

PROTOCOALE DE COMUNICAȚIE : LABORATOR 1

Socketi UDP

Responsabil: Cătălin LEORDEANU

Cuprins

Obiective	1
Noțiuni teoretice	2
Nivelul transport	2
Porturi	2
UDP	2
Header UDP	2
Calcul Checksum	3
API	3
Network byte-order	3
Socket	4
Structuri de date	4
struct sockaddr	4
struct sockaddr_in	4
struct in_addr	5
Utilizarea adreselor IPv4	5
inet_addr()	5
inet_aton()	5
inet_ntoa()	6
Funcții socketi UDP	6
socket()	6
bind()	7
close() / shutdown()	8
recvfrom()/sendto()	8
Apeluri extra	9
sendmsg()/recvmsg()	9
gethostbyname()/ gethostbyaddr()	9
Exerciții	10
Comenzi utile	11

Obiective

În urma parcurgerii acestui laborator studentul va fi capabil să:

- explice în ce constă protocolul de transport UDP
- scrie un program care folosește socketi UDP

Noțiuni teoretice

Nivelul transport

Nivelul transport oferă multiplexare la nivel de aplicații. Astfel 2 aplicații aflate pe două mașini diferite pot comunica prin intermediul acestui layer.

Porturi

Multiplexarea la nivelul transport se asigură prin porturi. Acestea sunt reprezentate pe 2 octeți, așadar numerele care sunt asignate porturilor sunt de la 0 la 65535, însă cele până la 1024 sunt rezervate pentru aplicații standard, precum:

- FTP: 20,21
- SSH: 22
- Telnet: 23
- SMTP: 25
- HTTP: 80
- POP3: 110
- IMAP: 143

O listă completă a porturilor rezervate găsiți la IANA: <http://www.iana.org/assignments/port-numbers>

UDP

UDP (User Datagram Protocol) este un protocol ce trimite pachete independente de date, numite datagrame, de la un calculator către altul fără a garanta în vreun fel ajungerea acestora la destinație. Acest protocol nu stabilește o conexiune permanentă între cele două calculatoare. Este descris în RFC 768 (<http://tools.ietf.org/html/rfc768>)

Protocolul UDP are următoarele proprietăți:

- nu se stabilește o conexiune între client-server. Așadar, serverul nu va aștepta apeluri de conexiune, ci așteaptă direct datagrame de la clienți. Acest tip de comunicare este întâlnit în sistemele client-server în care se transmit puține mesaje și în general prea rar pentru a menține activă o conexiune între cele două entități. (Un exemplu în acest caz este DNS-ul)
- nu se garantează ordinea primirii mesajelor și nici prevenirea pierderilor pachetelor. UDP-ul se utilizează mai ales în rețelele în care există o pierdere foarte mică de pachete și în cadrul aplicațiilor pentru care pierderea unui pachet nu este foarte gravă (Un exemplu: aplicațiile peer-to-peer din cadrul unei rețele locale).
- are un overhead foarte mic, în comparație cu celelalte protocoale de transport (Are un header de 8 Bytes, în comparație cu TCP-ul care are minim 20Bytes).

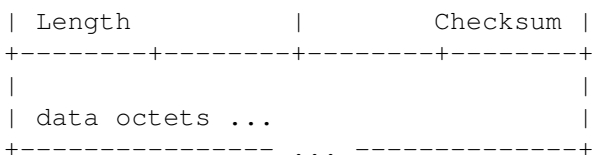
Header UDP

Orice pachet UDP are următorul header:

```

0          7 8          15 16          23 24          31
+-----+-----+-----+-----+
|                Source | Destination          |
|                Port  | Port                  |
+-----+-----+-----+-----+
|                |                |

```

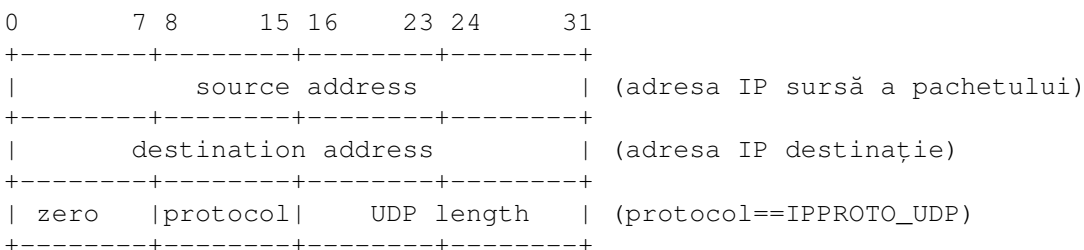


Explicații header:

- Portul sursă este ales random de către mașina sursă a pachetului dintre porturile libere existente pe acea mașină.
- Portul destinație este portul pe care mașina destinație poate recepționa pachete.
- Lungime este lungimea în octeți a datagramă (header+date).
- Checksum este valoarea sumei de verificare pentru datagramă. În secțiunea următoare, este prezentat modul de calcul a checksum-ului pentru un pachet UDP.

Calcul Checksum

Pentru a calcula acest număr header-ului UDP i se adaugă un *pseudo-header* format din :



Pe acest pseudo-header, împreună cu datagram UDP (header+date) se calculează suma în următorul mod:

- Se formează cuvinte de 16 biți. Dacă numărul de octeți al datelor este impar, atunci se mai adaugă la sfârșit un octet 0.
- Se calculează suma complementelor față de 1 al tuturor cuvintelor.
- Se calculează complementul față de 1 al sumei de la pasul anterior.

Checksum-ul datagramă este numărul returnat de ultimul pas.

API

Network byte-order

Un sistem de operare poate ordona octeții într-un cuvânt în următoarele variantele:

- Big Endian - cel mai semnificativ octet primul
- Little Endian - cel mai puțin semnificativ octet primul

Datorită acestei împărțiri există o problemă dacă vrem să interconectăm două sisteme de operare diferite din punct de vedere al reprezentării datelor. În acest sens, Internet-ul a impus o secvență standard pentru stocarea datelor numerice, numită **network byte-order**. Spre deosebire, secvența octeților pentru calculatorul gazdă se numește **host byte-order**.

Indiferent de ordonarea octeților pe o mașină, în momentul în care se trimit date în rețea acestea trebuie să fie în formatul Network Byte Order.

Pentru conversie există funcții pentru două tipuri: short (2 octeți) și long (4 octeți). Aceste funcții sunt valabile și pentru variantele unsigned.

```
1 #include <arpa/inet.h>
2
3 uint32_t htonl(uint32_t hostlong);
4 uint16_t htons(uint16_t hostshort);
5 uint32_t ntohl(uint32_t netlong);
6 uint16_t ntohs(uint16_t netshort);
```

Socket

Un "socket" reprezintă un canal generalizat de comunicare între procese, reprezentat în Unix printr-un descriptor de fișier. El oferă posibilitatea de comunicare între procese aflate pe mașini diferite într-o rețea.

Structuri de date

struct sockaddr

În structura struct sockaddr definită în <sys/socket.h> se țin informațiile referitoare la adresa socket-urilor:

```
1 #include <sys/socket.h>
2
3 struct sockaddr
4 {
5     unsigned short sa_family;
6     char sa_data[14];
7 }
```

Explicații:

- sa_family - indica un format particular de adresa. Valori uzuale: **AF_INET** (protocol IPv4)
- sa_data - adresa utilizată

struct sockaddr_in

În cazul comunicației prin Internet, toate adresele socket-urilor sunt compuse dintr-un număr de port și adresa IP a calculatorului respectiv. Pentru specificarea acestor adrese se utilizează în locul structurii sockaddr, structura sockaddr_in ("in" de la Internet):

```
1 #include <netinet/net.h>
2 struct sockaddr_in
3 {
4     unsigned short sin_family;
5     unsigned short int sin_port;
6     struct in_addr sin_addr;
7 }
```

Explicații:

- sin_family - aceeași semnificație ca sa_family. Valoare constantă AF_INET
- sin_port - portul utilizat (Network Byte Order)
- sin_addr - adresa IP (Network Byte Order)

struct in_addr

Reține o adresă IPv4, care are 32 biți

```
1 struct in_addr
2 {
3     uint32_t s_addr;
4 }
```

Pentru această structură există câteva constante speciale:

- INADDR_LOOPBACK - Reprezintă adresa mașinii, '127.0.0.1', care mai este numită și 'localhost'. Această constantă poate fi folosită în loc de a obține adresa propriului calculator.
- INADDR_ANY - Reprezintă orice adresă și este folosită ca valoare pentru sin_addr. atunci când se dorește să fie acceptate conexiuni.
- INADDR_BROADCAST - Adresa folosită pentru broadcast.
- INADDR_NONE - Adresa returnată de unele funcții pentru a indica eroare.

Utilizarea adreselor IPv4

Funcțiile pentru manipularea adreselor Internet sunt definite în headerul <arpa/inet.h> și sunt explicate în secțiunea 3 a man-ului. (man 3 inet_addr)

inet_addr()

Convertește o adresă IPv4 din formatul standard (exemplu 10.2.5.160) în forma binară.

```
1 #include <sys/socket.h>
2 #include <netinet/in.h>
3 #include <arpa/inet.h>
4 [...]
5 struct sockaddr_in my_addr;
6 [...]
7 my_addr.sin_addr.s_addr = inet_addr("10.2.5.160");
8 [...]
```

Problema care apare la folosirea inet_addr() este ca în cazul în care apare o eroare se întoarce INADDR_NONE care este -1 și reprezintă adresa "255.255.255.255", care este o adresă validă.

inet_aton()

De aceea, este recomandat folosirea funcției inet_aton() (ASCII to network):

```
1 int inet_aton (const char *NAME, struct in_addr *ADDR);
```

Rezultatul întors este diferit de zero dacă adresa este validă și zero în caz de eroare

```
1 inet_aton("10.2.5.160", &(my_addr.sin_addr));
```

inet_ntoa()

Dacă se dorește ca dintr-o structură `in_addr` să se afișeze adresa IPv4 în formatul standard se folosește funcția inversă `inet_ntoa()` (network to ASCII):

```
1 char * inet_ntoa (struct in_addr ADDR);
```

Atenție! Rezultatul este întors într-un pointer alocat static, astfel că apeluri repetate ale acestei funcții vor scrie în același buffer, de aceea este bine ca rezultatul să fie copiat dacă se dorește a fi păstrat.

Funcții socketi UDP

În cele ce urmează se vor prezenta principalele funcții pentru manipularea socketilor.

Schema pentru o comunicare tipică client-server este:

```
Server: socket(...) -> bind(...) -> recvfrom(...) -> recvfrom(...) / sendto(...) ...  
Client: socket(...) -> [bind(...)] ---> sendto(...) -> recvfrom(...) / sendto(...) ...
```

Atât serverul cât și clientul apelează `socket()`. Serverul este cel care așteaptă date, prin `recvfrom()`. Clientul trimite o cerere la server prin `sendto()`. Ulterior, schimburile între cele două procese sunt în funcție de natura aplicației.

Clientul poate face opțional `bind`, dacă vrea să specifice o interfață pe care pleacă pachetele sau un port sursă, dar de cele mai multe ori acest aspect cade în sarcina sistemului de operare.

socket()

Pentru obținerea descriptorului de fișier se folosește funcția:

```
1 #include <sys/types.h>  
2 #include <sys/socket.h>  
3 [...]  
4 int socket(int domain, int type, int protocol);  
5 [...]
```

Explicații:

- `domain` - reprezintă familia protocoalelor pe care urmează să le utilizăm în transferul informației. Valori uzuale :
 - PF_INET (IPv4)
 - PF_INET6 (IPv6)
- `type` - tipul socketului. Valori uzuale:
 - SOCK_STREAM - Indică stabilirea unei comunicații bazată pe construirea unei conexiuni între sursă și destinație. Comunicația este FIFO, fiabilă și sigură.
 - SOCK_DGRAM - Oferă un flux de date bidirecțional dar care nu promite să fie sigur, în secvență sau neduplicat. Adică, un proces care recepționează mesaje pe un socket datagramă, poate găsi mesaje duplicate și posibil într-o ordine diferită față de cea în care au fost trimise.

- SOCK_RAW - Permite accesul utilizatorului la protocoalele de comunicație inferioare care suportă abstractizarea socketilor. Acești socketi sunt orientați datagramă, deși caracteristicile lor sunt dependente de interfața oferită de protocol. Socketii RAW nu sunt pentru utilizatorul general. Ele sunt oferite în mod deosebit celor interesați în dezvoltarea de noi protocoale de comunicație sau celor care doresc acces la niște facilități mai rar folosite ale unui protocol existent.
- protocol - specifică protocolul de transport utilizat. Fiecare tip de socketi are un protocol specific și de aceea valoarea câmpului este bine să fie 0 pentru ca să se aleagă protocolul corect în funcție de argumentul type specificat.
- link: man 2 socket <http://www.kernel.org/doc/man-pages/online/pages/man2/socket.2.html>

În cazul UDP se utilizează tipul SOCK_DGRAM.

Funcția întoarce descriptorul socket-ului sau -1 în caz de eroare setând corespunzător variabila `errno`.

bind()

Odata ce am obținut un socket, trebuie să îi asociem un port pe mașina locală (acest lucru este uzual în cazul în care se dorește așteptarea de conexiuni pe un anumit port).

```
1 #include <sys/types.h>
2 #include <sys/socket.h>
3
4 int bind(int sockfd, struct sockaddr *my_addr, int addrlen);
```

Explicații:

- `sockfd` - descriptorul de fișier returnat de funcția `socket()`
- `my_addr` - conține informații despre adresa IP și port
- `addrlen` - lungimea celui de-al doilea parametru poate fi setat ca fiind `sizeof(struct sockaddr)`
- link-man 2 bind <http://www.kernel.org/doc/man-pages/online/pages/man2/bind.2.html>

În urma apelului, socketului `sockfd` i se asignează adresa din `my_addr`. În caz de succes returnează zero, iar în caz de eroare -1 și setează corespunzător `errno`.

Dacă în structura `my_addr`, portul este 0 (`my_addr.sin_port = 0`) atunci se alege orice port disponibil din cele nerezervate.

Observație: Câteodată, când se încearcă să se repornească un server, `bind()` nu reușește, iar eroarea este "Address already in use". Asta înseamnă că un socket care a fost conectat pe acel port încă mai este agățat și utilizează portul.

În această situație fie se poate aștepta să se deconecteze câteva minute sau se poate specifica în program să se reutilizeze portul:

```
1 [...]
2 int yes = 1;
3 if (setsockopt(listener, SOL_SOCKET, SO_REUSEADDR, &yes, sizeof(int)) == -1) {
4     perror("setsockopt");
5     exit(1);
6 }
7 [...]
```

O alternativă este să aflați ce proces utilizează acel port și să-l terminați manual. Comanda `lsof -i :1337` vă va întoarce informații despre procesul care ascultă pe portul 1337. Puteți apoi să folosiți comanda `kill` pentru a îl opri.

(Dacă știți că un port e folosit, dar `lsof` nu afișează informații despre niciun proces, cel mai probabil nu aveți drepturi suficiente; rulați `lsof` cu `sudo`.)

close() / shutdown()

Pentru a închide un socket se folosește funcția de închidere a unui descriptor de fișier din Unix:

```
1 #include <unistd.h>
2
3 int close(int fd);
```

Acest lucru va împiedica alte citiri și scrieri din socket. Orice încercare de citire/scriere din acest socket va întoarce eroare. În cazul în care se dorește mai mult control asupra socketului care urmează să fie închis, se poate folosi funcția `shutdown()` care permite întreruperea comunicației într-un sens sau în ambele direcții (caz asemănător cu `close()`).

```
1 #include <sys/socket.h>
2
3 int shutdown(int sockfd, int how);
```

Explicații:

- `sockfd` - descriptorul de fișier returnat de funcția `socket()`
- `how` - specifică modul de închidere. Valori luate:
 - 0 - Utilizatorul nu mai este interesat să citească datele.
 - 1 - Nu se mai pot face transmisiuni de date.
 - 2 - ambele (caz asemănător cu `close()`)
- link: `man 2 shutdown`

Observație: `shutdown()` nu închide de fapt un descriptor de fișier, ci doar îi schimbă modul de utilizare. Pentru a elibera un socket trebuie folosit `close()`.

recvfrom()/sendto()

Prin intermediul acestor apeluri specifice socketelor UDP, se pot trimite/recepționa date.

```
1 #include <sys/types.h>
2 #include <sys/socket.h>
3 int sendto(int sockfd, char* buff, int nbytes, int flags, struct sockaddr *to, int
  addrln);
4 int recvfrom(int sockfd, char* buff, int nbytes, int flags, struct sockaddr *from, int
  addrln);
```

Explicații:

- sockfd: socketul prin care se realizează comunicarea.
- buff: bufferul ce conține datele care se trimit/recepționează.
- flags: specifică condiții de efectuare a transmisiei/recepției.
- to/from: Structură ce indică adresa unde se trimit/de unde se primesc datele. (În cazul recvfrom() se populează de funcție).
- addrlen: Lungimea structurii to/from (în octeți).
- link: man 2 recvfrom - <http://www.kernel.org/doc/man-pages/online/pages/man2/recvfrom.2.html>

Dacă nu apar erori, apelurile întorc lungimea mesajului în octeți. La eroare, se întoarce -1.

Apeluri extra

sendmsg()/recvmsg()

Apelurile de sistem folosite în comunicația fără conexiune sunt sendmsg() și recvmsg():

```
1 #include <sys/types.h>
2 #include <sys/socket.h>
3
4 int sendmsg(int sockfd, struct msghdr* msg, int flags);
5 int recvmsg(int sockfd, struct msghdr* msg, int flags);
```

Explicații:

- sockfd - identificatorul socketului
- msg - structura care se va trimite/recepționa

```
1      #include <sys/types.h>
2      struct msghdr {
3          void *msg_name; /* optional address */
4          socklen_t msg_namelen; /* size of address */
5          struct iovec *msg_iov; /* scatter/gather array */
6          size_t msg_iovlen; /* # elements in msg_iov */
7          void *msg_control; /* ancillary data, see below */
8          size_t msg_controllen; /* ancillary data buffer len */
9          int msg_flags; /* flags on received message */
10     };
11
```

- flags - aceeași semnificație ca la sendto()/recvfrom()
- link - man 2 sendmsg

gethostbyname()/ gethostbyaddr()

Funcțiile prezentate până acum se bazează pe faptul că știm adresa IP a entității la care ne conectăm. De cele mai multe ori, nu cunoaștem această adresă, ci doar numele sub care ea este recunoscută în rețea. (Este mult mai ușor de ținut minte un nume decât un număr, de aceea s-a și inventat și protocolul DNS).

Funcțiile prezentate în această secțiune rezolvă acest neajuns, returnând adresele și numele sub care este recunoscută o mașină în rețea.

```
1 #include <netdb.h>
2 #include <sys/socket.h> /* for AF_INET */
3
4 extern int h_errno;
5 struct hostent *gethostbyname(const char *name);
6 struct hostent *gethostbyaddr(const void *addr, socklen_t len, int type);
```

Explicații:

- name - numele host-ului despre care se cer informații
- addr - un pointer la un buffer ce conține adresa, dată în formatul specific familiei de adrese din care face parte.
- len - lungimea bufferului anterior.
- type - familia de adrese (AF_INET pentru adrese de tip internet).
- link - man 3 gethostbyname: <http://www.kernel.org/doc/man-pages/online/pages/man3/gethostbyname.3.html>

Structura hostent este definită astfel:

```
1 #include <netdb.h>
2
3 struct hostent
4 {
5     char *h_name;
6     char **h_aliases;
7     int h_addrtype;
8     int h_length;
9     char **h_addr_list;
10 }
```

Explicații:

- h_name: Reprezintă numele oficial al mașinii respective,
- h_aliases: Reprezintă alte nume sub care ea este cunoscută (alias-uri)
- h_addrtype: Reprezintă tipul adreselor acelei mașini (pentru adrese IP conține valoarea constantei AF_INET).
- h_length: Indică lungimea binară a tipului de adresă utilizat (în cazul adreselor IP această lungime este de 4 octeți).
- h_addr_list: Reprezintă un vector de adrese sub care aceasta mașină e recunoscută în rețea. (fiecare adresă are h_length octeți).

Exerciții

Pe baza scheletului de cod atașat laboratorului, veți implementa un mini server de backup fișiere și un client al acestui server.

Detalii implementare:

- Vă veți grupa în echipe de 2 persoane, una implementând serverul și cealaltă clientul; (la calculatoare diferite)
- Clientul va trimite un fișier binar prin UDP, iar serverul îl va recepționa.
- Nu trimiteți numele fișierului, ci primiți-l ca argument în cele 2 fișiere.
- Presupuneți că clientul va transmite următorul calup de date disponibil din fișier (setați dimensiunea calupului de date maxim la MTU-ul rețelei (în cazul practic 1500 octeți))

- Serverul va recepționa datele și le va scrie în fișierul lui.
- La final rulați md5sum pe cele 2 fișiere pentru a vedea dacă au fost recepționate corect.
- Pentru compilare & rulare & testare verificați fișierul Makefile.

Bonus: Implementați mecanismul de preluare a numelui fișierului la server și permiteți recepționarea mai multor fișiere consecutiv.

Comenzi utile

Următoarele comenzi de shell vă pot fi utile în rezolvarea laboratorului:

- netstat

```
1 $ netstat -ulpn
2 (Not all processes could be identified, non-owned process info
3 will not be shown, you would have to be root to see it all.)
4 Active Internet connections (only servers)
5 Proto Recv-Q Send-Q Local Address           Foreign Address         State       PID/Program name
6 udp        0      0 0.0.0.0:5355            0.0.0.0:*                -
7 udp        0      0 127.0.0.53:53           0.0.0.0:*                -
8 udp        0      0 192.168.0.220:68        0.0.0.0:*                -
9 udp        0      0 0.0.0.0:1337            0.0.0.0:*                25278/./server
10 udp6       0      0 :::5355                 :::*                     -
```

Comanda de mai sus afișează detaliat informații despre procesele din sistem care folosesc protocolul UDP.

După cum vă informează și primele două linii de output, trebuie să rulați cu sudo pentru a putea vedea informații despre toate procesele, inclusive cele pe care nu le dețineți.

- tcpdump

```
1 $ sudo tcpdump -i any -n dst port 1337
2 tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
3 listening on any, link-type LINUX_SLL (Linux cooked), capture size 262144 bytes
4 22:11:13.144809 IP 127.0.0.1.59153 > 127.0.0.1.2222: UDP, length 1500
5 22:11:13.144835 IP 127.0.0.1.59153 > 127.0.0.1.2222: UDP, length 1500
```

Comanda monitorizează tot traficul care are ca destinație portul 1337 și afișează informații.