید که این اطلاعات را در قالب یک رشته به شما می‌دهد.

۱.۶. برقراری ارتباط

شروع برقراری ارتباط با دستور start connection on interface <I> 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>
```

نام حالت را به صورت IDLE, CONNECT, ACTIVESTATE, OPENSTATE, OPENCONFIRM, ESTABLISHED با حروف تماماً بزرگ چاپ کنید.

۲.۶. تبلیغ مسیرها

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

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

۳.۶. چاپ مسیرها

در صورت دریافت دستور print routes to <PREFIX> تمامی مسیرهای یاد گرفته شده به این پیش‌وند را به همان ترتیب گفته‌شده در بخش انتخاب مسیر چاپ کنید. دقت کنید که باید هر مسیر را در یک خط چاپ کنید. در صورتی

که هیچ مسیری به این پیش‌وند وجود نداشته باشد خروجی زیر چاپ می‌شود:

```
no routes found for <PREFIX>
```

برای چاپ یک مسیر نیز شماره‌ی AS های آن را به ترتیب بنویسید و در انتها پیشوند را چاپ کنید؛ مثلاً:

```
1 3 4 4.0.0.0/8
```

۴.۶. تغییر اولویت‌ها

در صورت دریافت دستور priority of <I> is <X> اولویت مسیرهایی که شروع آن‌ها از واسط I است را به X 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 برابر۸۰ است.

۵.۶. خراب‌کاری در شبکه

در صورت دریافت دستور hijack <PREFIX> 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.

۶.۶. حذف مسیرها

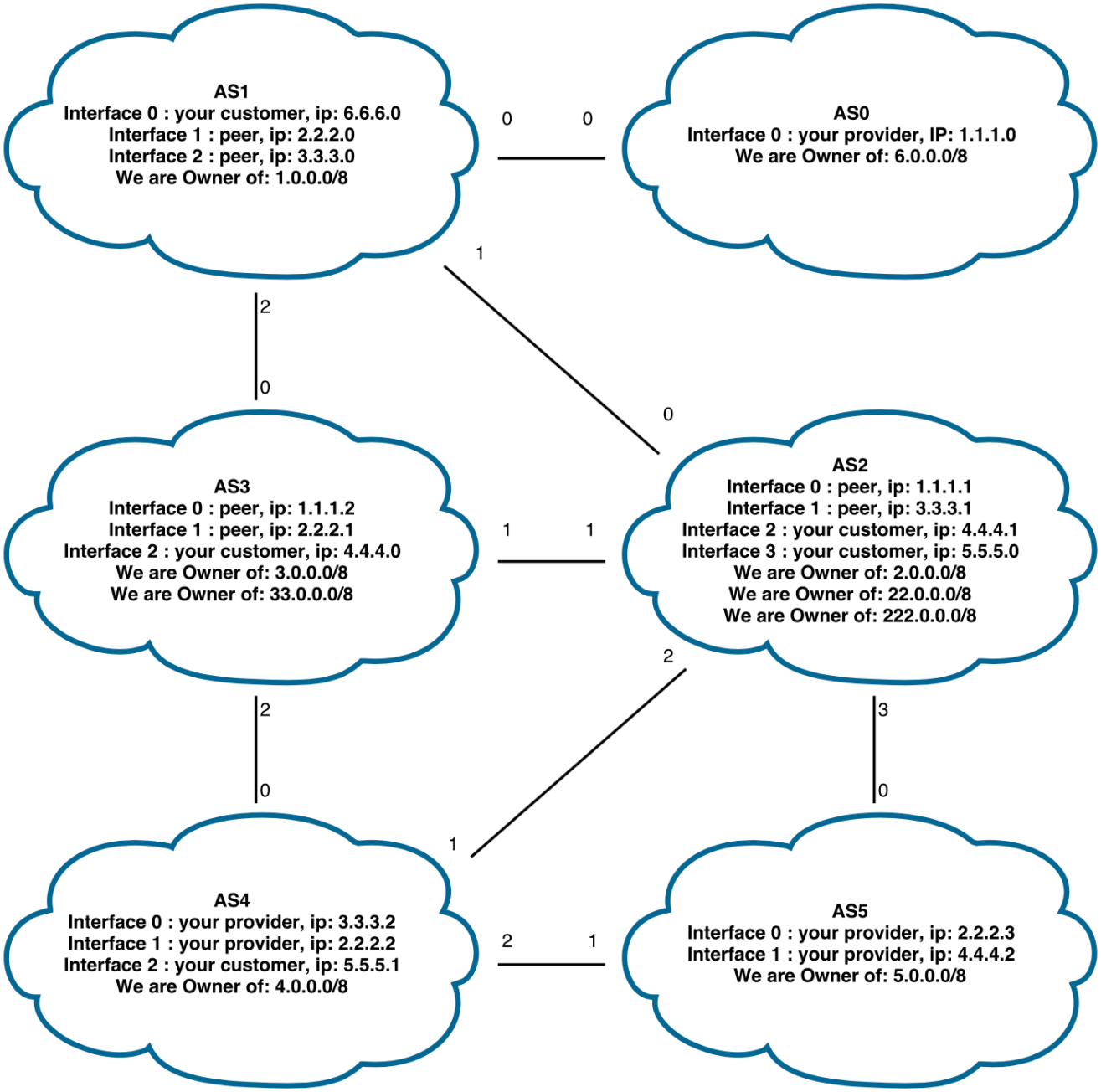
در صورت دریافت دستور withdraw <PREFIX> این پیشوند را پاک کرده و پیام UPDATE مربوط به این مساله را به تمامی همسایه‌ها ارسال کنید. ابتدا از واسط شماره صفر خود شروع کنید. همسایه‌ها نیز به محض دریافت این پیام، مسیری را که از ما به این پیش‌وند می‌رسیده را پاک کرده و این مساله را به همسایه‌های دیگر خود اعلام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

۷. نقشه‌ی تمرین و آزمون‌های نمره‌دهی

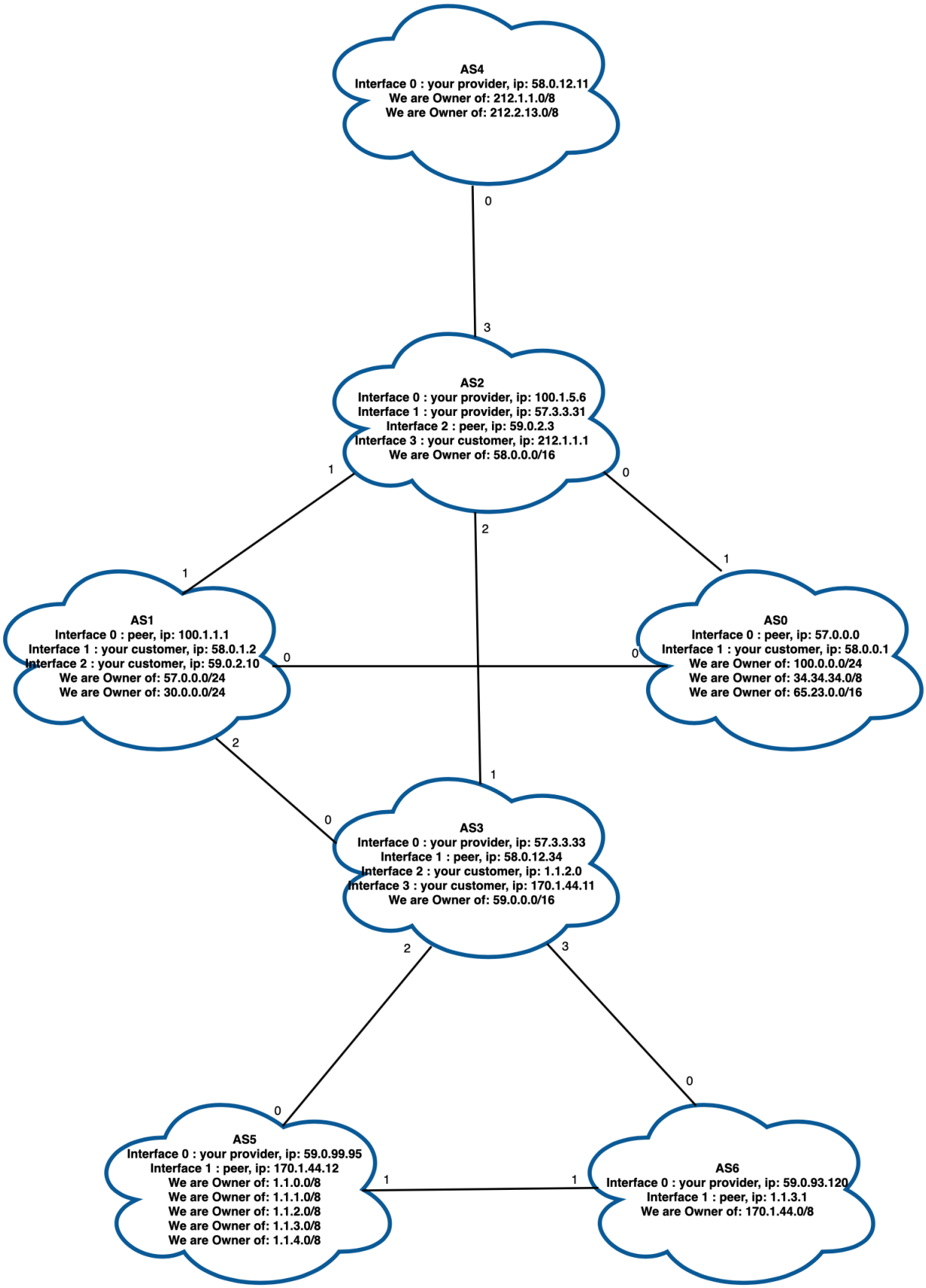
در صفحه‌های بعد می‌توانید نقشه‌هایی که در اختیار شما قرار داده شده تا کدهای خود را تست کنید را مشاهده فرمایید. برای استفاده از نقشه‌ی اول در فایل info.sh نام نقشه را `BGP_Simple` قرار دهید و برای استفاده از نقشه‌ی دوم این نام را به `BGP_New` تغییر دهید

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

- برقراری ارتباط: در این بخش تنها برقراری ارتباط بین AS ها و درستی آنها سنجیدشده 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 ASها 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 ۷: نقشه BGP_Simple



شکل ۸: نقشه BGP_New

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 صورتی‌که 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 زیپی 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 .
- نسخهٔ نهایی تمرین را در مخزنِ خود در وب سایت طرشت بارگذاری نمایید.