genua.

A Packet's Journey Through the OpenBSD Network Stack

Alexander Bluhm

September 2024

genua.

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

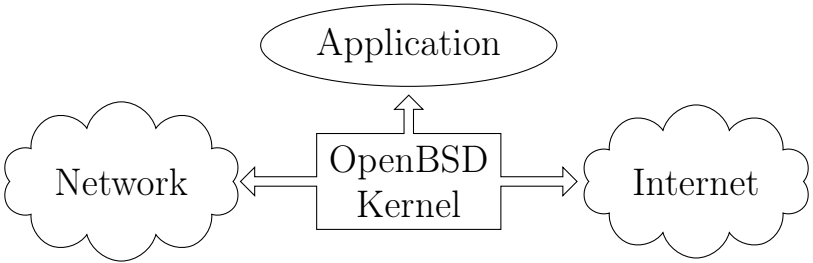
IPsec

Tips & Tricks

Final Words

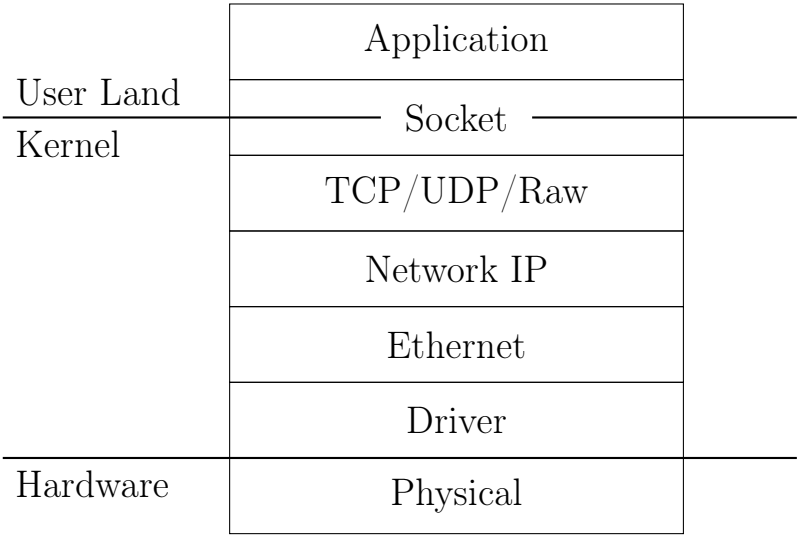
genua.

Environment



genua.

Layer


Application
User Land
Socket
Kernel
TCP/UDP/Raw
Network IP
Ethernet
Driver
Hardware
Physical

genua.

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

Tips & Tricks

Final Words

genua.

Protocol Statistics Counter

netstat -ss

```
ip:
    2208239 total packets received
    2204405 packets for this host
    3717 packets for unknown/unsupported protocol
    101 packets not forwardable
tcp:
    1542163 packets sent
        616615 data packets (165690394 bytes)
        333 data packets (296898 bytes) retransmitted
        740242 ack-only packets (646590 delayed)
        153288 window update packets
```

genua.

pf State

pfctl -s states

```
em0 tcp 80.154.94.48:22 <- 84.191.39.187:60436
       ESTABLISHED:ESTABLISHED
em0 tcp 192.168.0.31:22 (80.154.94.48:10031) <- 80.154.94.10:19890
       ESTABLISHED:ESTABLISHED
em1 tcp 80.154.94.10:19890 -> 192.168.0.31:22
       ESTABLISHED:ESTABLISHED
em0 tcp 192.168.0.31:443 (80.154.94.48:443) <- 193.30.66.53:51727
       CLOSING:ESTABLISHED
```

genua.

Network Sockets

netstat -an

```
Active Internet connections (including servers)
Proto   Recv-Q Send-Q  Local Address   Foreign Address  TCP-State
tcp          0      0  10.0.1.37.22    10.0.1.2.30176   ESTABLISHED
tcp          0   1144  10.0.1.37.22    10.0.1.1.27069   ESTABLISHED
tcp          0     44  10.0.1.37.22    10.0.1.2.49002   ESTABLISHED
tcp          0      0  *.22            *.*              LISTEN
Active Internet connections (including servers)
Proto   Recv-Q Send-Q  Local Address   Foreign Address
udp          0      0  10.0.1.37.7040  10.0.1.1.123
udp          0      0  10.0.1.37.161   *.*
udp          0      0  *.*             *.*
```

genua.

Tcpdump Traffic

tcpdump -ni enc0 -v

```
18:45:13.665393 (unprotected): SPI 0x00006861:
  fdd7:e83e:66bc:100::17 > fdd7:e83e:66bc:100::70:
  10.188.168.17 > 10.188.175.72:
  icmp: echo request
  (id:c52b seq:0) (ttl 255, id 8945, len 1028) (len 1028, hlim 64)
18:45:13.665849 (unprotected): SPI 0x00006862:
  fdd7:e83e:66bc:100::70 > fdd7:e83e:66bc:100::17:
  10.188.175.72 > 10.188.168.17:
  icmp: echo reply
  (id:c52b seq:0) (ttl 253, id 7503, len 1028) (len 1028, hlim 64)
```

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

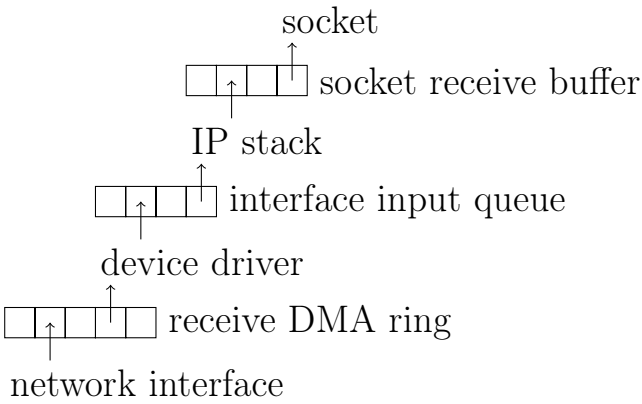
Tips & Tricks

Final Words

genua.

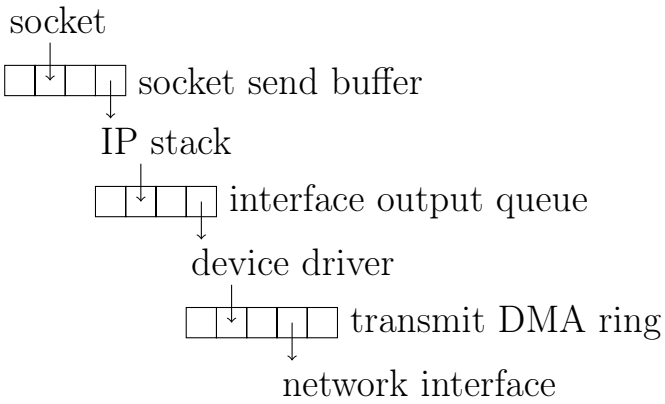
genua.

Input Queues



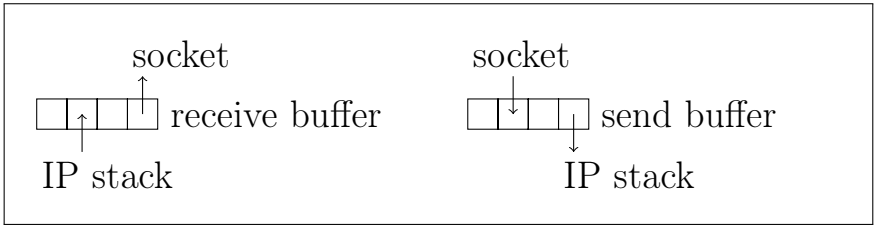
genua.

Output Queues



genua.

Socket Buffer Recv-Q Send-Q

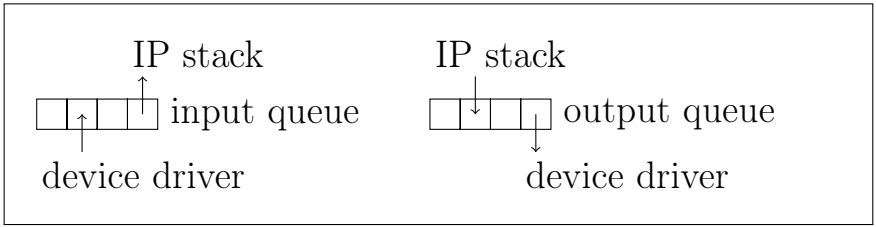


netstat -n -p tcp

```
Active Internet connections (including servers)
Proto  Recv-Q Send-Q  Local Address      Foreign Address    TCP-State
tcp     65160      0  10.10.11.2.35477   10.10.11.1.41996   ESTABLISHED
tcp         0 131732  10.10.12.3.44648   10.10.12.4.12345   ESTABLISHED
```

genua.

Interface Queue qdrops



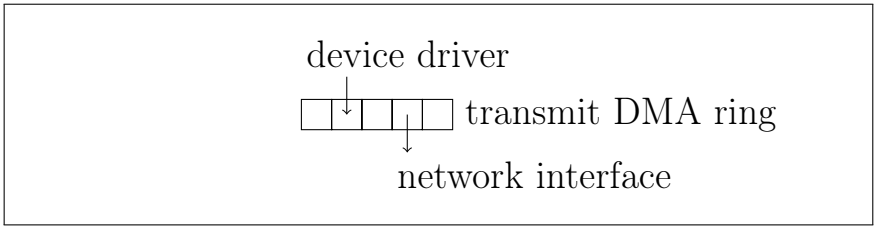
kstat

```
ix0:0:rxq:0
     packets: 158069114 packets
       bytes: 300328302395 bytes
      qdrops: 7902615 packets
        qlen: 0 packets
    enqueues: 6947931
    dequeues: 5729695
```

```
ix0:0:txq:0
     packets: 11094621 packets
       bytes: 894601044 bytes
      qdrops: 0 packets
        qlen: 0 packets
    maxqlen: 255 packets
```

genua.

Device Driver OACTIVE



ifconfig em1

```
em1: flags=8c43<UP,BROADCAST,RUNNING,OACTIVE,SIMPLEX,MULTICAST> mtu 1500
        lladdr 0c:c4:7a:78:f3:55
        description: Intel I210
        index 6 priority 0 llprio 3
        media: Ethernet autoselect (1000baseT full-duplex)
        status: active
        inet 10.10.12.3 netmask 0xffffff00 broadcast 10.10.12.255
```

genua.

Transmit Rings oactives

dmesg

```
ix1 at pci1 dev 0 function 1 "Intel X550T" rev 0x01, msix,
  4 queues, address a0:36:9f:e0:52:55
```

kstat txq

```
ix1:0:txq:0
        qlen: 90 packets
      maxqlen: 255 packets
      oactive: true
     oactives: 1919595

ix1:0:txq:1
        qlen: 141 packets
      maxqlen: 255 packets
      oactive: true
     oactives: 2251633

ix1:0:txq:2
        qlen: 6 packets
      maxqlen: 255 packets
      oactive: true
     oactives: 1788971

ix1:0:txq:3
        qlen: 55 packets
      maxqlen: 255 packets
      oactive: true
     oactives: 1925063
```

genua.

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

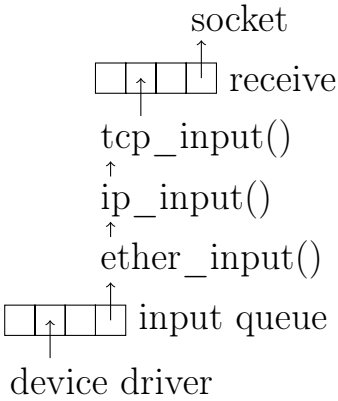
IPsec

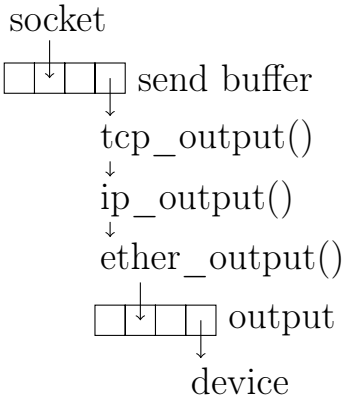
Tips & Tricks

Final Words

genua.

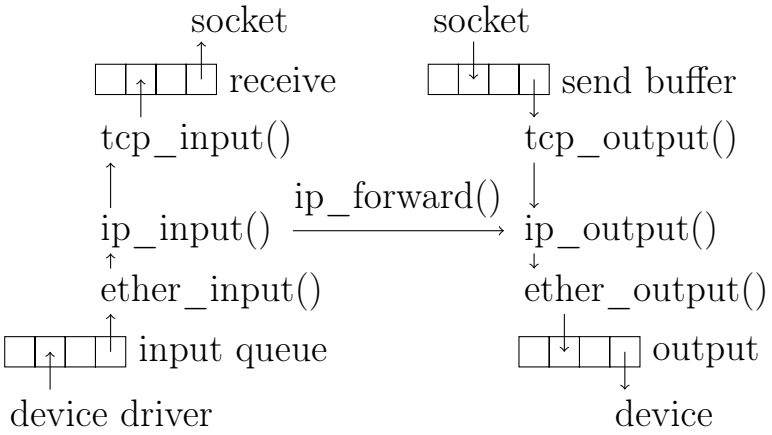
Protocol Stack





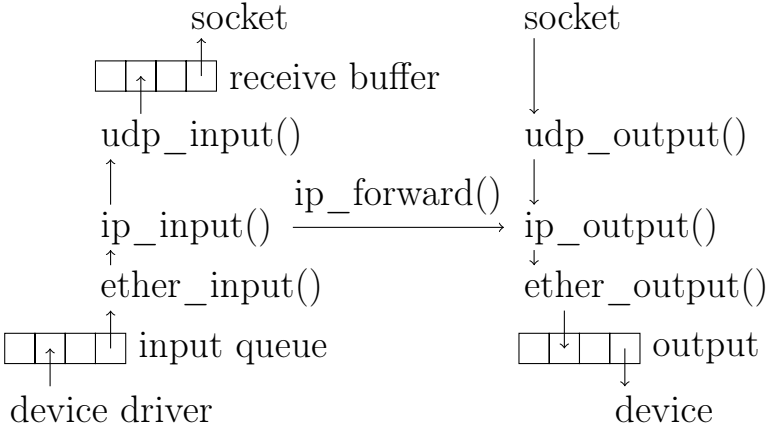
genua.

IP Forwarding



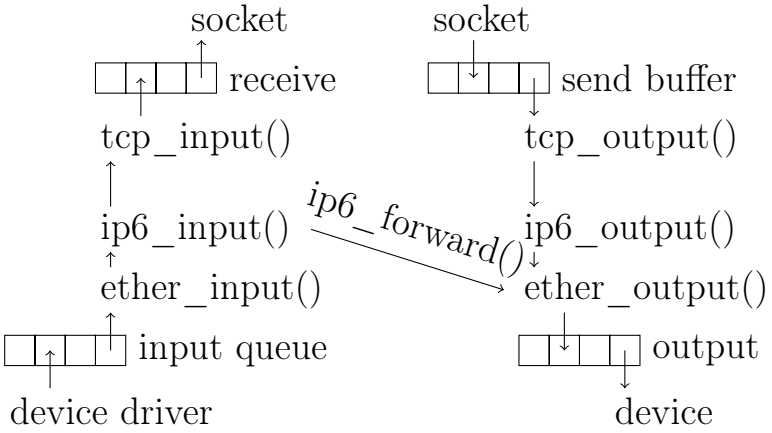
genua.

UDP


socket
receive buffer
udp_input()
ip_input()
ether_input()
input queue
device driver
ip_forward()
socket
udp_output()
ip_output()
ether_output()
output
device

genua.

IPv6


socket
receive
tcp_input()
ip6_input()
ether_input()
input queue
device driver
socket
send buffer
tcp_output()
ip6_output()
ether_output()
output
device
ip6_forward()

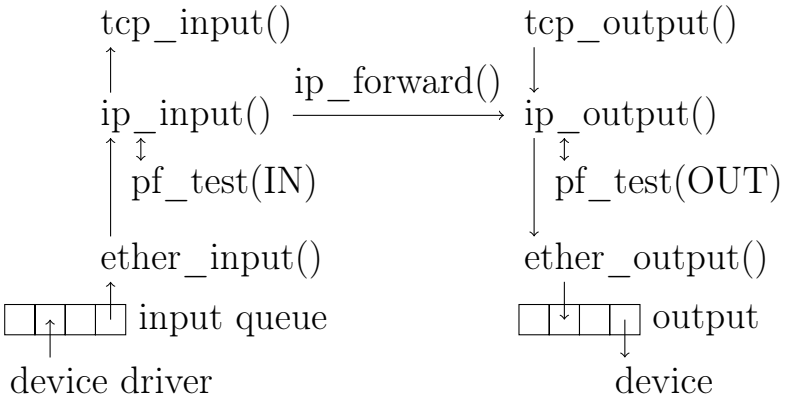
genua.

Netstat Counter
netstat -ss

```
ip:
        393285673 total packets received
        246049103 packets for this host
        126364372 packets forwarded
        242653961 packets sent from this host
tcp:
        86520236 packets sent
        123883097 packets received
udp:
        400260722 datagrams received
        82147309 dropped due to full socket buffers
        318113273 delivered
        396391499 datagrams output
ip6:
        434697159 total packets received
        278214439 packets for this host
        146543897 packets forwarded
        240423068 packets sent from this host
```

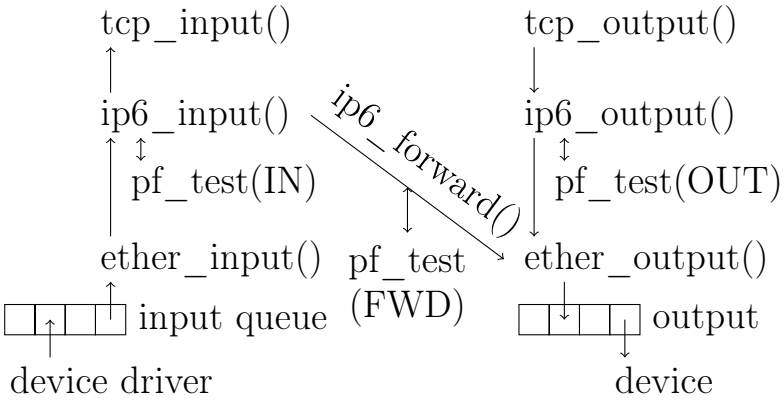
genua.

pf Test



genua.

pf IPv6



genua.

pf Rules

pfctl -s rules

```
block return all
pass all flags S/SA
block return in on ! lo0 proto tcp from any to any port 6000:6010
block return out log proto tcp all user = 55
block return out log proto udp all user = 55
```

genua.

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

Tips & Tricks

Final Words

genua.

Create Device

ifconfig -C

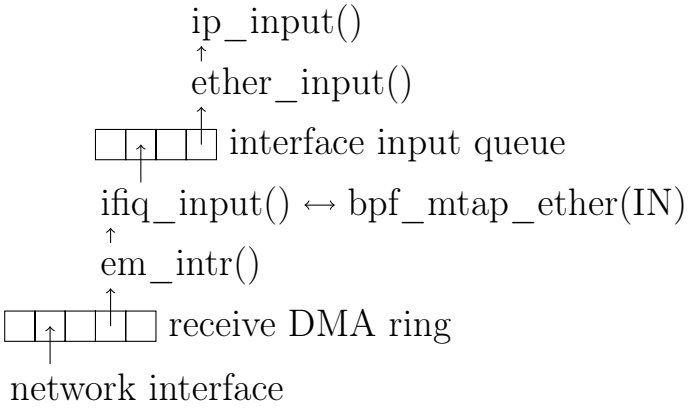
```
aggr bpe bridge carp egre enc eoip etherip gif gre lo mgre mpe mpip
mpw nvgre pair pflog pflow pfsync ppp pppoe rport sec svlan tap
tpmr trunk tun veb vether vlan vport vxlan wg
```

ifconfig vlan0 create vnetid 42 parent em0 up

```
vlan0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        lladdr 0c:c4:7a:78:f3:54
        index 22 priority 0 llprio 3
        encap: vnetid 42 parent em0 txprio packet rxprio outer
        groups: vlan
        media: Ethernet autoselect (1000baseT full-duplex)
        status: active
```

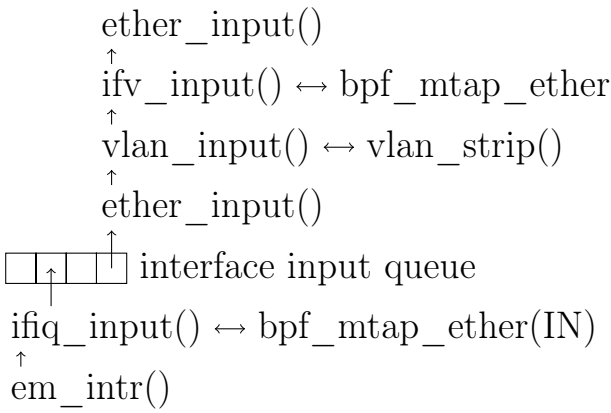
genua.

Hardware Receive and Tcpdump



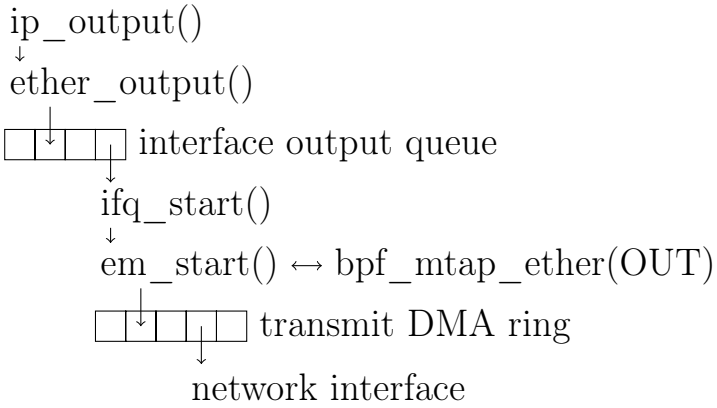
genua.

VLan Receive and Tcpdump



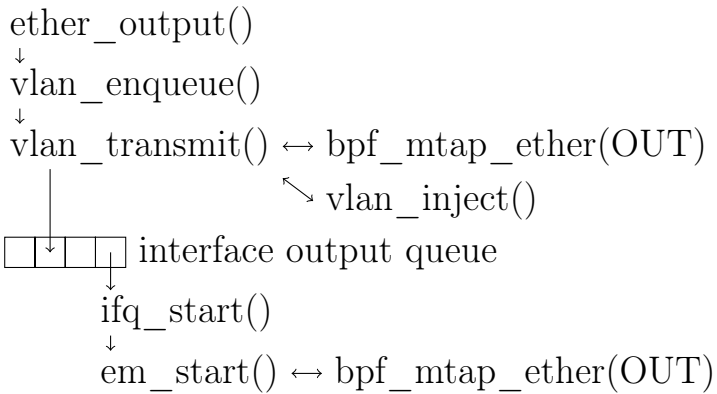
genua.

Hardware Transmit and Tcpdump



genua.

VLan Transmit and Tcpdump



Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

Tips & Tricks

Final Words

genua.

genua.

IPsec Flow and SA

ipsecctl -s flow

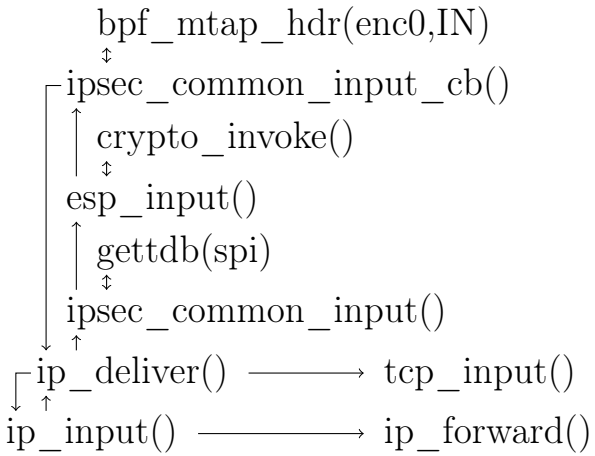
```
flow esp in
    from 10.188.108.0/24 to 10.188.115.0/24
    local fdd7:e83e:66bc:100::70 peer fdd7:e83e:66bc:100::17
flow esp out
    from 10.188.115.0/24 to 10.188.108.0/24
    local fdd7:e83e:66bc:100::70 peer fdd7:e83e:66bc:100::17
```

ipsecctl -s sa

```
esp tunnel
    from fdd7:e83e:66bc:100::17 to fdd7:e83e:66bc:100::70
    spi 0x10000861
esp tunnel
    from fdd7:e83e:66bc:100::70 to fdd7:e83e:66bc:100::17
    spi 0x10000862
```

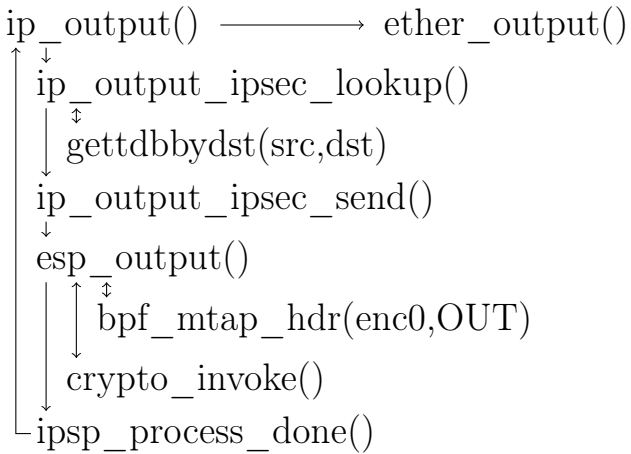
genua.

IPsec Input



genua.

IPsec Output



Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

Tips & Tricks

Final Words

genua.

genua.

Netstat Diff

- netstat -ss >before
- do some networking
- netstat -ss >after
- diff -up before after

```
 tcp:
-       1024475 packets sent
-               829253 data packets (10075190801 bytes)
+       1024496 packets sent
+               829269 data packets (10075194870 bytes)
                8767 data packets (35352640 bytes) retransmitted
-               162711 ack-only packets (163607 delayed)
+               162715 ack-only packets (163619 delayed)
                21271 window update packets
```

genua.

Grep Inet Counter

netstat -ss

```
ip:
        81 packets not forwardable
```

grep 'not forwardable' /usr/src/usr.bin/netstat/inet.c

```
        p(ips\_cantforward, "\t%lu packet%s not forwardable\n");
```

grep -l 'ips_cantforward' /usr/src/sys/*/*

```
        /usr/src/sys/net/pf.c
        /usr/src/sys/netinet/ip_input.c
        /usr/src/sys/netinet/ip_var.h
```

view /usr/src/sys/netinet/ip_input.c

```
        if (!ISSET(flags, IP_FORWARDING)) {
                ipstat_inc(ips_cantforward);
                goto bad;
        }
```

genua.

Counters

Developers should

- check for error
- count errors
- use unique counter

Agenda

Networking

Measurement

Packet Queues

IP Stack

Pseudo Devices

IPsec

Tips & Tricks

Final Words

genua.

genua.

Open Topics

- routing
- arp, neighbor discovery
- bridge, veb
- hardware offloading

genua.

Links

- these slides
  https://github.com/bluhm/talk-packetflow

genua.

Questions

?