

Od nazwy do konkretnej maszyny

Marek MARCZYKOWSKI*

1. Wprowadzenie

Od wpisania adresu w przeglądarce internetowej do wyświetlenia strony musi być wykonany szereg mniej lub bardziej skomplikowanych czynności. W tym artykule omówię najważniejsze z nich: tłumaczenie nazwy na adres IP oraz lokalizację, na podstawie adresu IP, fizycznej maszyny, do której on należy (routing). To ostatnie jest szczególnie interesujące w tak rozległej sieci, jaką jest internet.

2. DNS

DNS (*Domain Name System*) jest rozproszonym systemem służącym do tłumaczenia nazw na adresy IP. Wbrew temu, co normalnie widzimy, każda nazwa jest zakończona kropką. System ten ma budowę hierarchiczną (struktura drzewa). Na samym szczycie znajduje się najkrótsza możliwa domena – kropka, obsługiwana przez tzw. root servery. Znają one adresy serwerów obsługujących domeny najwyższego poziomu (ang. *Top Level Domain*, TLD), czyli np. *com.*, *pl.*. Te z kolei znają servery obsługujące domeny pod nimi, np. *com.pl.*, *edu.pl.* i tak dalej (patrz rysunek). Każda taka domena składa się (jako węzeł drzewa) z ciągu rekordów zawierających różne dane powiązane z daną nazwą, np. adres IP, opis, alternatywna nazwa, serwer pocztowy obsługujący daną domenę.

Z punktu widzenia użytkownika końcowego całą tę operację odpytywania po kolei *.*, *pl.*, *edu.pl.*, *mimuw.edu.pl.*, *www.mimuw.edu.pl.* wykonuje serwer dostawcy internetowego, który przy okazji zapamiętuje ostatnio używane nazwy. Dlatego w konfiguracji sieci wystarczy wpisać (lub samo się konfiguruje przez DHCP) tylko pojedyncze adresy serwerów DNS.

3. Routing

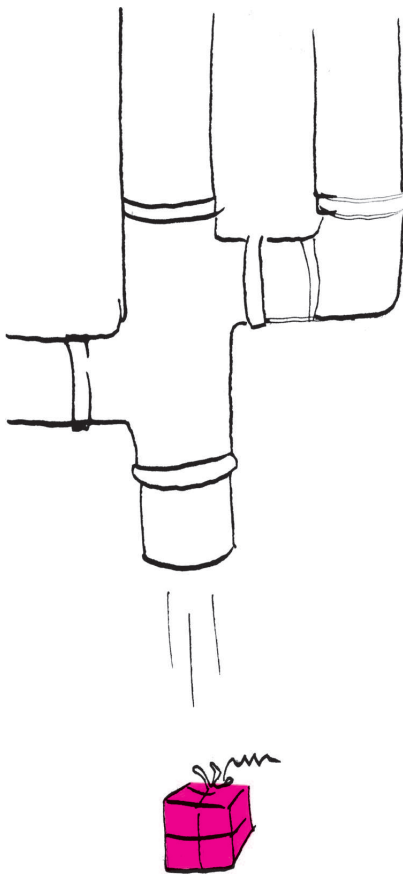
3.1. Podstawowy routing. Wiemy już, jak uzyskać adres IP na podstawie nazwy, ale co dalej? Z punktu widzenia końcowego użytkownika sytuacja jest prosta: ma jedną domyślną bramę (do dostawcy internetowego), do której wysyła wszystkie pakiety, a dostawca już jakoś sobie z tym dalej radzi. W dalszej części omówię właśnie radzenie sobie dostawcy z tym problemem. Wprowadźmy najpierw kilka pojęć, które będą przydatne przy dalszym opisie:

- Prefiks – zakres adresów IP w formie: adres i maska. Do takiego zakresu należą te adresy, w których zapisie binarnym bity wskazane przez maskę są takie same jak w adresie wyznaczonym przez prefiks. Maska zazwyczaj składa się z pewnej liczby jedynek, po których następują już same zera. Np. prefiks 192.168.0.0/24 oznacza wszystkie adresy, w których pierwsze 24 bity (czyli 3 oktety) są równe 192.168.0.
- System autonomiczny (ang. *autonomous system*, AS) – sieć należąca do jednego podmiotu, mająca wspólną politykę zarządzania, np. sieć Telekomunikacji Polskiej.
- Tablica routingu – tablica, na podstawie której system operacyjny routera podejmuje decyzję, dokąd wysłać dany pakiet.

Kierowaniem pakietów odpowiednim łączem i do odpowiedniego operatora zajmują się routery. Każdy router ma tablicę routingu. Taka tablica składa się z ciągu prefiksów i powiązanych z nimi akcji (np. kolejny router, do którego ma zostać wysłany pakiet). Prefiksy są przeglądane w kolejności od najdłuższego (najbardziej szczegółowego) do najkrótszego (najogólniejszego) aż do napotkania pasującego do pakietu, co do którego jest podejmowana decyzja. Np. dla pakietu przesyłanego do 192.168.2.24 i tablicy routingu:

```
127.0.0.0/8 via 127.0.0.1 dev lo
192.168.0.0/16 via 192.168.1.10 dev eth0
192.168.2.0/24 via 192.168.1.11 dev eth0
192.168.1.0/24 dev eth0
```

*student i administrator,
Wydział Matematyki, Informatyki
i Mechaniki, Uniwersytet Warszawski



Rozwiązanie zadania F 760.

Kawalki balonika wyrzucone pod kątem α do poziomu spadną w odległości

$$l = \frac{v^2}{g} \sin 2\alpha$$

od balonika. Znajdą się one na zewnątrz koła o promieniu R , jeśli

$$\frac{1}{2} \arcsin \frac{Rg}{v^2} \leq \alpha \leq \frac{\pi}{2} - \frac{1}{2} \arcsin \frac{Rg}{v^2}.$$

Stosunek masy strzępków wylatujących na odległość większą niż R do masy całkowitej balonika jest równy stosunkowi powierzchni S pasa na powierzchni wybuchającego balonika, leżącego między kątami

$$\frac{1}{2} \arcsin \frac{Rg}{v^2} \quad \text{oraz} \quad \left(\frac{\pi}{2} - \frac{1}{2} \arcsin \frac{Rg}{v^2} \right),$$

do całkowitej powierzchni balonika S_0 . Oznaczając przez r promień balonika, otrzymujemy:

$$\begin{aligned} \frac{S}{S_0} &= \frac{2\pi r}{4\pi r^2} (h_2 - h_1) = \\ &= \frac{1}{2} \left[\sin \left(\frac{\pi}{2} - \frac{1}{2} \arcsin \frac{Rg}{v^2} \right) - \right. \\ &\quad \left. - \sin \frac{1}{2} \arcsin \frac{Rg}{v^2} \right] \\ &= \frac{\sqrt{2}}{2} \sin \left(\frac{\pi}{4} - \frac{1}{2} \arcsin \frac{Rg}{v^2} \right). \end{aligned}$$

Zatem

$$m = \frac{M}{\sqrt{2}} \sin \left(\frac{\pi}{4} - \frac{1}{2} \arcsin \frac{Rg}{v^2} \right).$$

pakiet zostanie przesłany do routera 192.168.1.11 przez interfejs (kartę sieciową) *eth0*, natomiast skierowany do 192.168.1.32 będzie wysłany bezpośrednio do celu przez interfejs *eth0*.

3.2. Routing dynamiczny. W praktyce tablica routingu każdego większego routera jest generowana dynamicznie w zależności od tego, które łącza działają, którądy będzie najkrótsza trasa itp. W internecie służy do tego protokół **BGP** (*Border Gateway Protocol*). Podstawowa koncepcja polega na podziale całej sieci (internetu) na systemy autonomiczne. Każdy z takich systemów ma jakieś adresy IP w postaci zbioru prefiksów i ogłasza ten fakt systemom, z którymi jest bezpośrednio połączony. Te systemy z kolei ogłaszają swoim „sąsiadom”, że przez nie można dostać się do określonych prefiksów należących do konkretnego AS-u, i tak dalej. Wszystko wydaje się proste, gdyby nie fakt, że wielokrotnie dany prefiks jest osiągalny na więcej niż jeden sposób. Tu właśnie objawia się potęga protokołu BGP.

Każdy router BGP ma tablicę, w której dla każdego prefiksu są przechowywane trasy, którymi jest on osiągalny (trasa to ciąg systemów autonomicznych). Od każdego sąsiada może on otrzymać jedną taką trasę (najkorzystniejszą dla niego) i powiązać z nią pewne dodatkowe informacje (o tym dalej). Podstawowe kryteria, na podstawie których jest wybierana najlepsza trasa, to:

1. waga trasy (ustawienie lokalne dla routera),
2. lokalna preferencja (*localpref*) – wspólne dla całego AS-u,
3. pochodzenie trasy (z wewnętrznego systemu, zewnętrznego itp.),
4. długość ścieżki,
5. sugerowana metryka.

Pierwsze dwa kryteria zależą bezpośrednio od administratora i są wewnętrzne dla danego AS (nie są przekazywane na zewnątrz) – są one nadawane na podstawie reguł zdefiniowanych w routerze (najczęściej na podstawie tego, od którego routera trasa przyszła oraz którym łączem). W ten sposób w przypadku prefiksów dostępnych przez AS *A* oraz AS *B* można wymusić wysyłanie pakietów do tego pierwszego (gdyż np. mamy z nim lepsze połączenie). Tak można decydować, którądy wysyłać pakiety, a co z pakietami powrotnymi? Tu też można manipulować. Załóżmy, że jesteśmy w AS-ie *X* i spośród dwóch AS-ów pośrednich *A* i *B* chcielibyśmy zawsze dostawać pakiety tylko za pośrednictwem *A*. Wystarczy wówczas operatorowi *B* ogłaszać trasy mniej korzystne – konkretnie przez wydłużenie ścieżki, dodając siebie (tj. *X*) wielokrotnie na końcu. W ten sposób zdalny router, mając zdecydować, której trasy użyć, użyje krótszej (czyli przez AS *A*) – oczywiście, o ile nie zmieni tego ważniejszymi kryteriami.

Kolejnym mechanizmem pozwalającym kształtować przepływ ruchu są etykiety (*community*), którymi można dodatkowo oznaczać trasy. Mają one dwa zastosowania:

- dodanie dodatkowych informacji dla kolejnych routerów (np. skąd daną trasę otrzymano),
- zmianę traktowania danej trasy (np. ustawienie innych lokalnych preferencji, przekazywanie wydłużonej trasy do jednego z sąsiadów lub nieprzekazywanie trasy dalej).

Mając do dyspozycji tak potężne narzędzie, można bardzo precyzyjnie modelować, którądy jakie pakiety będą przepływać. Mówi się, że BGP służy do zarządzania w warstwie ósmej sieci – politycznej. Jest w tym sporo prawdy, ponieważ większość ustawień i nietypowych zachowań jest tutaj skutkiem umów między operatorami (gdzie taniej, gdzie lepsze warunki i gwarancja jakości łącza, oraz, oczywiście, gdzie większa przepływność czy przepustowość). Zdarza się, że np. ruch z Uniwersytetu Warszawskiego do polskich adresów jest kierowany przez jednych sąsiadów, do zagranicznych przez innych, a ruch do Telekomunikacji Polskiej (TPNET) przez jeszcze innych (tak – to nie żart).



Rozwiązanie zadania M 1270.

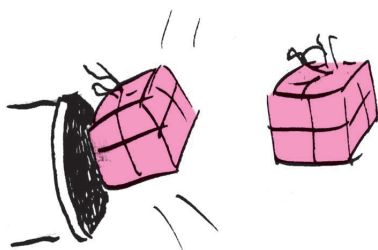
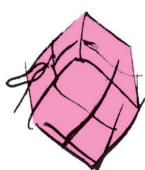
Zauważmy,

$$\text{że } 1 + p^2 = pq + qr + rp + p^2 = \\ = (p + q)(p + r) \text{ oraz analogicznie}$$

$$1 + q^2 = (q + p)(q + r),$$

$$1 + r^2 = (r + p)(r + q).$$

Wobec tego $(1 + p^2)(1 + q^2)(1 + r^2) = \\ = ((p + q)(p + r)(r + p))^2$, skąd bezpośrednio wynika teza.



Przy tak istotnym dla działania całego internetu mechanizmie istotne jest również dbanie o bezpieczeństwo, niezawodność oraz wydajność. To pierwsze w większości sprowadza się do stosowania filtrów, kto może, a kto nie, rozgłaszać trasy do konkretnych prefiksów (na podstawie danych urzędów przyznających adresy, np. RIPE). Niedopatrzenie w tej dziedzinie jakiś czas temu spowodowało, że pewien pakistański operator, chcąc zablokować dostęp do serwisu YouTube dla swoich klientów, zablokował go dla sporej części świata na kilka godzin (artykuł na ten temat: <http://www.ripe.net/news/study-youtube-hijacking.html>). W kwestii niezawodności i wydajności istnieje szereg mechanizmów. Oto kilka z nich:

BGP Route Flap Damping. Polega on na tym, że gdy jakiś router wysyła zbyt często informacje o zmianie konkretnej trasy, to sąsiad je otrzymujący przestaje na nie na pewien czas reagować, aby nie generować ciągłych zmian w tablicach routingu na wielu routerach w całej sieci, co mogłoby prowadzić do istotnego zwiększenia ich obciążenia oraz, w dużym stopniu, do utraty pakietów i ciągłej zmiany ich kolejności.

Agregacja prefiksów. Ponieważ pełna tablica routingu BGP wymaga coraz więcej zasobów (pamięci, procesora), stosuje się łączenie kilku wpisów w jeden. Np. jeśli jakiś AS rozgłasza adresy 193.0.96.0/24, 193.0.97.0/24, 193.0.98.0/24 oraz 193.0.99.0/24, to można to zwinąć do 193.0.96.0/22 obejmującego wszystkie te adresy. W takiej sytuacji cała podsieć jest rozgłaszana, jeśli przynajmniej jeden z jej elementów jest osiągalny. Dzięki temu redukuje się wielokrotne wpisy z taką samą trasą i zastępuje jednym.

„Miękka” rekonfiguracja. Przy zmianie konfiguracji (np. ustawianiu parametrów *localpref*) istnieje konieczność ponownego przetworzenia tras, których ta zmiana dotyczy. Najprościej by było zrestartować proces, co spowodowałoby wyczyszczenie tablicy oraz zerwanie sesji BGP z sąsiednimi routerami, a następnie zestawienie ich z powrotem i pobranie aktualnych informacji. Taka operacja powoduje w praktyce, że router na kilka minut znika z internetu oraz wszystkie routery, do których był podłączony, mają dodatkową pracę w postaci ponownego przetworzenia tras, które przez niego przechodziły. Aby temu zapobiec, wprowadzono mechanizm umożliwiający tylko częściowe odświeżenie tablicy, np. ponowne przesłanie tras przechodzących przez dany AS (nie ruszając pozostałych).

4. Podsumowanie

Większość użytkowników internetu nie ma bezpośredniej styczności z głównym tematem tego artykułu, ale wydaje mi się, że warto wiedzieć, jak to działa, że to nie jest jakieś magiczne pudełko, które robi cuda, tylko efekt wieloletniej pracy inżynierów sieciowych, która teraz pozwala funkcjonować sieci na tak olbrzymią skalę, jaką jest cała Ziemia.

Reguła czy przypadek?

Gdy do wielomianu $x^2 + x + 11$ podstawiać będziemy kolejno liczby 0, 1, 2, ..., 9, otrzymamy 11, 13, 17, 23, 31, 41, 53, 67, 83, 101 – same liczby pierwsze. Dla 10 już tak nie będzie: wartość wielomianu dla $x = 10$ dzieli się przez 11.

To akurat dziwić nie powinno – każdy wielomian $x^2 + x + a$ dla $x = a - 1$ dzieli się przez a , ponieważ $(a - 1)^2 + (a - 1) + a = a^2$.

Ale Czytelnik Dociekliwy zauważy, że różnice kolejnych wartości to kolejne liczby parzyste: 2, 4, 6, ..., 18. Czy to też reguła? I to okazuje się także prawidłowością, bowiem $k^2 + k = k(k + 1) = 2 \cdot (1 + 2 + \dots + (k - 1) + k)$, a więc wartości każdego wielomianu $x^2 + x + a$ przy

podstawieniu kolejnych liczb naturalnych m i $m + 1$ różnić się będą o $2(m + 1)$.

No to może wreszcie i fakt, że podstawiając kolejne liczby naturalne 0, 1, ..., $(p - 2)$ do wielomianu $x^2 + x + p$, gdzie p jest – jak 11 – liczbą pierwszą, otrzymujemy same liczby pierwsze, też jest prawidłowością? Np. dla 3 czy 5 się zgadza!

Okazuje się, że nie: poza wymienionymi trzema liczbami pierwszymi jest tak jeszcze tylko dla 17 i 41. Tyle że nie jest to, oczywiście, przypadek. Ale dowód tego faktu nie jest prosty, choć – zapewne – może sprawić Czytelnikowi Studiującemu wiele satysfakcji.

M. K.