



Virtualno Računalo  
(Virtual) NIC/LAN  
TX RX

Aplikacije/programi

Socket buffer

Socket Mrežni stôg

UDP TCP  
Window scaling  
Congestion Control  
Timers  
TCP queue & Limits

IP

RFS GRO RPS  
Network scheduler (qdisc)

QEMU NIC/LAN  
Za mrežno sučelje

TAP

TX RX

R X T X

Mrežna kartica (NIC/LAN) [eth0]  
TCP/UDP/IP Checksum  
LRO TSO  
R X T X  
RSS  
aRFS  
RX TX  
VLAN  
Ethernet Checksum  
RX TX

Sysctl variable

net.core.rmem\_max  
net.core.rmem\_default

net.core.wmem\_max  
net.core.wmem\_default

TCP

UDP

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

Mrežne statistike na ovoj razini mreže se mogu dobiti iz datoteka: /proc/softirqs, /proc/net/softnet\_stat, /proc/net/dev te /proc/net/netstat i /proc/net/snmp  
Odnosno sa sljedećim naredbama:  
netstat -i, netstat -s te ss kao i naredbom: nstat

Skaliranje TCP prozora (uklj/isklj) je postavljeno u:  
net.ipv4.tcp\_window\_scaling  
Odabrani algoritam za nadzor zagušenja postavljen je 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  
Iskorištavanje veze u Time\_wait stanju:  
net.ipv4.tcp\_tw\_reuse  
Ograničenje ukupnog broja TCP konekcija u stanju otvaranja: net.core.somaxconn  
TCP Syn queue memorija (max broj konekcija koje su u stanju otvaranja za pojedini port):  
net.ipv4.tcp\_max\_syn\_backlog

Opseg portova dostupnih za uporabu za sve mrežne konekcije na računalo: net.ipv4.ip\_local\_port\_range  
Omogućavanje usmjeravanja (routinga) između mrežnih sučelja računala/poslužitelja: net.ipv4.ip\_forward  
Omogućavanje upotrebe IP adrese koja nije lokalna (pr. za VRRP): net.ipv4.ip\_nonlocal\_bind, ...

Druge funkcionalnosti na ovom sloju mreže (potpuno softverske ili djelomično softverske):  
• 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

Odabrani scheduler za sva mrežna sučelja je postavljen u: net.core.default\_qdisc  
Za listu svih dostupnih schedulera (pr. code1, 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/schedulera:  
tc qdisc add dev eth0 root fq

Međumemorija za sve opcije koje se snimaju uz mrežne pakete (pr. mrežno sučelje, zaglavlja, greške, ...):  
net.core.optmem\_max

(RX) Međumemorija zajednička za sva mrežna sučelja na sustavu: net.core.netdev\_max\_backlog

TXqueuelen međumemorija  
Moguće je mijenjati veličinu (xx):  
ip link set eth0 txqueuelen XX

NAPI Pooling  
Broj paketa koji se mogu spremati u međumemoriju prije NAPI dohvaćanja. Ako imamo uključene LRO i GRO onda se to odnosi na agregirane/spojene pakete: net.core.dev\_weight  
Za NAPI upravljačke programe; broj paketa koji se mogu dohvatiti odjednom, zajedničko za sva mrežna sučelja:  
net.core.netdev\_budget  
Vrijeme u µs dozvoljeno za dohvaćanje (pooling) tih paketa:  
net.core.netdev\_budget\_usecs

Za mrežne statistike na ovoj razini (za eth0), možemo koristiti naredbe poput:  
ip -s link te sar -n EDEV kao i ethtool -S eth0 za najnižu razinu mreže.  
Moguće je gledati statistike i direktno iz datoteka: /proc/net/dev ili /proc/net/netstat ili /proc/net/snmp

Moguće je povećanje/smanjenje ove međumemorije do određene hardverske granice (za RX=YY te za TX=ZZ) sa sljedećom naredbom:  
ethtool -G eth0 rx YY tx ZZ

Hardverski akcelerirane (ubrzane) funkcionalnosti mrežnih kartica (ako ih imaju):  
Receive Side Scaling (RSS) - za povećanje RX i TX nîzova na XX i YY: ethtool -L eth0 rx XX tx YY  
Hardverski Receive Flow Steering (aRFS) - 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 eth0 sg on  
Hardverska podrška za VLAN (802.1Q) za TX i RX: ethtool -K eth0 rxvlan on txvlan on  
Postavke mrežne kartice, poput brzine i duplex načina rada; pr. 1000Mbps (1Gbps), full duplex:  
ethtool -s eth0 speed 1000 duplex full  
Optimizacije signala prekida (IRQ) prema CPU afinitetu (za IRQ broja XY) se podešavaju u:  
/proc/irq/XY/smp\_affinity

Fizička mrežna sučelja (pr. eth0) se sa TAP sučeljem spajaju preko bridge sučelja (pr. br0) !

GPL licenca  
Hrvoje Horvat  
v.1.03