

Autor: Dominik Mańkowski

KIEROWANIE ORAZ MAPOWANIE ETYKIET

LEGENDA

IP = Internet Protocol

MPLS = MultiProtocol Label Switching

LSR = Label Switching Router

LER = Label Edge Router

LSP = Label Switched Path

Ingress LER = pierwszy router na LSP

Egress LER = ostatni router na LSP

FEC = Forwarding Equivalence Class = klasa, do której należy pakiet; lokalne ID tunelu, do którego zostanie skierowany pakiet;

MPLS-FIB = MPLS-Forwarding Information Base = tabela mapująca np. prefix adresu docelowego do FEC (zamiast prefixu adresu docelowego może być coś innego, np. pełny adres docelowy, port wejściowy routera, adres źródłowy pakietu, prefix tego adresu lub inne)

IP-FIB = tabela routingu po IP (czyli najzwyczajny w świecie Longest Matching Prefix); mapuje prefix lub całość adresu docelowego do portu wyjściowego

NHLFE = Next Hop Label Forwarding Entry = zawiera informacje takie jak (ID, outPort, outLabel, operation = pop/push/swap, ptr_to_next_ID)

FTN = FEC To NHLFE = tablica występująca tylko w ingress LER; dzięki niej możemy dla danego FEC odnaleźć odpowiedni zbiór wskaźników do wpisów NHLFE

ILM = Incoming Label Map = słownik (incPort, incLabel, poppedLabels) do zbioru wskaźników do wpisów NHLFE

1) Wchodzi goły pakiet IP do sieci MPLS

2) Ingress LER przegląda tablicę MPLS-FIB, aby dowiedzieć się, jaka jest FEC pakietu. Ta tabela wygląda w ten sposób: [parametr_pakietu, FEC], gdzie parametrem pakietu może być np. adres docelowy, adres źródłowy, prefix któregoś z tych adresów, port wejściowy routera lub inne. Jeżeli FEC == 0, to routuj po IP, czyli idź do tabeli IP-FIB i tam znajdź port wyjściowy dla pakietu. Jeśli nie znaleziono FEC odpowiadające takiemu parametrowi pakietu, to odrzuć pakiet. A jeśli FEC != 0, to...

	MPLS-FIB	
Router LER	destAddress	FEC
R1	172.16.222.34	0
R1	172.16.222.33	0
R1	188.88.141.12	31
R1	119.44.55.200	41
R1	83.139.22.49	51

	IP-FIB	
Router LER	destAddress	outPort
R1	172.16.222.34	20300
R1	172.16.222.33	20400
R3	188.88.141.12	11111
R5	83.139.22.49	23184
R6	119.44.55.200	5432

- 3) Ingress LER przegląda tablicę FTN, aby odnaleźć odpowiedni zbiór wskaźników do wpisów NHLFE dla znalezionej wcześniej FEC. Z tego zbioru musi wybrać jeden, konkretny NHLFE (w projekcie zawsze będzie tylko 1 wpis NHLFE, ale np. jakbyśmy chcieli zrobić LSP Multicast, które może służyć np. do load-balancingu, to wtedy potrzebowalibyśmy tych wpisów wiele).

	FTN	
Router LER	FEC	NHLFE_ID
R1	31	1
R1	41	2
R1	51	3
R3	13	31
R3	43	32
R3	53	33

- 4) Mamy więc wskaźnik do jakiegoś wpisu NHLFE, idziemy tam, stamtąd patrzymy na akcję - może to być tylko push (bo mamy goły pakiet). Czyli przykładowo, na rysunku u dołu idziemy do NHLFE_ID=1, operacja to PUSH (mamy tylko 3 operacje: POP, PUSH i SWAP) - wsadzamy etykietę 17 na szczyt stosu i patrzymy na nextNHLFE_ID - jakieś jest, więc tam idziemy (a zatem outPort to NULL, bo i tak z niego teraz nie skorzystamy). Pod ID=5 mamy akcję dołóż etykietę 44 na stos etykiet i wyślij pakiet portem 13997

		NHLFE			
Router	NHLFE_ID	action	outLabel	outPort	nextID
R1	1	PUSH	17	NULL	5
R1	2	PUSH	66	13999	NULL
R1	3	PUSH	22	NULL	5
R1	4	POP	NULL	NULL	NULL
R1	5	PUSH	44	13997	NULL

- 5) Trzeba pamiętać, że etykieta na najniższym poziomie (tj. ta na samym dole stosu, poziom 1) ma ustawiony bit „s” - bottom of the stack - na 1. W każdej etykiecie na wyższym poziomie - na 0. Parametr ten indykuje, kiedy router ma skończyć zczytywać bity należące do tego pakietu [bo cały pakiet wygląda tak: część IP, etykieta na poziomie „m”, „m-1”..., „1”]
- 6) ...oraz należy pamiętać, że pierwsze 16 (licząc od 0) etykiet jest zarezerwowanych.
- 7) I taki pakiet MPLS-owy idzie do routera LSR.
- 8) LSR to odbiera, sprawdza w tablicy ILM jaki jest wskaźnik do wpisu NHLFE (a właściwie to zbiór wskaźników, ale znowu, wybierzemy i tak tylko jeden). Samym kluczem w tym słowniku jest konkatenacja [incPort, incLabel, poppedLabelStack], a wartością NHLFE_ID. Wyobraźmy sobie taką sytuację:
Przychodzi pakiet z portu 100 w tunelu o etykiecie 50, w którym zagnieżdżony jest tunel o etykiecie 40, w którym zagnieżdżony jest tunel o etykiecie 30. I jednocześnie przychodzi pakiet - również do portu 100 - o etykiecie 30 (tj. niezagnieżdżonej nigdzie). I powiedzmy, że zdejmujemy te dwie etykiety 50 i 40. Jak teraz odróżnić te 30? Trzeba patrzeć na to, co zostało

zdjęte. Czyli jak przyjdzie tunel z etykietą 50, to najpierw znajdzie mu wpis w ILM:
 [100, 50, -, 66], czyli port wejściowy 100, etykieta wejściowa 50, brak zdjętych etykiet i NHLFE_ID to 66. I pod tym NHLFE_ID=66 powiedzmy, że będzie tylko POP, nic więcej - czyli mamy zakończenie tunelu. Nie możemy dać w tym NHLFE wpisu next_NHLFE_ID, ponieważ jeśli by była sytuacja, że obok tunelu 40 idzie 41, i jeden chciałby iść w prawo, a drugi w lewo - to mamy problem.

No to musimy wrócić zatem do ILM. I tutaj znajdujemy wpis:

[100, 40, 50, 66] - i znowu zdejmujemy jedną etykietę. I znowu idziemy do ILM:

[100, 30, 50-40, 67] - i ten wpis odróżniamy od innych „30” na podstawie tego, że zdjęliśmy z niego dwie etykiety. I idziemy do NHLFE_ID=67, gdzie będzie jakiś SWAP, a tam może być już next_NHLFE_ID=NULL, czyli będzie jakiś outPort i to wypływamy, albo outPort=NULL i next_NHLFE_ID = np. 68, gdzie będzie np. PUSH i outPort, albo nextNHLFE_ID, itd.

	ILM			
Router	incPort	incLabel	poppedLabelStack	NHLFE_ID
R1	13999	66	-	4
R1	13997	17	-	4
R1	13997	22	-	4
R2	1025	44	-	21
R2	1026	93	-	22
R2	1026	28	-	23
R3	44444	58	-	34
R3	44444	17	58	34
R3	44444	22	58	35

- 9) Jeśli LSR nie znajdzie takiego klucza w ILM, to pakiet zostaje odrzucony (*dropped*) - nie można routować po IP, bo by mogła powstać pętla (chyba, że jesteśmy w stanie zapewnić, że nie powstanie) - ale domyślnie odrzucamy. Jeśli jednak by pętla powstała, to wówczas, po krótkim czasie, wartość TTL spadnie do 0 - i wtedy pakiet zostanie definitywnie odrzucony (i jeszcze router będzie mógł zajrzeć w głębi nagłówka IP pakietu, aby - ewentualnie - wysłać wiadomość o błędzie ICMP do adresu źródłowego)
- 10) I wysyłamy ten pakiet dalej LSP
- 11) Pakiet dochodzi do egress LER, ten sprawdza w ILM co ma zrobić. I tam znajduje wpis NHLFE, który będzie miał akcję POP (i całą resztę równą NULL) - i będziemy mieli goły pakiet IP teraz
- 12) Patrzymy więc do MPLS-FIB, tam szukamy FEC
- 13) Jeśli FEC == 0, to idziemy do IP-FIB i szukamy portu wyjściowego
- 14) Jeśli FEC != 0, to oznacza to, że wyszliśmy z jednego tunelu, wchodzimy w inny tunel; czyli idziemy do FTN, stąd NHLFE_ID, i reszta tak jak wyżej.