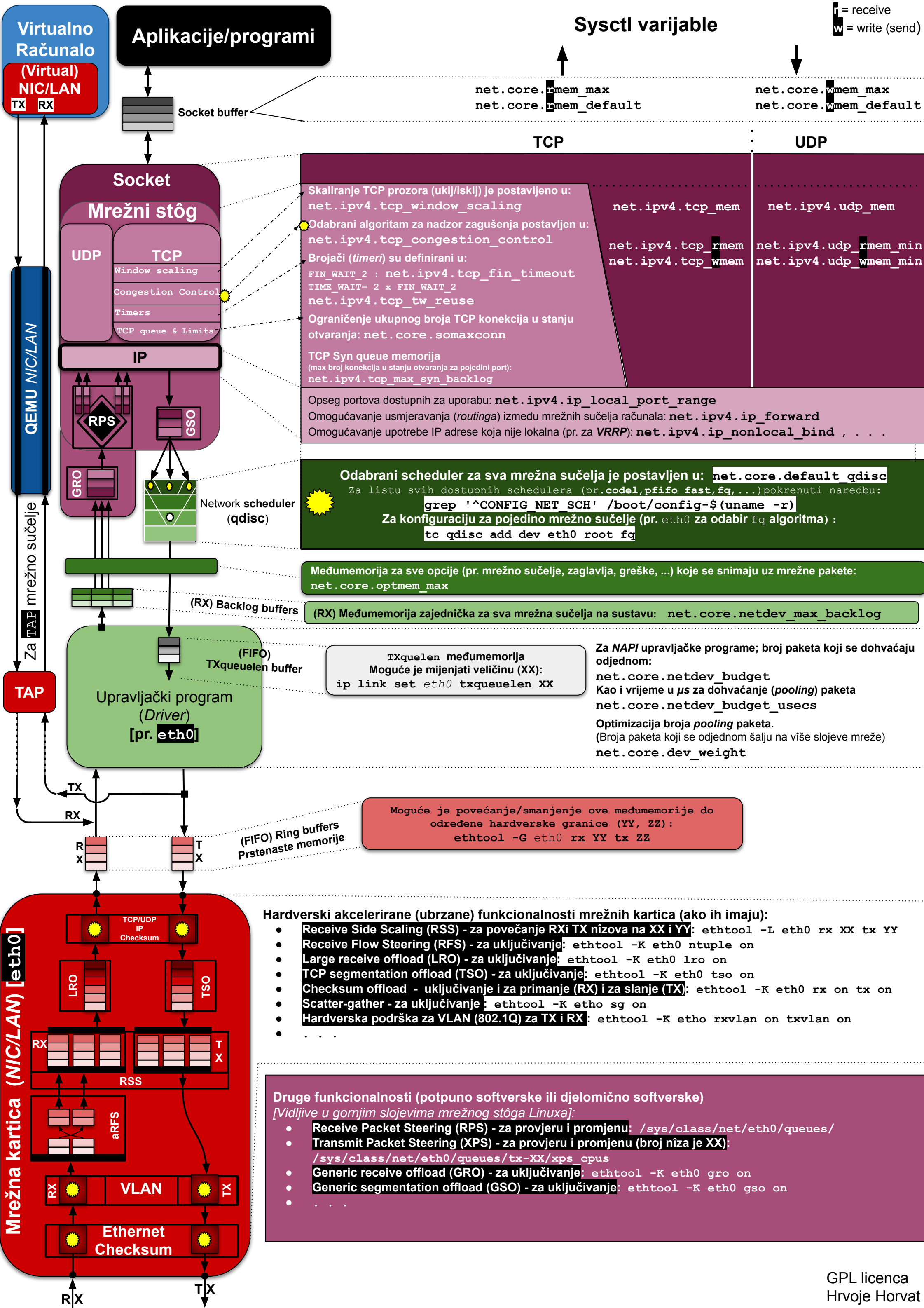
Virtualno Računalo
(Virtual) NIC/LAN
TX RX
Aplikacije/programi
Sysctl varijable
r = receive
w = write (send)
Socket buffer
net.core.rmem_max
net.core.rmem_default
net.core.wmem_max
net.core.wmem_default
TCP
UDP
Socket
Mrežni stôg
UDP
TCP
Window scaling
Congestion Control
Timers
TCP queue & Limits
IP
RPS
GSO
GRO
QEMU NIC/LAN
Za TAP mrežno sučelje
TAP
Skaliranje TCP prozora (uklj/isklj) je postavljeno u:
net.ipv4.tcp_window_scaling
Odabrani algoritam za nadzor zagušenja postavljen u:
net.ipv4.tcp_congestion_control
Brojači (timeri) su definirani u:
FIN_WAIT_2 : net.ipv4.tcp_fin_timeout
TIME_WAIT= 2 x FIN_WAIT_2
net.ipv4.tcp_tw_reuse
Ograničenje ukupnog broja TCP konekcija u stanju otvaranja: net.core.somaxconn
TCP Syn queue memorija
(max broj konekcija u stanju otvaranja za pojedini port):
net.ipv4.tcp_max_syn_backlog
net.ipv4.tcp_mem
net.ipv4.tcp_rmem
net.ipv4.tcp_wmem
net.ipv4.udp_mem
net.ipv4.udp_rmem_min
net.ipv4.udp_wmem_min
Opseg portova dostupnih za uporabu: net.ipv4.ip_local_port_range
Omogućavanje usmjeravanja (routinga) između mrežnih sučelja računala: net.ipv4.ip_forward
Omogućavanje upotrebe IP adrese koja nije lokalna (pr. za VRRP): net.ipv4.ip_nonlocal_bind , . . .
Network scheduler (qdisc)
Odabrani scheduler za sva mrežna sučelja je postavljen u: net.core.default_qdisc
Za listu svih dostupnih schedulera (pr.codel,pfifo fast,fq,...)pokrenuti naredbu:
grep '^CONFIG_NET_SCH' /boot/config-$(uname -r)
Za konfiguraciju za pojedino mrežno sučelje (pr. eth0 za odabir fq algoritma) :
tc qdisc add dev eth0 root fq
Međumemorija za sve opcije (pr. mrežno sučelje, zaglavlja, greške, ...) koje se snimaju uz mrežne pakete:
net.core.optmem_max
(RX) Backlog buffers
(RX) Međumemorija zajednička za sva mrežna sučelja na sustavu: net.core.netdev_max_backlog
(FIFO) TXqueuelen buffer
TXquelen međumemorija
Moguće je mijenjati veličinu (XX):
ip link set eth0 txqueuelen XX
Za NAPI upravljačke programe; broj paketa koji se dohvaćaju odjednom:
net.core.netdev_budget
Kao i vrijeme u μs za dohvaćanje (pooling) paketa
net.core.netdev_budget_usecs
Optimizacija broja pooling paketa.
(Broja paketa koji se odjednom šalju na víše slojeve mreže)
net.core.dev_weight
Upravljački program
(Driver)
[pr. eth0]
TX
RX
R X
T X
(FIFO) Ring buffers
Prstenaste memorije
Moguće je povećanje/smanjenje ove međumemorije do određene hardverske granice (YY, ZZ):
ethtool -G eth0 rx YY tx ZZ
Mrežna kartica (NIC/LAN) [eth0]
TCP/UDP IP Checksum
LRO
TSO
RX
T X
RSS
aRFS
VLAN
Ethernet Checksum
R X
T X
Hardverski akcelerirane (ubrzane) funkcionalnosti mrežnih kartica (ako ih imaju):
• Receive Side Scaling (RSS) - za povečanje RXi TX nîzova na XX i YY: ethtool -L eth0 rx XX tx YY
• Receive Flow Steering (RFS) - za uključivanje: ethtool -K eth0 ntuple on
• Large receive offload (LRO) - za uključivanje: ethtool -K eth0 lro on
• TCP segmentation offload (TSO) - za uključivanje: ethtool -K eth0 tso on
• Checksum offload - uključivanje i za primanje (RX) i za slanje (TX): ethtool -K eth0 rx on tx on
• Scatter-gather - za uključivanje : ethtool -K etho sg on
• Hardverska podrška za VLAN (802.1Q) za TX i RX : ethtool -K etho rxvlan on txvlan on
• . . .
Druge funkcionalnosti (potpuno softverske ili djelomično softverske)
[Vidljive u gornjim slojevima mrežnog stôga Linuxa]:
• Receive Packet Steering (RPS) - za provjeru i promjenu: /sys/class/net/eth0/queues/
• Transmit Packet Steering (XPS) - za provjeru i promjenu (broj nîza je XX):
/sys/class/net/eth0/queues/tx-XX/xps_cpus
• Generic receive offload (GRO) - za uključivanje: ethtool -K eth0 gro on
• Generic segmentation offload (GSO) - za uključivanje: ethtool -K eth0 gso on
• . . .
GPL licenca
Hrvoje Horvat