

# Packet Crafting

OpenFest, Sofia  
November, 2013

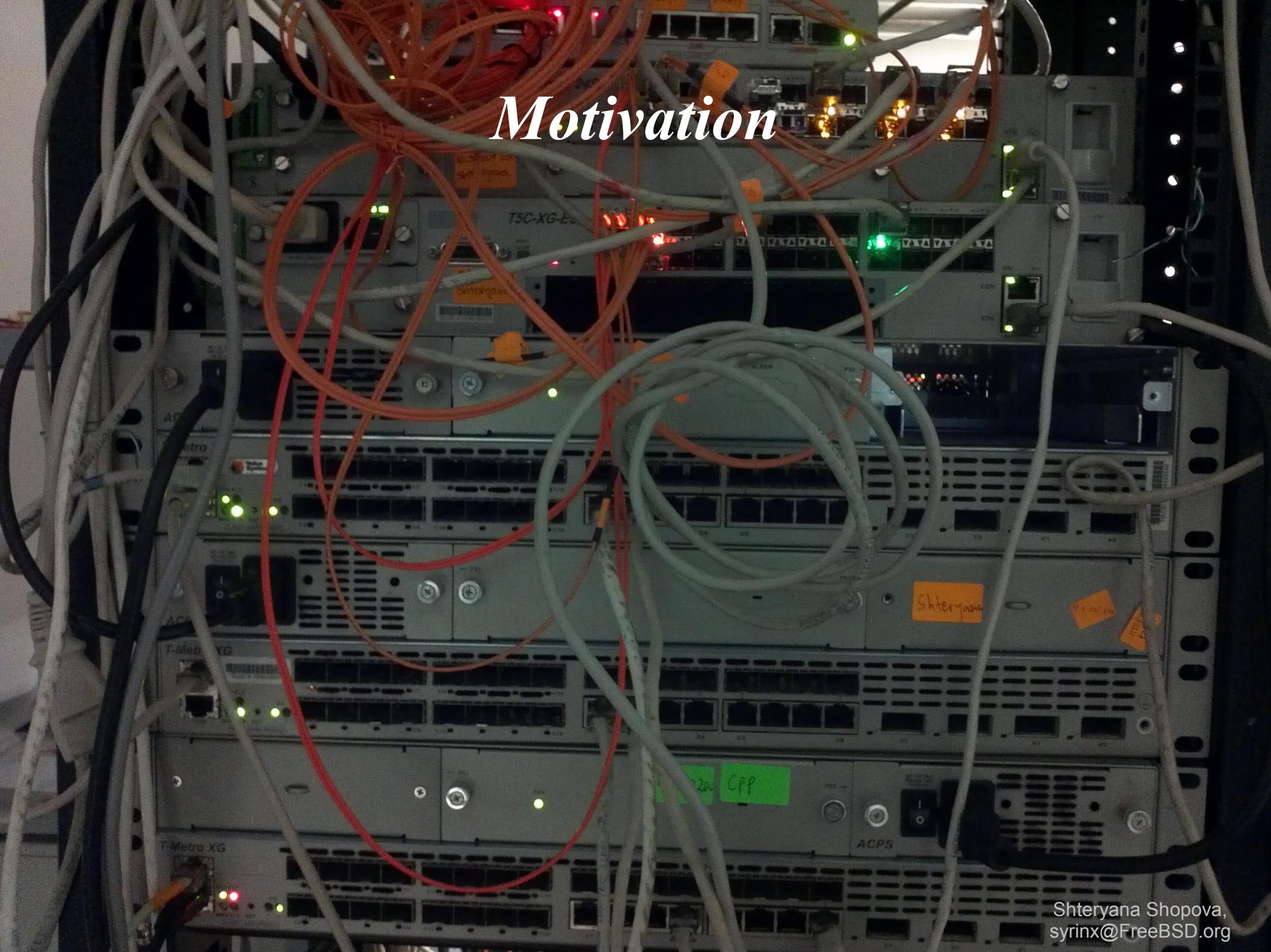


# Disclaimer

!!! Use at your own risk !!!



# *Motivation*



# *Motivation (2)*



Shteryana Shopova,  
syrinx@FreeBSD.org

# *Agenda*

- ping(8), traceroute(8), telnet(1), nmap(1)
- ng\_source(4), tcpdump(1)
- bridge(4), vlan(4), svlan(4), trunk(4)/lagg(4)
- yersinia(8)
- nemesis(1), libnet
- hyenae
- pierf, Scapy
- netmap(4)
- iperf
- PF\_PACKET, etc... sockets, C code
- OS network stack & daemons



# *Let's start*

- ping(8)
  - options: flood, quiet, multicast, source route, etc
- traceroute(8)
  - incrementing IP TTL probes
  - UDP, but may be random IP protocol number
- telnet(1)
  - TCP port 23, unless custom specified
- nmap(1)
  - network security scanner
- and more
  - arping, nslookup/host/dig, ...

```
RX packets:200 errors:0 dropped:0 overruns:0 frame:0  
TX packets:200 errors:0 dropped:0 overruns:0 carrier:0  
collisions:0 txqueuelen:0  
RX bytes:18564 (18.5 KB) TX bytes:18564 (18.5 KB)
```

```
sotir@sotir-HP-550:~$ nslookup yahoo.com  
;; connection timed out; no servers could be reached
```

```
sotir@sotir-HP-550:~$ cat /etc/resolv.conf  
# Dynamic resolv.conf(5) file for glibc resolver(3) generated by  
#           DO NOT EDIT THIS FILE BY HAND -- YOUR CHANGES WILL BE OVERWRITTEN  
nameserver 127.0.1.1  
sotir@sotir-HP-550:~$ ping 8.8.8.8  
PING 8.8.8.8 (8.8.8.8) 56(84) bytes of data.
```

^C

```
The language support files are incomplete. You can install them now or click "Run this action now" and follow the instructions. If you would like to install language support instead (click the "Run this action now" button), click "Close" and then run "sudo apt-get install language-support".
```

```
sotir@sotir-HP-550:~$ traceroute 8.8.8.8
```

The program 'traceroute' can be found in the following packages:

- \* inetutils-traceroute
- \* traceroute

Try: sudo apt-get install <selected package>

```
sotir@sotir-HP-550:~$
```

Run this action now

Close

# *ng\_source(4)*

- `tcpdump(8)`

```
0x0000. 0a00
11:40:23.411341 00:e0:0c:11:95:22 > 00:00:00:77:77:78, ethertype MPLS unicast (0x8847), length 82: MPLS (label 28673,
 0x0000: 0000 0077 7778 00e0 0c11 9522 8847 0700
 0x0010: 11ff 0100 5e02 020b 0040 9530 450b 8100
 0x0020: 0001 0800 4500 002e fdc5 0000 4011 35b9
 0x0030: 6401 0132 e002 020b f28f 4908 001a 9d1b
 0x0040: 476f 7420 4d75 6c74 6963 6173 743f 213f
 0x0050: 0a00
```

- there's wireshark too
- `netgraph(3)`
  - graph based kernel networking subsystem of FreeBSD
- `ng_ether(4)`
  - a node automatically created for every Ethernet interface in the system

# *ng\_source(4) contd.*

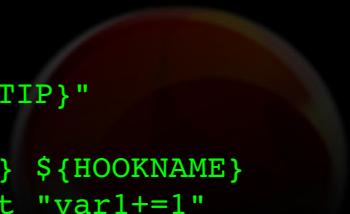
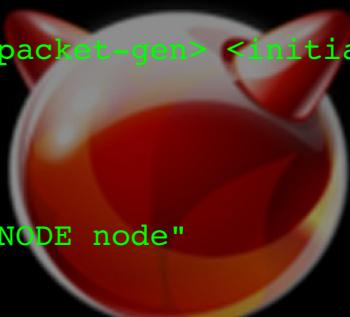
```
#!/bin/sh
#####
# TODO: insert some nice comment here
#####
ECHO=/bin/echo
NGHOOK=/usr/sbin/nghook
HOOKNAME=input

usage () {
$ECHO      "Usage: $0 <ng-node> <count> <packet-gen> <initial-arg> <increment-gen>"
}

...
$ECHO "Injecting $LIMIT packets into $NGNODE node"

# Loop from 1 to 255
while [ "$var1" -le $LIMIT ]
do
DSTIP=${ARGS}
$ECHO "Injecting packet with DST IP ${DSTIP}"

${PACKGEN} ${DSTIP} | ${NGHOOK} ${NGNODE} ${HOOKNAME}
let "var1 = var1 + 1" >> /dev/null  # let "var1+=1"
ARGS=`${INCGEN} ${DSTIP}`
done
exit 0
```



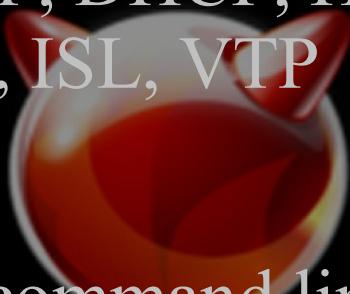
# *Ether-like interfaces*

- bridge(4)
- vlan(4), svlan(4)
- trunk(4)/lagg(4)
- tap(4)



# *Yersinia*

- <http://www.yersinia.net>
- Supported protocols
  - STP, CDP, DTP, DHCP, HSRP, IEEE 802.1Q, IEEE 802.1X, ISL, VTP
- UI
  - GTK, ncurses, command line, netclient
- predefined scenarios
- netclient
  - TCP port 12000
  - passwd/enable passwd : root/tomac



# *nemesis(1)*

NEMESIS -- The NEMESIS Project Version 1.4 (Build 26)

NEMESIS Usage:

```
nemesis [mode] [options]
```

NEMESIS modes:

- arp
- dns
- ethernet
- icmp
- igmp
- ip
- ospf (currently non-functional)
- rip
- tcp
- udp



```
mv -f .deps/hyenaes-uncp.Tpo .deps/hyenaes-uncp.Po
gcc -DHAVE_CONFIG_H -I. -I.. -g -O2 -MT hyenaes-hsrp.o -MD -MP -MF .deps/hyenaes-hsrp.Tpo -c -o hyenaes-hsrp.o hyenaes-hsrp.c
mv -f .deps/hyenaes-hsrp.Tpo .deps/hyenaes-hsrp.Po
gcc -DHAVE_CONFIG_H -I. -I.. -g -O2 -MT hyenaes-attack.o -MD -MP -MF .deps/hyenaes-attack.Tpo -c -o hyenaes-attack.o hyenaes-attack.c
mv -f .deps/hyenaes-attack.Tpo .deps/hyenaes-attack.Po
gcc -DHAVE_CONFIG_H -I. -I.. -g -O2 -MT hyenaes-patterns.o -MD -MP -MF .deps/hyenaes-patterns.Tpo -c -o hyenaes-patterns.o hyenaes-patterns.c
mv -f .deps/hyenaes-patterns.Tpo .deps/hyenaes-patterns.Po
gcc -DHAVE_CONFIG_H -I. -I.. -g -O2 -MT hyenaes-config.o -MD -MP -MF .deps/hyenaes-config.Tpo -c -o hyenaes-config.o hyenaes-config.c
mv -f .deps/hyenaes-config.Tpo .deps/hyenaes-config.Po
gcc -DHAVE_CONFIG_H -I. -I.. -g -O2 -MT hyenaes-base.o -MD -MP -MF .deps/hyenaes-base.Tpo -c -o hyenaes-base.o hyenaes-base.c
mv -f .deps/hyenaes-base.Tpo .deps/hyenaes-base.Po
gcc -g -O2 -o hyenae hyenae.o hyenae-assistant.o hyenae-remote.o hyenae-common.o hyenae-protocol.o hyenae-eth.o hyenae-arp.o hyenae-pppoe.o hyenae-ip.o hyenae-icmp.o hyenae-tcp.o hyenae-udp.o hyenae-dns.o hyenae-bootp.o hyenae-nrhp.o hyenaes-hsrp.o hyenaes-attack.o hyenaes-patterns.o hyenaes-config.o hyenaes-base.o -lpcap -ldnet -lpthread -lpthread -lpcap -ldnet -lpcap -ldnet -lpthread -lpcap -ldnet
```

• No \*BSD packages

# No \*BSD packages

- available on SF
- assumes /usr/include & /usr/lib

# *hyenae (contd.)*

- “ERROR: Root privileges required”
- Supported protocols
  - ARP, PPPoE, ICMP, TCP, UDP, DNS, DHCP, HSRP
- UI
  - command line, QT-frontend
- Example

```
$ sudo hyenae -a pppoe-discover -i em0 -s 01:00:82:00:00:c2 -d 01:00:82:00:00:c3
* Initializing
* Opening network interface (em0)
* Launching attack
```

Press any key to stop

# *Scapy*

- “Powerful interactive packet manipulation program in python”
- Python interpreter disguised as a Domain Specific Language
- Fast Packet Designing
- Interactive packet and result manupulation
- Fast packet generator ? :dd



# *Scapy - ls()*

```
Welcome to Scapy (2.2.0)
>>>
>>> ls()
ARP          : ARP
ASN1_Packet : None
BOOTP       : BOOTP
CookedLinux : cooked linux
DHCP        : DHCP options
DHCP6       : DHCPv6 Generic Message)
DHCP6OptAuth : DHCP6 Option - Authentication
DHCP6OptBCMCSDomains : DHCP6 Option - BCMCS Domain Name List
DHCP6OptBCMCServers : DHCP6 Option - BCMCS Addresses List
DHCP6OptClientFQDN : DHCP6 Option - Client FQDN
DHCP6OptClientId : DHCP6 Client Identifier Option
DHCP6OptDNSDomains : DHCP6 Option - Domain Search List option
DHCP6OptDNSServers : DHCP6 Option - DNS Recursive Name Server
DHCP6OptElapsedTime : DHCP6 Elapsed Time Option
DHCP6OptGeoConf :
DHCP6OptIAAddress : DHCP6 IA Address Option (IA_TA or IA_NA suboption)
...
<6 more screens of output>
```

# *Scapy - lcs()*

```
>>> lsc()
bind_layers      : Bind 2 layers on some specific fields' values
corrupt_bits     : Flip a given percentage or number of bits from a string
corrupt_bytes    : Corrupt a given percentage or number of bytes from a string
defrag           : defrag(plist) -> ([not fragmented], [defragmented],
                           : defrag(plist) -> plist defragmented as much as possible
                           : Exploit Etherleak flaw
                           : Fragment a big IP datagram
                           : Transform a layer into a fuzzy layer by replacing some default values by
random objects   : Show differences between 2 binary strings
hexdiff          : --
hexdump          : --
hexedit          : --

rdpcap            : Read a pcap file and return a packet list
send              : Send packets at layer 3
sendp             : Send packets at layer 2
sendpfast         : Send packets at layer 2 using tcpreplay for performance
sniff             : Sniff packets
split_layers      : Split 2 layers previously bound
sr                : Send and receive packets at layer 3
srbt              : send and receive using a bluetooth socket
srflood           : Flood and receive packets at layer 3
srloop            : Send a packet at layer 3 in loop and print the answer each time
srp               : Send and receive packets at layer 2
traceroute        : Instant TCP traceroute
wrpcap            : Write a list of packets to a pcap file

<some commands omitted, such as wireshark>
```

# *Scapy – sending a packet*

```
Welcome to Scapy (2.2.0)
>>> a=Ether()/IP(dst="www.slashdot.org")/TCP()/"GET /index.html
HTTP/1.0 \n\n"
>>> hexdump(a)
0000  00 22 68 5A AA E4 74 E5  0B E3 39 68 08 00 45 00  ."hZ..t...9h..E.
0010  00 43 00 01 00 00 40 06  41 3B AC 14 00 12 D8 22  .C.....@.A;...."
0020  B5 30 00 14 00 50 00 00  00 00 00 00 00 00 50 02  .0...P.....P.
0030  20 00 84 38 00 00 47 45  54 20 2F 69 6E 64 65 78  ..8..GET /index
0040  2E 68 74 6D 6C 20 48 54  54 50 2F 31 2E 30 20 0A  .html HTTP/1.0 .
0050  0A
>>> sendp(Ether()/IP(dst="1.2.3.4",ttl=(1,4)), iface="wlan0")
...
Sent 4 packets.
```

# *Scapy – writing your own class*

- TBD
- (homework)



# *netmap(4)*

- fast packet I/O framework
- reduce syscalls to improve packet processing performance
- in many cases, shared buffers between user and kernel space
- open("/dev/netmap")
- ioctl(.., NIOCREGIF, ..) to bind to an interface
- fill in available buffers with data
- non-blocking ioctl(.., NIOCTXSYNC) /ioctl(.., NIOCRXSYNC) to transmit/receive data
- supported drivers - em(4), igb(4), ixgbe(4), re(4)



```
shteryana@aphrodite:/usr/src/tools/tools/netmap %  
shteryana@aphrodite:/usr/src/tools/tools/netmap % ./pkt-gen
```

```
main [1461] missing ifname
```

## Usage: *netmap(4)'s pkt-gen- example*

pkt-gen arguments

-i interface	interface name
-f function	tx rx ping pong
-n count	number of iterations (can be 0)
-t pkts_to_send	also forces tx mode
-r pkts_to_receive	also forces rx mode
-l pkts_size	in bytes excluding CRC
-d dst-ip	end with %n to sweep n addresses
-s src-ip	end with %n to sweep n addresses
-D dst-mac	end with %n to sweep n addresses
-S src-mac	end with %n to sweep n addresses
-a cpu_id	use setaffinity
-b burst size	testing, mostly
-c cores	cores to use
-p threads	processes/threads to use
-T report_ms	milliseconds between reports
-P	use libpcap instead of netmap
-w wait_for_link_time	in seconds

```
shteryana@aphrodite:/usr/src/tools/tools/netmap %
```

# *netmap(4) - example*

```
char *buf;
int fd;
struct pollfd      fds;
struct nmreq       nmr;
struct netmap_if *nifp;

fd = open("/dev/netmap", O_RDWR);

strlcpy(nmr.nr_name, "em0", sizeof(nmr.nr_name));
nmr.nr_version = NETMAP_API;
ioctl(ng.fds.fd, NIOCGINFO, &nmr);

buf = mmap(NULL, nmr.nr_memsize, PROT_READ | PROT_WRITE,
           MAP_SHARED, fd, 0));

ioctl(fd, NIOCREGIF, &nmr);

fds.events = POLLOUT | POLLWRNORM | POLLWRBAND;
fds.fd = fd;

/* prepare packet in buf */
poll(&fds, 1, INFTIM);

ioctl(fd, NIOCTXSYNC, NULL); /* optional */
```

# *iperf*

- commonly used network testing tool
- creates UDP/TCP streams to measure throughput
- iperf (July, 2010) vs iperf3 (March, 2013)
- client and server
- uni-/bi-directional tests



# PF\_PACKET socket & friends

## 29.3 Datalink Provider Interface (DPLI)

SVR4 provides datalink access through DPLI. DPLI is a protocol-independent interface designed by AT&T that interfaces to the service provided by the datalink layer [Unix International 1991]. Access to DPLI is by sending and receiving STREAMS messages.

There are two styles of DPLI. In one style, there is a single device to open, and the desired interface is specified using a DPLI `DI_ATTACH_REQ` request. In the other style, the application simply opens the device (e.g., `/dev/bt`). But for efficient operation, two additional STREAMS modules are normally pushed onto the stream: `pifmod`, which performs packet filtering within the kernel, and `bifmod`, which buffers the data destined for the application. We show this in Figure 29.2.

Conceptually, this is similar to what is described in the previous section. The `bifmod` supports buffering within the kernel, so each message can be reduced to the amount of data and number of bytes it supports, a small max length and a read timeout.

One interesting difference, however, is the type of pseudomachine supported by the BPF and `pifmod` filters. The BPF filter is a directed acyclic control flow graph (CGC), while `pifmod` uses a Boolean expression tree. The main important difference is that the register machine used by the BPF is natural for a hardware or a software implementation [McNamee and Jacobson 1993]. Figure 29.2 shows that the DPLI interface is very similar to the BPF interface, except that the BPF interface is expression tree, depending on the complexity of the filter.

Another difference is that BPF allows the filtering decision before copying the packet, if copied to user-space, the BPF filter will discard. Depending on the BPF implementation, the process may be delayed by up to 100ms, with a 10ms discard.

- PF\_PACKET
  - Linux-only
  - SOCK\_DGRAM or SOCK\_RAW modes

## 29.4 DPLI, SOCK\_PACKET and PF\_PACKET

There are two methods of receiving packets from the datalink layer, the original method, which is more widely available but less flexible, since it uses type `SOCK_PACKET`. The newer method, which introduces more performance features, is to create a socket of family `PF_PACKET`. To do this, sufficient privileges (similar to creating a raw socket), and the backlog must be a nonzero value specifying the maximum number of pending sockets, the second argument to `socket` can be `SOCK_DGRAM` packets with the link-layer header removed or `SOCK_RAW` link-layer packets. `SOCK_PACKET` sockets only return the complete frame to receive all frames from the datalink, we need to use `recvfrom`, `SOCK_DGRAM`, when `PF_PACKET` is used.

If we want to only BPF from a specific interface (e.g., `eth0`), we can use `SO_ATTACH_FILTER` or `SO_BROADCAST` (e.g., `SO_BROADCAST = 1`) to filter options.

Other constants for the final argument are `ETH_P_ARP` and `ETH_P_IGMP`, for example. Specifying a protocol of `ETH_P_ARP` tells the datalink which frame types to pass to the socket for the frames the datalink receives. If the datalink supports a promiscuous mode (e.g., an Ethernet), then the device must be put into a promiscuous mode. The code will then receive all frames, even if they are not destined for the socket. The `SO_ATTACH_FILTER` option can be combined by an OR of `SO_DONTFLOOD` to flag the flags, setting the `IFF_PROMISC` flag, and then setting the flags with `SO_DONTFLOOD`. Unfortunately, with this method, multiple promiscuous listeners can interfere with each other and a buggy program can leave promiscuous mode on even after it exits.

Some differences are evident when comparing this Linux feature to BPF.

• The Linux feature provides no kernel buffering and filtering is only available on newer systems (via the `SO_ATTACH_FILTER` socket option). It is a normal socket receive buffer, but multiple frames cannot be buffered together and passed to the application with a single read. This increases the overhead involved in copying the potentially voluminous amounts of data from the kernel to the application.

2. `SOCK_PACKET` provides no filtering by device. (`PF_PACKET` sockets can be limited to a device by calling `bndt`). If `ETH_P_IP` is specified in the call to `socket`, then all IPv4 packets from all devices (Ethernet, PPP links, SLIP links and the loopback device, for example) are passed to the socket. A generic socket address structure is returned by `recvfrom`, and the `sa_data` member contains the device name (e.g., `eth0`). The application must then discard data from any device in which it is not interested. The problem again is too much data can be returned to the application, which can get in the way when monitoring a big open network.

3. `libpcap`: Packet Capture Library

The packet capture library, `libpcap`, provides implementation-independent access to the underlying packet capture facility provided by the OS. Currently, it supports only the reading of packets (although adding a few lines of code to the library lets one write datagrams too on some systems). See the next section for a description of another library that supports not only writing datalink packets, but also constructing arbitrary packets.

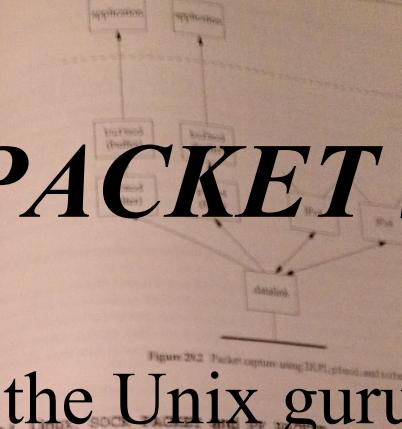


Figure 29.2 Packet capture using DPLI, pifmod, and bifmod

## Chapter 29. Datalink Access

Introduction	787
29.1	788
29.2	788
29.3	790
29.4	791
29.5	792
29.6	792
29.7	793
29.8	793
Summary	815

## Chapter 30. Client/Server Design Alternatives

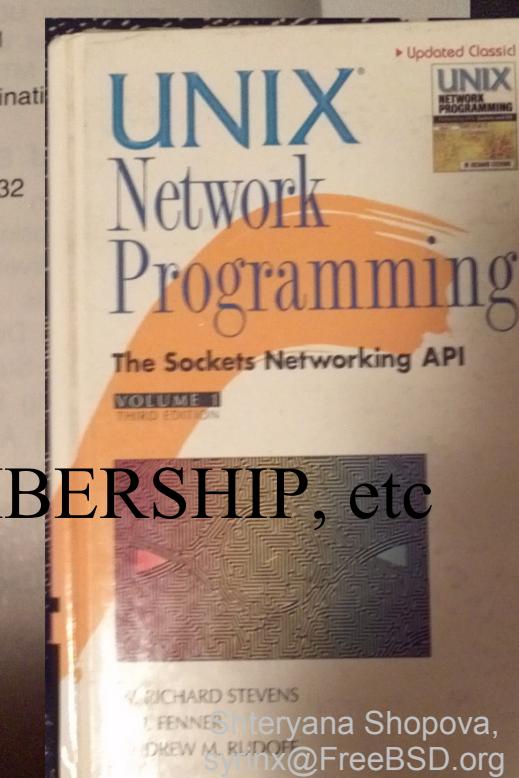
Introduction	817
30.1	817
30.2	819
30.3	820
30.4	821
30.5	822
30.6	822
30.7	822
30.8	822
TCP Concurrent Server, One Child per Client	822
TCP Preforked Server, No Locking Around accept	826
TCP Preforked Server, File Locking Around accept	832
TCP Preforked Server, Thread Locking Around accept	835
TCP Preforked Server, Descriptor Passing	836

## Chapter 27. IP Options

Introduction	709
27.1	709
27.2	709
27.3	711
27.4	719
27.5	725
27.6	731
27.7	731
27.8	733
27.9	732
Summary	733

## Chapter 28. Raw Sockets

Introduction	735
28.1	735
28.2	736
28.3	737
28.4	739
28.5	741
28.6	755
28.7	760
28.8	769
Summary	786



# *References*

<http://www.kohala.com/start/unpv12e.html>

<http://www.unpbook.com/src.html>

<http://www.secdev.org/projects/scapy/>

<http://info.iet.unipi.it/~luigi/netmap>

[http://www.secdev.org/conf/scapy\\_pacsec05.pdf](http://www.secdev.org/conf/scapy_pacsec05.pdf)

<http://docs.python.org/2/tutorial/introduction.html>

<http://www.secdev.org/projects/scapy/doc/usage.html#interactive-tu>

# Questions?



# Thank you!



\* sources from this presentation available at <http://people.freebsd.org/~syrinx/pktgen/>