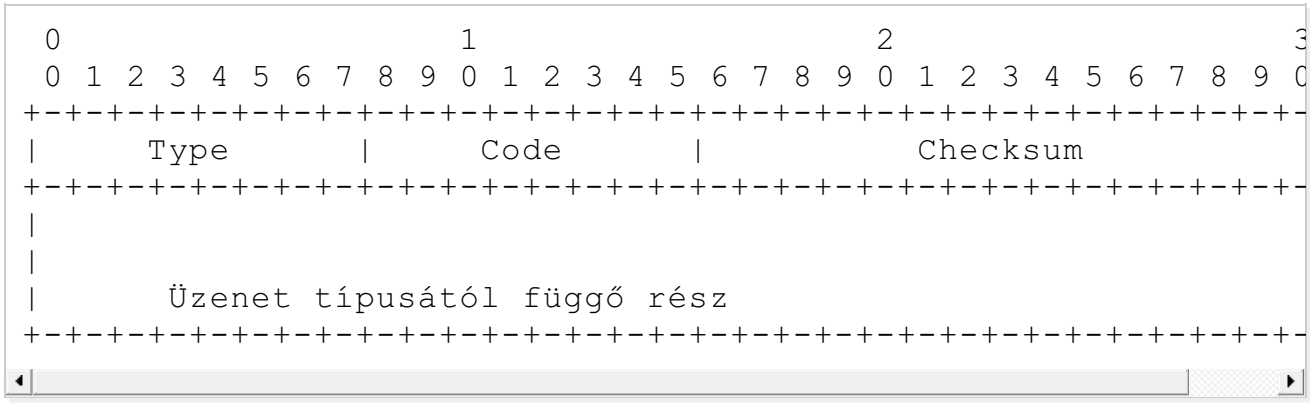
Számítógépes hálózatok — 6. előadás

ICMP, Internet Control Message Protocol

- ICMP - **RFC792**
- *To control* = *vezérel*
- STD 5 (IP-vel együtt)
  - Egy-egy RFC, vagy RFC-k egy halmaza lehet **internet standard**, (STD)
  - A szabvánnyá válás folyamatáról szól az **RFC2026**
  - Feltétel többek közt: két különböző, együttműködő implementáció
  - Az **STD-k** listája
- "IP companion protocol"
- Bizonyos szempontból az IP fölött levő réteg
  - IP csomagokat használ
  - IP protocol mező: `1`
- Bizonyos szempontból az IP alatti réteg
  - Az IP viselkedését (is) befolyásolja
    - Hibaüzenetek
    - Szolgálati üzenetek: routing, netmask stb.

ICMP formátum



- Type: az elsődleges információ, az üzenet típusát határozza meg
- Code: bizonyos üzeneteknél az üzenet altipusa
- Checksum: egyes komplemens összeadás, mint az IP-nél, az egész icmp üzenetre, 16 bites darabokban

ICMP üzenet-típusok

ICMP üzenet-típusok. Zárójelben a Stevens könyv fejezeteire való utalás

Type	Code	Description
0	0	*echo reply (Ping reply. Chapter 7)*
3		*destination unreachable:*
	0	network unreachable (Section 9.3)
	1	host unreachable (Section 9.3)
	2	protocol unreachable

	3	port unreachable (Section 6.5)
	4	fragmentation needed but don't-fragment bit se
	5	source route failed (Section 8.5)
	6	destination network unknown
	7	destination host unknown
	8	source host isolated (obsolete)
	9	destination network administratively prohibite
	10	destination host administratively prohibited
	11	network unreachable for TOS (Section 9.3)
	12	host unreachable for TOS (Section 9.3)
	13	communication administratively prohibited by f
	14	host precedence violation
	15	precedence cutoff in effect
4	0	*source quench (elementary flow control. Secti
5		*redirect (Section 9.5):*
	0	redirect for network
	1	redirect for host
	2	redirect for type-of-service and network
	3	redirect for type-of-service and host
8	0	*echo request (Ping request. Chapter 7)*
9	0	*router advertisement (Section 9.6)*
10	0	*router solicitation (Section 9.6)*
11		*time exceeded:*
	0	time-to-live equals 0 during transit (Tracerou
	1	time-to-live equals 0 during reassembly (Secti
12		*parameter problem:*
	0	IP header bad (catchall error)
	1	required option missing
13	0	*timestamp request (Section 6.4)*
14	0	*timestamp reply (Section 6.4)*
15	0	*information request (obsolete)*
16	0	*information reply (obsolete)*
17	0	*address mask request (Section 6.3)*
18	0	*address mask reply (Section 6.3)*

- Két fő csoport
  - Hibaüzenetek
  - Vezérlő, konfiguráló üzenetek

A hibaüzenetekre nagyon szigorú szabályok vonatkoznak

Sose eredményezhet hibaüzenetet:

- ICMP hibaüzenet
- IP broadcast, vagy multicast
- Alacsonyabb (link layer) broadcast, vagy multicast
- Egy IP csomag többedik (nem első) fragmentuma
- Olyan IP csomag, aminek forráscíme nem egy host IP címe
- IGMP (Internet Group Management) üzenetek

A hibaüzenet mindig tartalmazza a kiváltó IP csomag lényeges részét

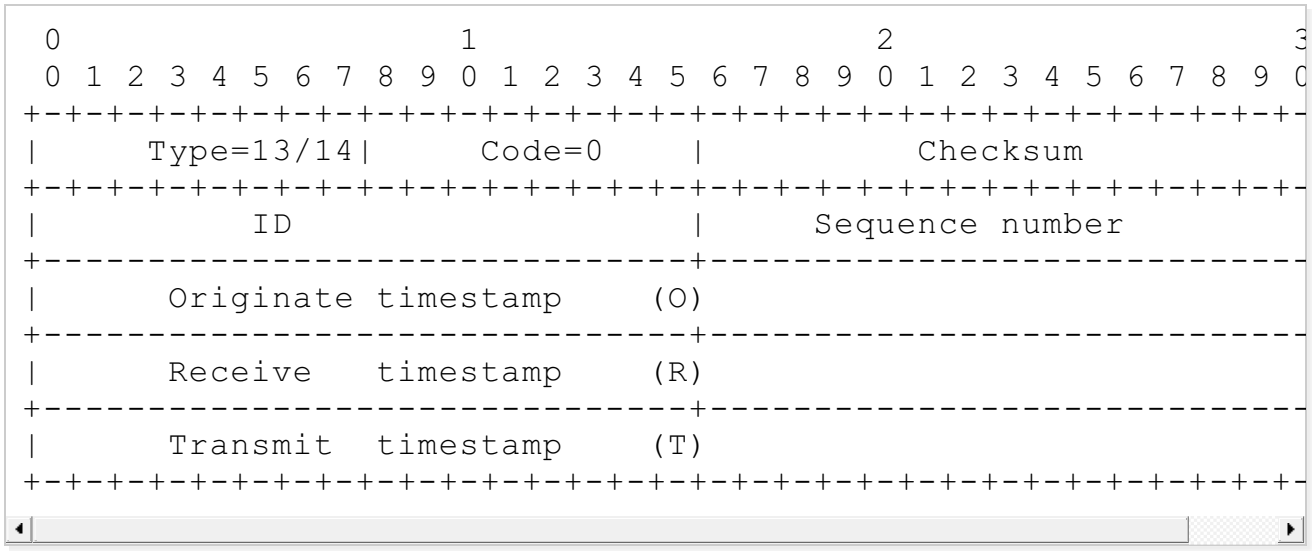
- A teljes IP fejrészet (20-60 byte)
- Az első 8 byte-ját az IP adat résznek
- TCP és UDP esetén ez tartalmazza a portokat
  - A vevő ebből tudja, hogy melyik programot érinti a hiba

ICMP address mask request/reply

- Request típus: `17`, broadcast
- Reply típus: `18`, unicast
- RARP-pal kapcsolatban használatos
- Több válasz is érkezhet
  - A RARP csak egy puszta címet ad
  - Nekem is az legyen a netmaskom, ami a többinek
- Kiment a divatból
  - DHCP betölti ezt a funkciót is

ICMP Timestamp request/reply

- Request típus: `13`
- Reply típus: `14`
- Mindkettő unicast



- Időegység: az UTC éjfél (Coordinated Universal Time) óta eltelt milliszekundumok
  - Ha a felső bit `0`
  - Ha a felső bit `1`, akkor más idő is lehet - implementáció függő
- A gyakorlatban - ma már - a receive és a transmit timestamp egyezik
- Óra beállításra alkalmas
  - Legyen *RTT* a request küldés és a reply fogadás közti idő

- Ha
  - egyformán jár az óránk, és
  - egyforma a késleltetés oda-vissza,
  - akkor *O+RTT/2=R*
  - Ha *R* ennél nagyobb, akkor a mi óránk késik, ha kisebb, akkor siet
- Alternatívák:
  - `13` TCP/UDP port: daytime. A helyi időt mutatja olvasható formában
  - Ravaszabb órabeállító protokoll: NTP (Network Time Protocol)

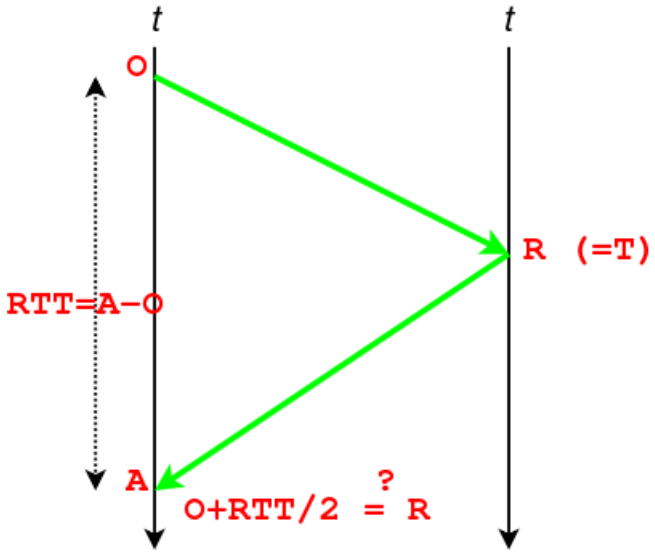


Figure 1: ICMP timestamp request/reply

NTP

- NTP: **RFC5905**, **http://www.ntp.org**
  - UDP `123`-as porton működik
  - A nyilvános interneten is ezredmásodperces pontosságúak lehetnek eszközeinkben a rendszerórák
  - Pl. szökőmásodperceket azonnal érvényesít
- Az NTP segítségével nem egymáshoz szinkronizáljuk az órákat, hanem referencia órához (reference clock) állítjuk
- A konfigurációban több NTP szervert adhatunk meg, mindegyiket kérdezgetjük, aztán választunk referenciát
- **Stratum:** az atomórával vagy GPS-sel közvetlen kapcsolatban levő ntp szerver stratum `1`
  - Ha egy NTP szerver stratum `n` szerverhez szinkronizál, akkor ő stratum `n+1`
- Az óra beállítása célszerűen nem pillanatszerű: az `adjtime` rendszerhívás fokozatosan szinkronizálja az órát
  - Így nem lesz kimaradás, sose jár hátrafele az óra.
- Publikus ntp szerverek listája: **http://psp2.ntp.org/bin/view/Servers/WebHome**
- Magyarországon népszerű: time.kfki.hu

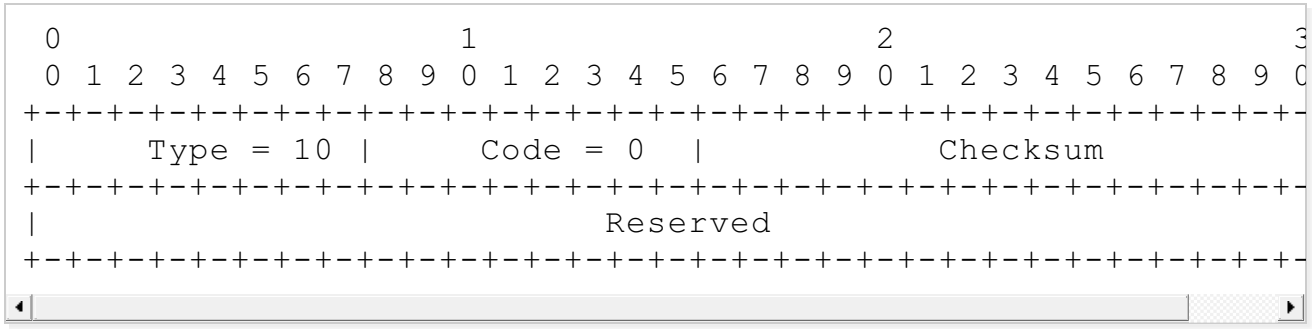
A pool.ntp.org projekt

- Önkéntesek által üzemeltetett **projekt**
- Sok ezer NTP szerver clustere
- A `pool.ntp.org` név véletlenszerűen választ a név alapján publikus ntp szerver(eke)t
- Egy linuxban ha települ az NTP csomag, ez az alapértelmezett „szerver”, amihez szinkronizál
- A pool-hoz bárki csatlakozhat, része lehet a clusternek

Router advertisement/router solicitation

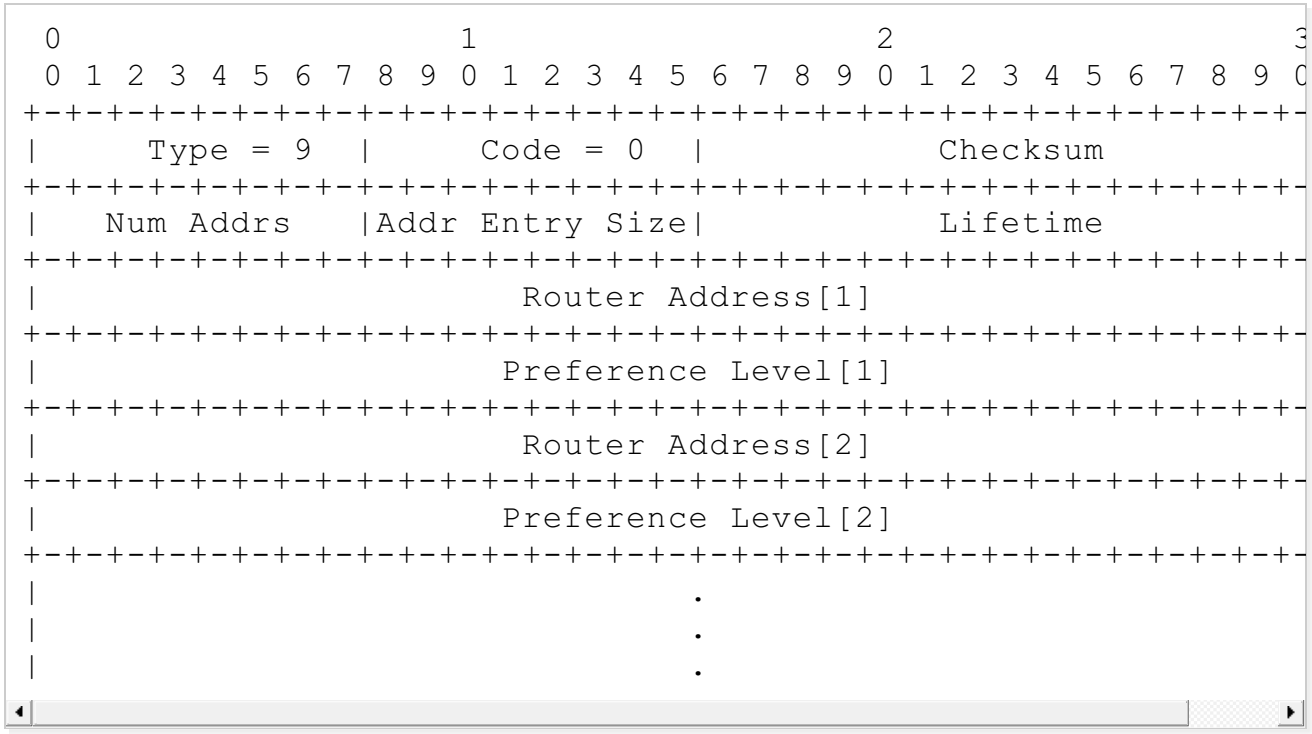
- **RFC1256**
- Dinamikusan, ICMP üzenetek által állít route-okat

Router Solicitation



- Type = `10`
- Multicast: `224.0.0.2` = all routers
- A routerek unicast-tal válaszolnak: router advertisement ICMP csomaggal

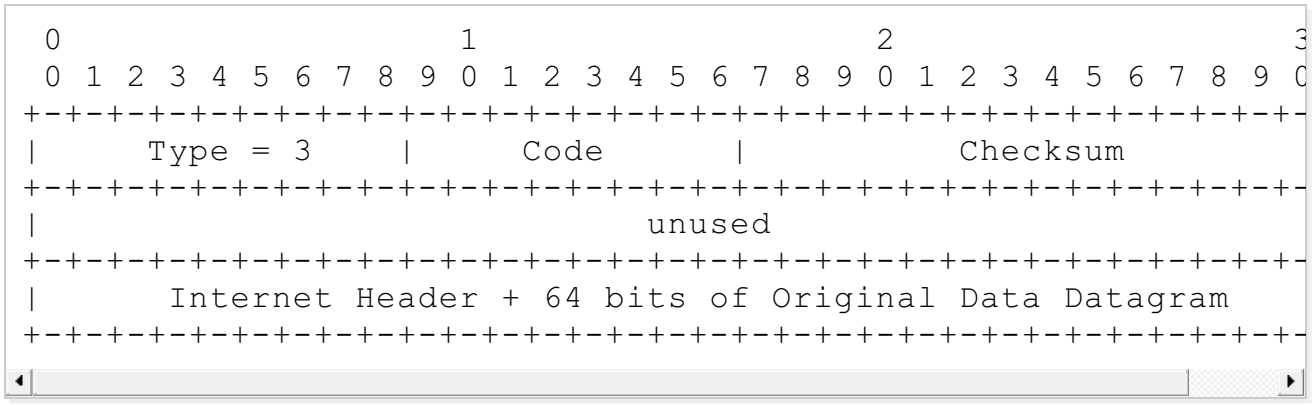
Router Advertisement



- A default router címét hirdeti
- Num Addrs: ennyi router címet hirdetek
- Addr entry size: ennyi 4 byte-os érték egy entry (=2)

- Lifetime: ennyi másodpercig érvényes ez a hirdetés
- Router Address: a router IP címe
- Preference Level: előjeles szám, minél nagyobb, annál jobban preferáld
- Nem csak solicite-ra válaszul, unicasttal, hanem multicasttal is
  - 8-10 percenként, véletlent belekeverve küldik
  - A `224.0.0.1` címre = all hosts
- Nem lehet hálózat/router vagy host/router hozzárendelést megadni

Destination unreachable

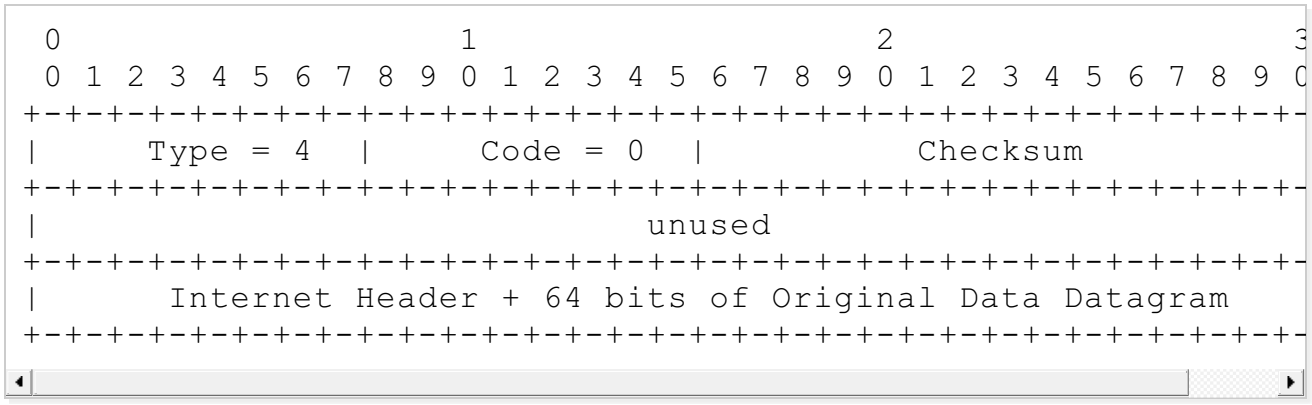


- Gyakori hibaüzenet
- Küldheti a címzett, vagy egy közbülső router
- Tűzfalak is küldhetik előbb, akár a címzettet mímelve
- A code mező mutatja a finomabb okot
  - Network unreachable: router küldi, ha zavarba jön, nem találja a címzettet
  - Host unreachable: az utolsó router küldi, aki úgy érzi, látnia kéne, de ilyen nincs
  - Protocol unreachable: többnyire nem UDP/TCP-vel kapcsolatos hiba
  - Port Unreachable
    - Jellemzően a címzettől jön
    - TCP-re nem jellemző
      - Ott RESET-tel bomlik a kapcsolat
  - Fragmentation needed but DF set = Túl nagy csomag
    - Egy közbülső router küldi
    - A következő MTU kisebb mint a csomag
    - A csomagban kérték, hogy DF: Don't Fragment
    - A második 4 byte-os szóban elküldheti a bajt okozó MTU-t
      - Ez eredetileg használatlan
      - **RFC1191** vezette be

Path MTU discovery

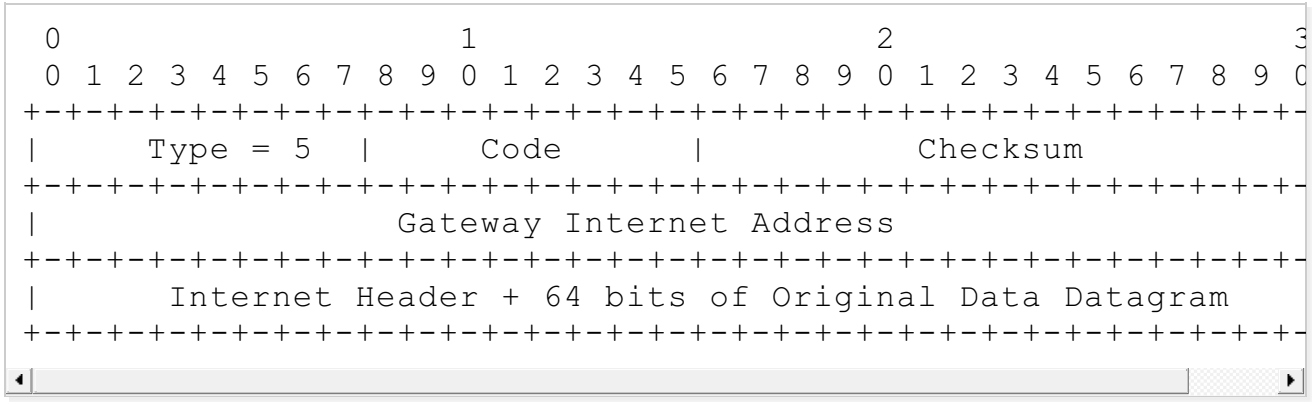
- Különösen TCP kapcsolatoknál fontos
- Lehetőleg nagy csomagokat akarunk küldeni
- De nem akarjuk, hogy közben fragmentumokra szedjék a csomagokat
- A cél címig mekkora a legkisebb MTU?
  - Időben akár változhat is !
- **RFC1191**
  - A küldőnek van egy változója a célcím-hez tartozó MTU-ról
  - Kezdetben a célcím route-jához tartozó MTU
  - Csökkenti, ha „túl nagy csomag” üzenetet kap

Source quench - Lassíts!



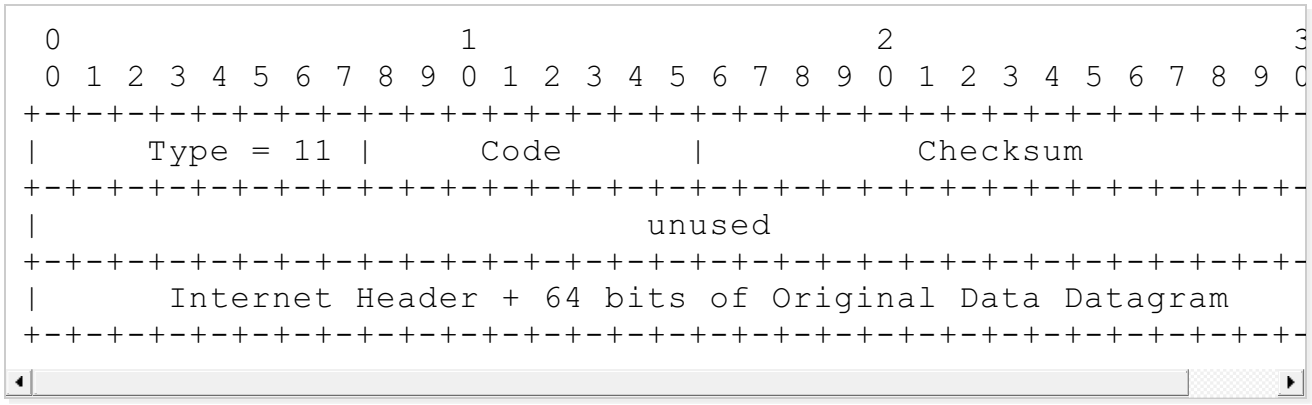
- Torlódás kezelés - congestion control
  - Ősi, nyers eszköz a torlódáskezelésre - (különösen TCP-ben) kifinomultabb eszközök vannak
- Egy közbülső router, vagy a célállomás küldheti
- Ha eldobta a csomagot, mert nem tudta már továbbítani
- Ha már közel van ehhez az állapothoz
- A küldő állomás visszafogja magát
  - A visszaküldött csomagból látszik, hogy hol is!
  - Egy idő után dönthet úgy, hogy újra erőteljesebben küld
- Tűzfalak kiszűrhetik
  - hiba
  - furcsa jelenségeket okozhat
  - esetleg sokáig nem vesszük észre

Redirect - A célállomást rövidebb úton is eléred, ha erre keresed



- Router küldi
- Megjelöl egy másik routert, ami kedvezőbb
- Csak akkor, ha látja, hogy a küldő és a kedvezőbb router egy hálózaton van
- A küldő állomás módosítja a routing tábláját
- Visszaélésre ad módot
  - A default routeren kívül mástól nem szabad elfogadni
  - Lehet, hogy még a dafault router-től sem!
- Redirect Code
  - 0 = Redirect datagrams for the Network.
  - 1 = Redirect datagrams for the Host.
  - 2 = Redirect datagrams for the Type of Service and Network.
  - 3 = Redirect datagrams for the Type of Service and Host.

Time exceeded

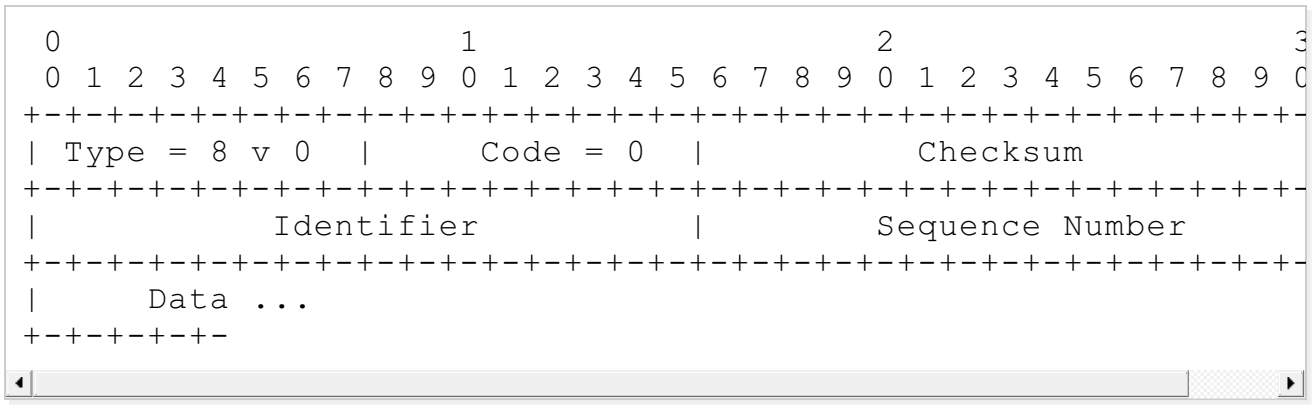


- Eldobtam a csomagod, mert mire ideért lejárt a TTL
  - Code = `0` : time to live exceeded in transit;
  - Code = `1` : fragment reassembly time exceeded.
    - Lejárt a timeout, és nem sikerült az egész (részekben küldött) datagram-ot összeállítani
    - Ha az első fragmentum nem érkezett meg, akkor nem!
- A visszaküldött csomagból látszik, hogy melyik kapcsolathoz tartozik
- Túlbuzgó tűzfalak ezt is kiszűrhetik - ez is nehezen kideríthető hibához vezethet

Tűzfalakon ajánlatos legalább is a következő ICMP üzeneteket átengedni

- Destination unreachable (type = `3`)
- Source quench (type = `4`)
- Time exceeded (type = `11`)

Echo request/reply - a ping program eszközei



- Type: request = `8`, reply = `0`
- ID: egy ping instanciát azonosít, (= processz ID)
- Sequence Number: egy instancián belül a sorszámot
  - Lehet, hogy más sorrendben kapjuk vissza az echot!
- Adat: a küldött adatot vissza kell kapjuk
- ping program
  - Klasszikus eszköz egy IP cím elérhetőségének vizsgálatára
  - `-c` kapcsoló (count): ennyi request-et küld
  - `-s` kapcsoló (size): ekkora adat részt küld (+8 byte ICMP fejrész)
  - `-f` kapcsoló (flood): gyorsan küld sokat egymás után
    - A küldött csomagokat pont (**.**), a vetteket backspace jelzi: mozi

        - Rendszergazdai jogok kellenek a használatához
    - Ha távoli gép nem elérhető, érdemes a default route-ot pingetni
    - `-R` kapcsoló (record route)
        - Az IP record route opcióját kapcsolja be
        - Mindkét irányban láthatjuk a közbülső routereket

IP record route opció

```
+--------+--------+--------+---------//--------+
|00000111| length | pointer|     route data    |
+--------+--------+--------+---------//--------+
  Type=7
```

- **length**: ennyi byte az opció
- a route data length-3 hosszú
- **pointer**: a következő IP cím helyét mutatja: először 4, legfeljebb 40
    - 40: tele az IP header
- Az adat 4 byte-os IP címekből áll

Traceroute

- Az IP record route opció legfeljebb 9 router címet tárol
- Nem mindenki engedi át
- 1, 2, 3,… TTL-lel küld UDP csomagokat
- Az i-edik menetben küldött csomagot az i-edik hop router utasítja el ICMP time exceeded üzenettel
- UDP `33434`-től egyre nagyobb portokra küld csomagokat
    - Ezzel azonosítja a választ
- Tűzfalak korlátozhatják
- Alternatívák
    - `traceroute -I`: ICMP csomagokat küld
    - `mtr` (my/Matt's traceroute): ICMP csomagok, látványos felület, mozi
    - `tcptraceroute`: TCP csomagok, a 80-as, vagy bármely más portra

Szerző: Pásztor Miklós, Máray Tamás