Protokół ICMP

Autor: Grzegorz Burgiel
4FDS

Streszczenie

Niniejsze opracowanie opisuje protokół ICMP : formaty komunikatów kontrolnych i zastosowanie protokołu.

Spis treści

1. Wstęp

Protokół IP (Internet Protocol) jest protokołem zawodnym, realizującym przenoszenie pakietów w sposób bezpołączeniowy. Pakiety przekazywane są w postaci datagramów. Istnieje wiele powodów w wyniku których datagramy nie zostaną dostarczone do miejsca docelowego np.: datagramy nie zostaną dostarczone, gdy maszyna do której presyłamy datagramy jest na stałe lub tylko czasowo odłączona od sieci, albo gdy routery będące na drodze datagramów są tak przeciążone , że nie mogą przetworzyć przybywających pakietów.

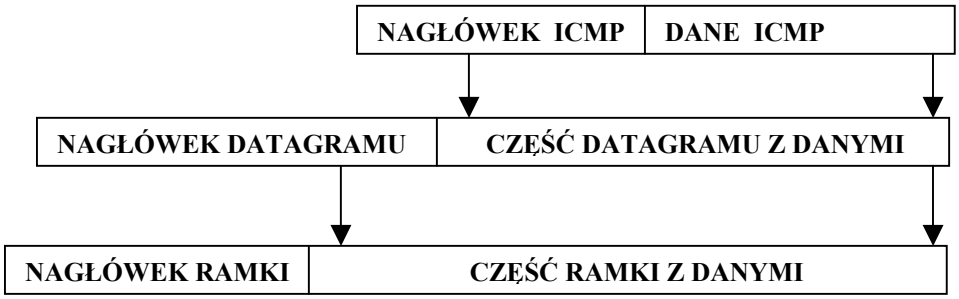
Protokół IP nie zwiera w sobie mechanizmów umożliwiających nadawcy sprawdzenie np.: stanu połączenia, czy wysłane datagramy dotarły do miejsca docelowego.

„Aby umożliwić routerom w intersieci oznajmianie o błędach oraz udostępnianie informacji o niespodziewanych sytuacjach”[1] do grupy protokołów TCP/IP wprowadzono protokół kontrolny – ICMP (ang. Internet Control Message Protocol).

Komunikaty przesyłane są poprzez sieć w częściach datagramów IP przeznaczonych na dane. „Odbiorcą końcowym komunikatu ICMP nie jest ani program użytkowy, ani użytkownik”[1], lecz moduł oprogramowania ICMP na danej maszynie. Pierwotnie ICMP powstał, aby „umożliwić routerom powiadamianie węzłów o powodach błędów w dostarczaniu datagramów”[1] jednak może być używany nie tylko przez routery. Każdy węzeł może za pomocą ICMP wymieniać informacje z routerem lub innym węzłem.

2. Dostarczanie komunikatów ICMP

Aby komunikat ICMP mógł zostać dostarczony wymaga dwóch poziomów kapsułkowania. (rys.2) Komunikat ICMP podróżuje przez sieć w części datagramu IP przeznaczonej na dane, a ten przemieszcza się przez sieć fizyczną w części dla danych ramki. Trasy datagramów przenoszących komunikaty ICMP są wyznaczane dokładnie w ten sam sposób jak dla datagramów przenoszących informacje - nie maja one żadnych dodatkowych priorytetów ani zabezpieczeń. Istnieje więc możliwość ich zagubienia albo zniszczenia. W sieci przeciążonej komunikat o błędzie może spowodować dodatkowe przeciążenie. W procedurach obsługi błędów istnieje następujący wyjątek: jeśli błąd został spowodowany przez datagram IP niosący komunikaty o błędach ICMP nie jest tworzony komunikat o błędzie.



Rys.2. Dwa poziomy kapsułkowania ICMP. [1]

3. Format komunikatu ICMP

Każdy komunikat ICMP ma własny format, jednak „wszystkie te formaty rozpoczynają się trzema takimi samymi polami: 8-bitowym polem TYP komunikatu, które identyfikuje komunikat, 8-bitowym polem KOD, które daje dalsze informacje na temat rodzaju komunikatu, oraz polem SUMA KONTROLNA (ICMP używa tego samego algorytmu wyliczenia sumy co IP, ale suma kontrolna ICMP odnosi się tylko do komunikatu ICMP)".[1] Komunikaty ICMP zawierają nagłówek i pierwsze 64 bity danych datagramu, z którym były problemy. Odesłanie nie tylko samego nagłówka datagramu umożliwia odbiorcy bardziej dokładne określenie, który protokół(-oły) lub który program użytkowy jest odpowiedzialny za dany datagram. Pole TYP w komunikacie ICMP definiuje zarówno jego znaczenie, jak i format. Wśród typów wyróżnia się :

Wartość w polu TYP	**Typ komunikatu ICMP**
0	**odpowiedź z echem**
3	**odbiorca nieosiągalny**
4	**tłumienie nadawcy**
5	**zmień trasowanie**
8	**prośba o echo**
11	**przekroczenie terminu datagramu**
12	**kłopot z parametrami datagramu**
13	**prośba o czas**
14	**odpowiedź z czasem**
15	**prośba o informację (przestarzałe)**
16	**odpowiedź z informacją (przestarzałe)**
17	**prośba o maskę ad**
18	**odpowiedź z maską adresową**

Poniżej został opisany każdy z tych komunikatów oraz odpowiadający mu format.

4. Sprawdzanie osiągalności odbiorcy (ang. ping)

„Jedno z najczęściej używanych narzędzi do badania błędów odwołuje się do komunikatów ICMP <prośba o echo> i <odpowiedź z echem>".[1] Węzeł albo router wysyła komunikat ICMP „prośba o echo" do określonego odbiorcy. Maszyna, która odbierze prośbę o echo, tworzy odpowiedź z echem i odsyła ją do pierwotnego nadawcy. „Prośba ta zawiera opcjonalne miejsce na dane, odpowiedź zaś kopię danych wysłanych w prośbie." [1] Prośba o echo i związana z nią odpowiedź używane są do sprawdzania, czy odbiorca jest osiągalny i czy odpowiada. „W związku z tym, że zarówno prośba, jak i odpowiedź wędrują w datagramach IP, pomyślne odebranie odpowiedzi oznacza, że główne części systemu przesyłania danych działają"[1] tzn.:

- oprogramowanie IP komputera prawidłowo skierowało datagram,
- routery pośrednie między nadawcą a odbiorcą prawidłowo skierowały datagramy,
- maszyna docelowa musiała być uruchomiona (reagować na przerwania),
- oprogramowanie ICMP, jak i IP musiało prawidłowo działać,
- routery wzdłuż ścieżki powrotnej musiały prawidłowo wyznaczać trasy.

Istnieją implementacje polecenia *ping* wysyłające serie próśb o echo, rejestrują odpowiedzi i udostępniają statystyki opisujące liczbę zagubionych datagramów. Umożliwiają użytkownikowi także ustawienie długości wysyłanych danych oraz czasu między prośbami.

4.1. Format komunikatów „prośba o echo" i „odpowiedź z echem"

Rysunek nr 4.1 przedstawia format komunikatu „prośba o echo" i „odpowiedź z echem". Pole OPCJONALNE DANE ma zmienną długość i zawiera dane, którymi odbiorca odpowiada nadawcy. „Komunikat <odpowiedź z echem> zawsze zawiera dokładnie te same dane, które przybyły w prośbie." Pola IDENTYFIKATOR i NUMER KOLEJNY są używane przez nadawcę do przyporządkowywania odpowiedzi prośbom. Wartość pola TYP służy w tym przypadku do określania, czy komunikat jest prośbą (8), czy odpowiedzią (0).

0	8	16	31
TYP (8 lub 0)	KOD	SUMA KONTROLNA	
IDENTYFIKATOR		NUMER KOLEJNY	
OPCJONALNE DANE			

Rys. 4.1 Format komunikatów „prośba echo” i „odpowiedź z echem”.

5. Powiadamianie o nieosiągalnych odbiorcach

W przypadku problemów przesyłem dalej lub dostarczeniem datagramu do odbiorcy router wysyła do nadawcy datagramu komunikat „odbiorca nieosiągalny". Format takiego komunikatu jest przedstawiony na rys. 5

0	8	16	31
TYP (3)	KOD(0-12)	SUMA KONTROLNA	
NIE UŻYWANE (MUSI BYĆ RÓWNE ZERU)			
NAGŁÓWEK + PIERWSZE 64 BITY INTERSIECIOWEGO DATAGRAMU			

Rys. 5. Format komunikatu ICMP „odbiorca” nieosiągalny”

Pole KOD komunikatu „odbiorca nieosiągalny" zawiera liczbę. Możliwe wartości to:

Wartość w polu KOD	**Znaczenie**
1	**sieć nieosiągalna**
2	**węzeł nieosiągalny**
3	**protokół nieosiągalny**
4	**konieczna fragmentacja przy ustawionym nie fragmentuj**
5	**błąd trasy nadawcy**
6	**nieznana sieć odbiorcy**
7	**nieznany węzeł odbiorcy**
8	**węzeł nadawcy odizolowany**

9	**komunikacja z siecią odbiorcy zabroniona administracyjnie**
10	**komunikacja z węzłem odbiorcy zabroniona administracyjnie**
11	**sieć niedostępna dla takiego rodzaju usług**
12	**węzeł niedostępny dla takiego rodzaju usług**

Router po wysłaniu komunikatu „odbiorca nieosiągalny" traci datagram. Kilka typów błędów wymaga szerszego komentarza :

- „błąd typu <sieć nieosiągalna> oznacza zazwyczaj kłopoty z wyznaczaniem tras"[1],
- błąd typu „węzeł nieosiągalny" – oznacza, że zawiódł mechanizm dostarczania pakietów.

Komunikat ICMP o błędzie zawiera krótki prefiks datagramu, który spowodował problemy, więc nadawca będzie wiedział dokładnie, jaki adres jest nieosiągalny.

Odbiorcy mogą być nieosiągalni z wielu względów np.:

- awarii sprzętu,
- nieistniejącego adresu odbiorcy,
- w rzadkich przypadkach nieznajomości adresu sieci odbiorcy przez ruter.

W sieciach Ethernet, osprzęt sieciowy nie zapewnia potwierdzeń stąd ruter może w dalszym ciągu wysyłać pakiety do odbiorcy, który nie odpowiada.

6. Kontrola przepływu datagramów

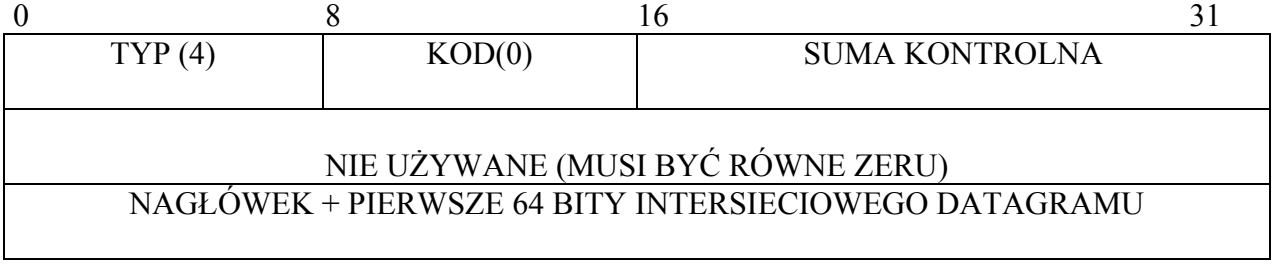
Protokół IP pracuje w trybie bezpołączeniowym, więc router nie może zarezerwować zasobów pamięci ani zasobów komunikacyjnych przed otrzymaniem datagramów. W wyniku tego routery mogą zostać przeciążone (ang. congestion) napływającymi datagramami. Przeciążenie pojawi się na routerze, który łączy LAN i WAN, bo datagramy przybywają szybciej niż mogą być wysłane. „Może się też zdarzyć, że wiele komputerów jednocześnie wysyła datagramy przez dany router, co może spowodować przeciążenie, mimo że żaden pojedynczy nadawca nie wywoła tego problemu."[1] Router używa komunikatu ICMP „tłumienie nadawcy" do powiadamiania o problemie z przeciążeniem. Komunikat „tłumienie nadawcy" to prośba o zmniejszenie liczby wysyłanych datagramów. Przeciążone routery wysyłają po jednym komunikacie „tłumienie nadawcy" dla każdego datagramu, który tracą. Istnieją także bardziej wyrafinowane metody kontroli przeciążenia. Routery badają przychodzące datagramy i wysyłają komunikat „tłumienie nadawcy" do tych, którzy przesyłają najwięcej datagramów. Inne aby

uniknąć przeciążenia, wysyłając prośby, gdy ich kolejki staja się długie, ale jeszcze nie przepełnione.

Nie ma komunikatu ICMP, który służy za odpowiedź na komunikat „tłumienie nadawcy". „Węzeł, który otrzymuje komunikaty tłumienia w związku z datagramami wysyłanymi do odbiorcy D zmniejsza liczbę datagramów wysyłanych do D, aż przestanie otrzymywać komunikaty <tłumienie nadawcy>".[1] Następnie może zwiększać liczbę wysyłanych datagramów, dopóki znów nie otrzyma komunikatu „tłumienie nadawcy".

6.1. Format komunikatu „tłumienie nadawcy"

Oprócz omówionych wcześniej pól oraz nie używanego 32-bitowego pola komunikaty „tłumienie nadawcy" mają pole przeznaczone na prefiks datagramu. Na rysunku nr 6.1 jest przedstawiony ten format.



Rys. 6.1. Format komunikatu „tłumienie nadawcy".

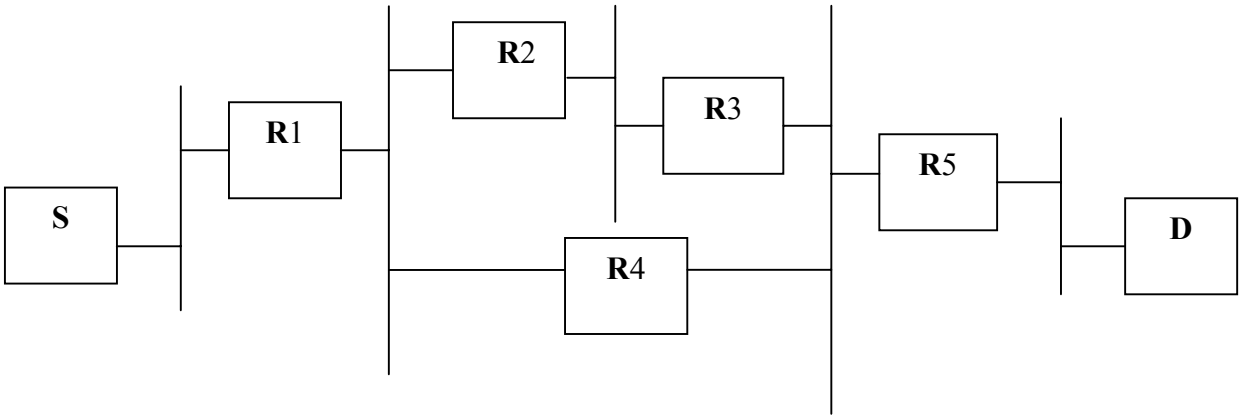
7. Zmiana trasowania

Pod czas startu systemu następuje inicjalizacja tablic tras z pliku konfiguracyjnego. Gdy zmienia się topologia sieci dane zawarte w pliku konfiguracyjnym mogą stać się nieaktualne. Dlatego routery okresowo wymieniają informacje o trasach, aby dostosowywać się do zmian w sieci i utrzymywać w tablicach aktualne dane.

Aby uniknąć duplikowania informacji o trasach w pliku konfiguracyjnym każdego węzła, z początkowej konfiguracji tras jest wyznaczane jest tylko minimum informacji konieczne do nawiązania komunikacji czyli adres jednego routera. W związku z tym węzeł rozpoczynający pracę z minimalną ilością informacji i przy uaktualnianiu swoich tablic tras polega na routerach i gdy router stwierdzi, że węzeł używa nieoptymalnych tras, jest wysyłany do takiego

węzła komunikat „zmień trasowanie", który stanowi prośbę do węzła o zmianę trasy. Dodatkowo przesyłany jest oryginalny datagram do jego odbiorcy.

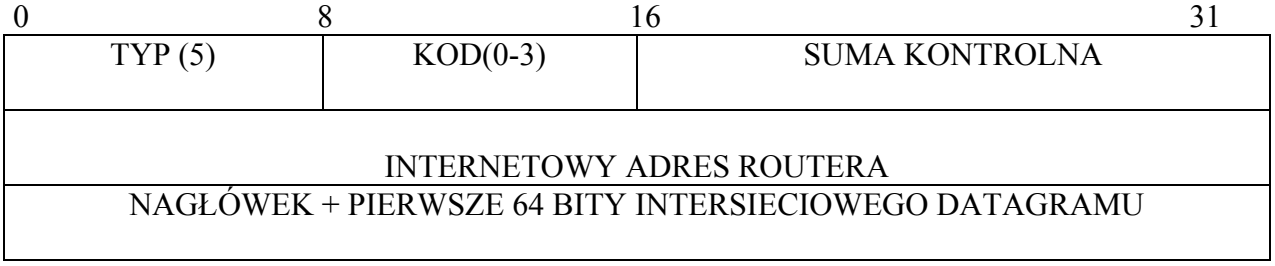
Zaletą schematu zmiany trasowania ICMP jest jego prostota: węzeł może startować, znając adres tylko jednego routera w lokalnej sieci. Ten początkowy router wysyła komunikaty ICMP „zmień trasowanie" za każdym razem, gdy dla danego datagramu istnieje lepsza trasa. Tablica tras węzła pozostaje mała, ale w dalszym ciągu zawiera optymalne trasy dla wszystkich używanych celów.



Rys. 7. Komunikaty ICPM „zmień trasowanie" nie umożliwiają poprawienia tras między routerami. Router R5 nie może nakazać routerowi R1 zmianę trasy ponieważ nie zna jego adresu.[1]

Prośby o zmianę trasy nie rozwiązują w ogólny sposób problemu przekazywania informacji o trasach, gdyż ograniczone są do operacji między routerem i węzłem połączonymi bezpośrednio jedną siecią. Rysunek nr 7 ilustruje to ograniczenie. „Przyjmijmy, że nadawca S z tego rysunku wysyła datagram do odbiorcy D. Przyjmijmy też, że router R1 nieprawidłowo wyznacza trasę tego datagramu w kierunku R2 zamiast R4 (tzn. R1nieprawidłowo wybiera ścieżkę dłuższą niż to konieczne). Gdy router R5, odbiera datagram, nie może wysłać komunikatu „zmień trasowanie" do R1 gdyż nie zna adresu tego rutera."[1]

Każdy komunikat „zmień trasowanie" zawiera 32-bitowe pole INTERSIECIOWY ADRES RUTERA oraz pole NAGŁÓWEK INTERSIECIOWY, jak na rys. 7.1. Pole INTERSIECIOWY ADRES RUTERA zawiera adres routera, do którego węzeł kieruje pakiety przeznaczone dla adresata wymienionego w nagłówku datagramu. NAGŁÓWEK zawiera nagłówek IP oraz 64 bity datagramu, który spowodował wysłanie komunikatu. Dzięki tym informacjom węzeł otrzymujący komunikat „zmień trasowanie" określa adres docelowy tego datagramu. Pole KOD komunikatu „zmień trasowanie" określa interpretacje adresu odbiorcy.



Rys 7.1. Format komunikatu „zmień trasowanie”

Wartość w polu KOD	**Znaczenie**
0	**zmień trasowanie datagramów do sieci (obecnie przestarzałe)**
1	**zmień trasowanie datagramów do węzła**
2	**zmień trasowanie datagramów o danym typie obsługi do sieci**
3	**zmień trasowanie datagramów o danym typie obsługi do węzła**

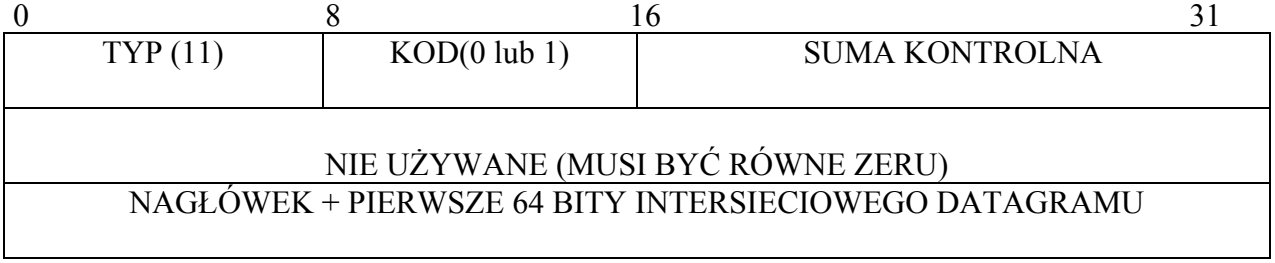
Generalnie routery wysyłają prośby ICMP o zmianę trasowania tylko do węzłów - nie wysyłają ich do routerów.

8. Wykrywanie zakleszczonych i zbyt długich tras

Routery intersieci obliczają adres następnego etapu przy użyciu własnych tablic. Błędy w tych tablicach mogą spowodować pętle w trasowaniu do jakiegoś odbiorcy D. Pętle występują gdy router, zamiast kierować datagramy do D kierują je do siebie nawzajem. Datagram, który wszedł do pętli w trasowaniu, porusza się po niej w nieskończoność. Aby temu zapobiec każdy datagram IP zawiera licznik czasu życia, jest to pole o długości 13 bitów, które wskazuje w sekundach czas, przez jaki datagram pozostanie w sieci zanim zostanie odrzucony. Ilekroć dany datagram przechodzi przez router czas istnienia (TTL) zostaje zmniejszony o co najmniej jedną sekundę. Ponieważ router normalnie przekazuje schemat IP w czasie krótszym niż jedna sekunda, ustawienie (TTL) staje się liczbą przeskoków (etapów).

„Gdy router porzuca datagram w wyniku wyczerpania się licznika etapów albo z powodu przekroczenia czasu oczekiwania na jego fragmenty, wysyła do nadawcy komunikat ICMP

<przekroczenie czasu>”[1] o formacie przedstawionym na rys. 8



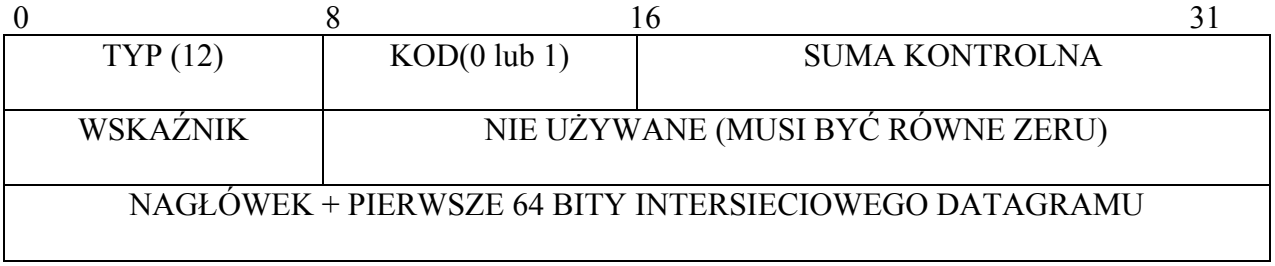
Rys 8. Format komunikatu ICMP „przekroczenie czasu".

W polu KOD zawiera informacje, o jakie przekroczenie czasu chodzi:

Wartość pola KOD	**Znaczenie**
0	**licznik czasu życia wyczerpany**
1	**przekroczony czas na składanie fragmentów**

9. Błąd parametrów

Gdy pojawi się błąd nieobsługiwany przez dotychczas opisane komunikaty ICMP o błędach router kieruje do nadawcy komunikat „błąd prametrów". „Jedna z możliwych przyczyn tego typu błędów ujawnia się, gdy argumenty opcji nie są poprawne."[1] Komunikat taki, w formacie jak na rys. 8, jest wysyłany tylko wtedy, kiedy błąd jest na tyle poważy, że datagram musi zostać porzucony. Jednoznaczność komunikatu, zapewnia pole o nazwie WSKAŹNIK służące do identyfikacji w datagramie oktetu, który spowodował kłopoty. Kod l jest używany do powiadamiania, że brakuje jakiejś opcji. Pole WSKAŹNIK nie jest używane, jeżeli wartość pola KOD jest równa l.



Rys 9. Format komunikatu „błąd parametrów"

10. Szacowanie czasu przesyłania pakietów

Zestaw protokołów TCP/IP zawiera wiele protokołów, które mogą być używane do synchronizowania zegarów. Jedna najprostszych metod polega na wykorzystaniu komunikatu ICMP do uzyskania informacji o czasie na innej maszynie. Maszyna, która potrzebuje danych o czasie, wysyła do innego komputera komunikat ICMP „prośba o czas", w którym prosi go o odesłanie bieżącego czasu. Maszyna otrzymująca taki komunikat odsyła maszynie proszącej „odpowiedź z czasem". Rys. nr 10 przedstawia format komunikatów „prośba o czas" i „odpowiedź z czasem".

Rys 10. Format komunikatu „prośba o czas” i „odpowiedź z czasem”.

Pole TYP określa rodzaj komunikatu: prośba (13) odpowiedź (14). Dzięki polom IDENTYFIKATOR oraz NUMER KOLEJNY nadawca może przyporządkować odpowiedzi prośbom. Pozostałe pola zawierają czasy wyrażone w milisekundach od północy czasu uniwersalnego. Pole CZAS POCZĄTKOWY wypełnia nadawca tuż przed wysłaniem prośby, pole CZAS OTRZYMANIA jest wypełniane przez odbiorcę zaraz po odebraniu komunikatu a pole CZAS ODESŁANIA jest wypełniane przed wysłaniem odpowiedzi.
Używane są trzy pola do szacowania czasu potrzebnego na przesyłanie datagramu między węzłami oraz do synchronizacji zegarów. Dzięki informacji zawartej w polu CZAS POCZĄTKOWY, węzeł może obliczyć czas potrzebny na dotarcie prośby do odbiorcy, przetworzenie jej na odpowiedź i powrót. Dodatkowo, odpowiedź zawiera zapis czasu, w którym prośba dotarta, jak i zapis czasu wysłania odpowiedzi, dzięki temu węzeł może obliczyć czas przesyłania przez sieć i za jego pomocą ocenić różnicę czasu między swoim zegarem a zegarem

na innej maszynie.

Dokładna ocena czasu przesyłu datagramu między węzłami jest trudna i znacznie ogranicza użyteczność komunikatów ICMP. Aby otrzymać dokładne oszacowanie czasu przesyłu, należałoby zrobić kilka pomiarów i je uśrednić. W dużej sieci otrzymane czasy różnią się, nawet jeżeli mierzone są w krótkich odstępach. Pamiętajmy, że IP jest protokołem zawodnym - datagramy mogą zostać zgubione, dostarczone z błędami lub dużym opóźnieniem. Zwielokrotnienie liczby pomiarów nie gwarantuje zatem poprawności wyników.

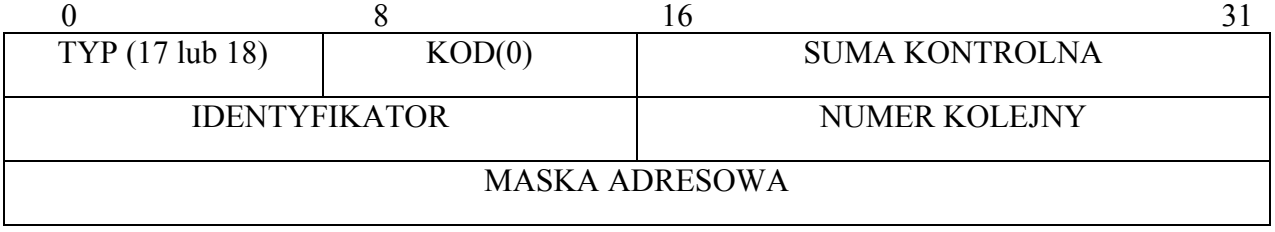
11. Komunikaty „prośba o informację" i „odpowiedź z informacją"

Komunikaty ICMP „prośba o informację" i „odpowiedź z informacją" (typ 15 i 16) wprowadzone, aby umożliwić węzłom dowiadywanie się o swoim adresie w sieci, zostały zastąpione przez protokoły: RARP i BOOTP i nie są obecnie używane.

12. Maska podsieci

Maska podsieci, umożliwia interpretację które bity 32-bitowego adresu sieciowego odpowiadają sieci fizycznej, a które odpowiadają identyfikatorowi węzła.

W celu uzyskania informacji o masce podsieci danej sieci fizycznej węzeł albo wysyła do routera (jeśli zna jego adres) komunikat „prośba o maskę adresową" albo rozgłasza go rys. 12 prezentuje format komunikatów związanych z maskami adresowymi.



Rys 12. Format komunikatu „ prośba o maskę adresową” i „odpowiedź z maską adresową”.

Pole TYP w komunikacie maski adresowej służy do określenia, czy jest to prośba (17), czy odpowiedź (18).

Literatura

[1] D. E. Comer „Sieci komputerowe TCP/IP zasady, protokoły i architektura” - WNT Warszawa 1997.
[2] J. Postel „RFC 792“ 1981.