

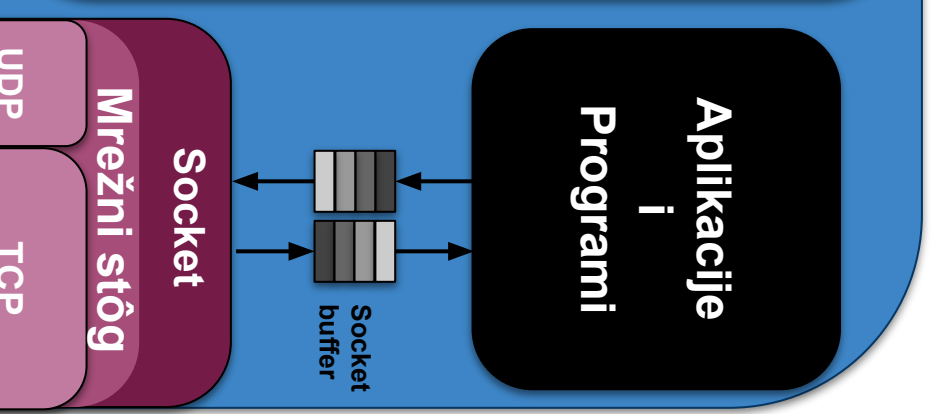
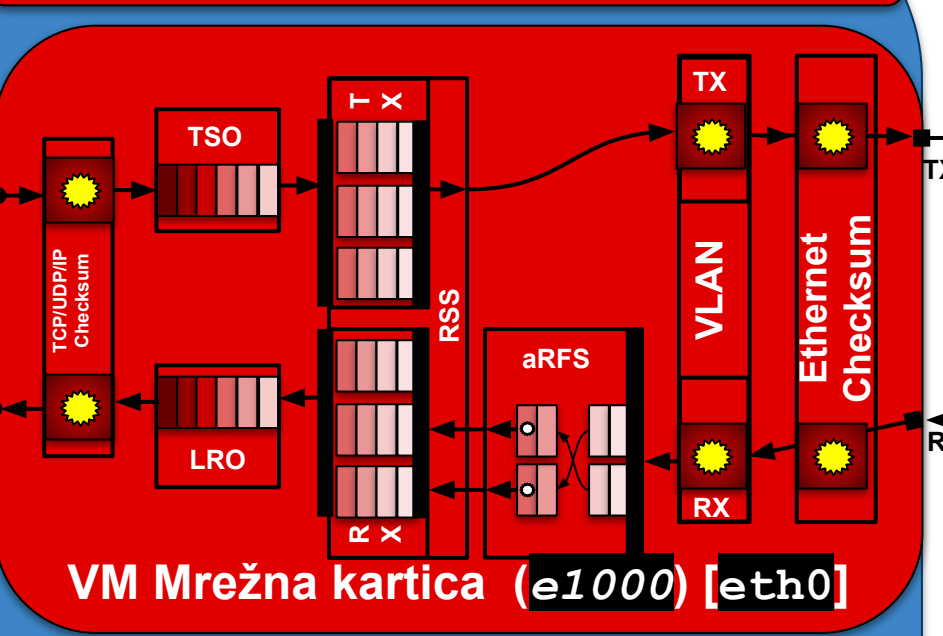
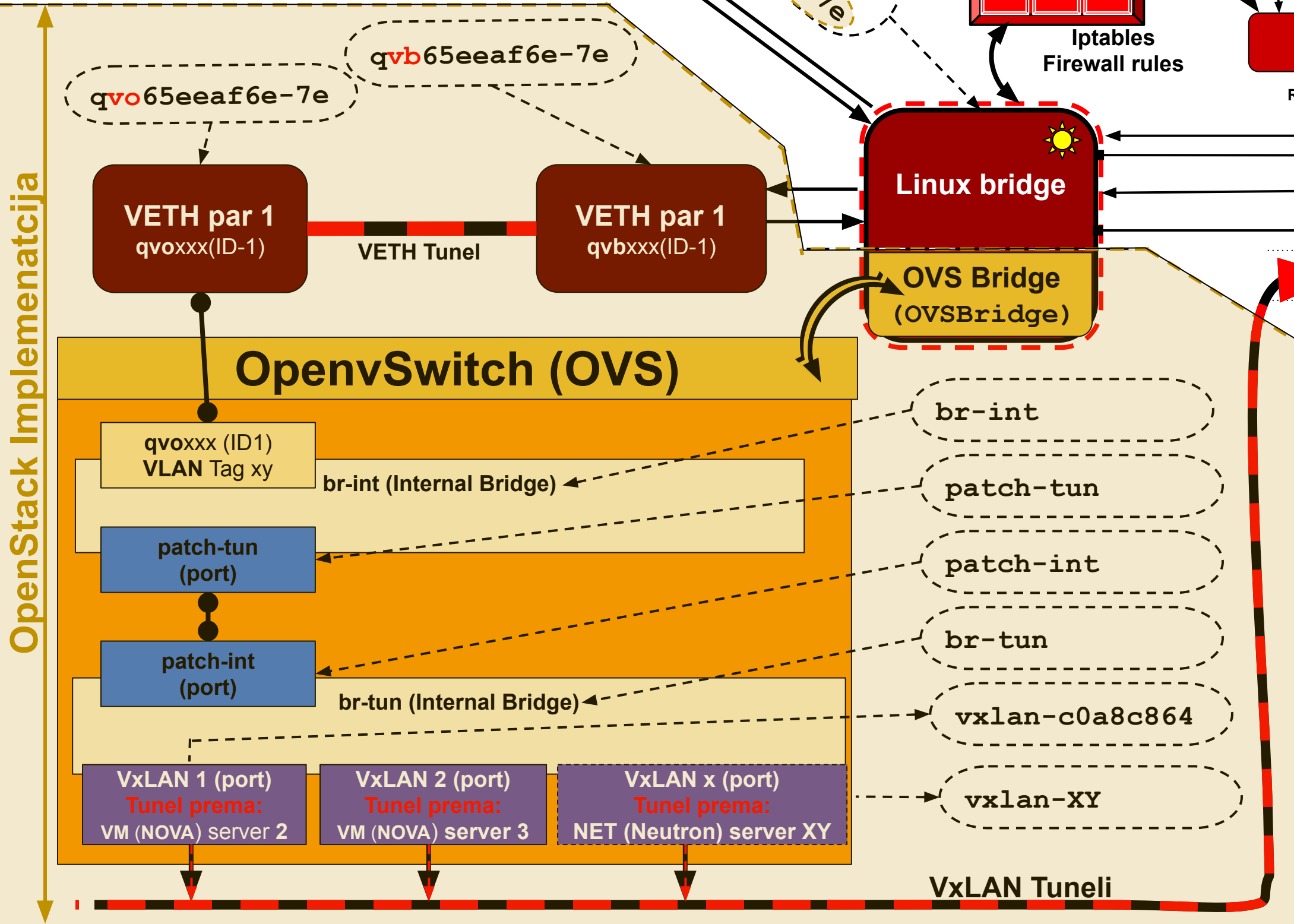
LXC kontejner

Linux unutar LXC kontejnera

- CGROUPS
- UTS namespace
- Process & PID namespace
- User namespace
- Mount namespace

VETH par 2
eth0@if13 (ID-2)

Network namespace

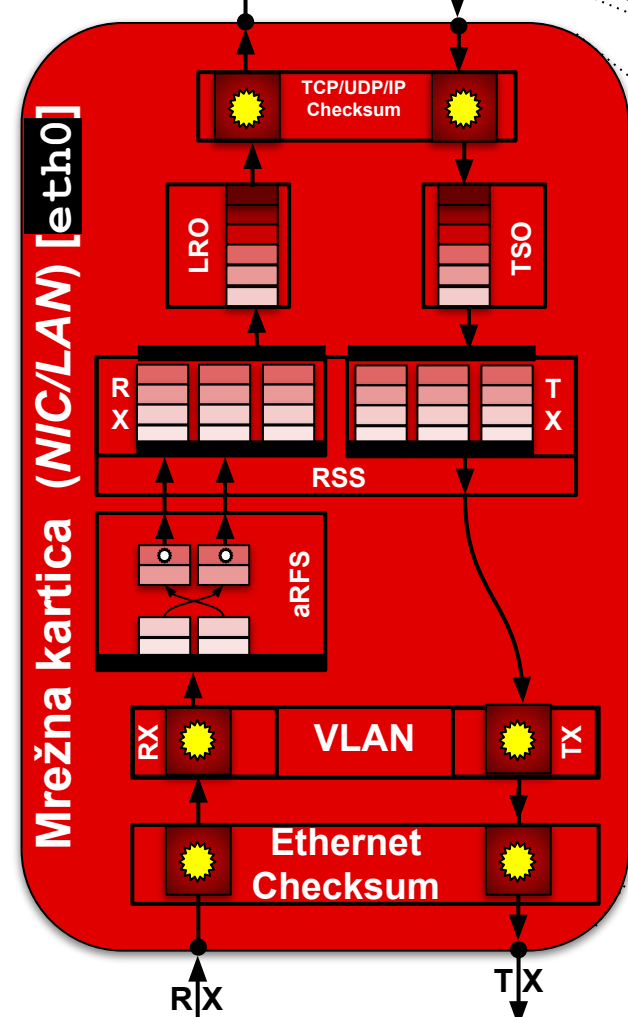
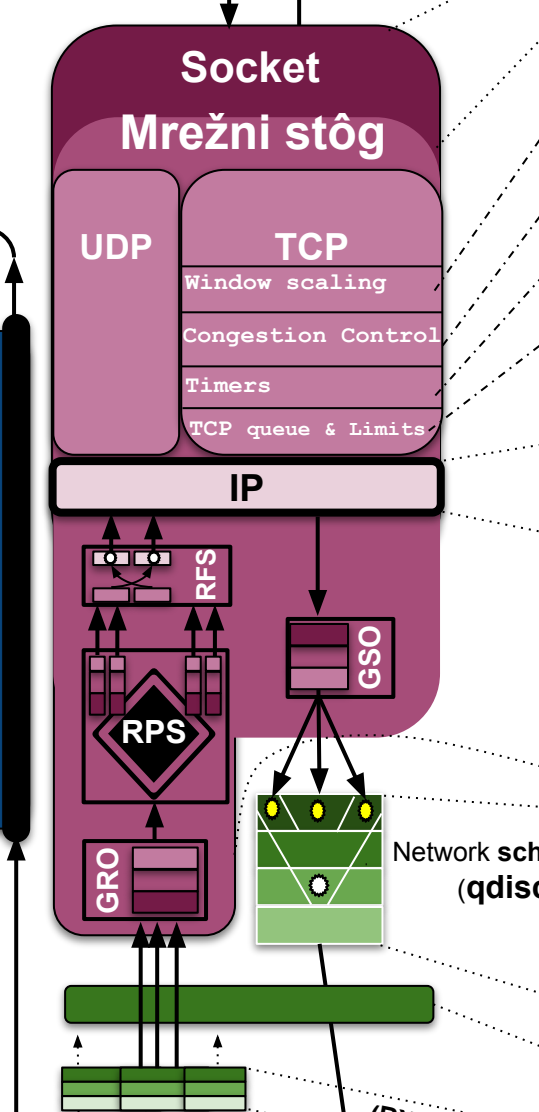


KVM (Hipervizor)

VirtIO BACKEND

Aplikacije/programi

Socket buffer



Sysctl varijable

TCP	UDP
<code>net.ipv4.tcp_mem</code>	<code>net.ipv4.udp_mem</code>
<code>net.ipv4.tcp_rmem</code>	<code>net.ipv4.udp_rmem_min</code>
<code>net.ipv4.tcp_wmem</code>	<code>net.ipv4.udp_wmem_min</code>

Skaliranje TCP prozora (uklj./tsklj) 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`

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`

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 niza 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. `code`, `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`

TXqueuen 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`

Ukupan broj paketa koji se mogu dohvatiti odjednom, zajednički za sva mrežna sučelja na sustavu:
`net.core.netdev_budget`

Vrijeme u μs dozvoljeno za dohvaćanje (pooling) 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 akcelrirane (ubrzane) funkcionalnosti mrežnih kartica (ako ih imaju):

Receive Side Scaling (RSS) - za povećanje RX i TX nizova 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 Linux bridge sučelja (pr. `br0`) ili OVS bridge sučelja!

VxLAN tunelima se povezuju računala (poslužitelji) među sobom, prema standardnoj implementaciji OpenStack-a!

GPL licenca
Hrvoje Horvat
v.1.01