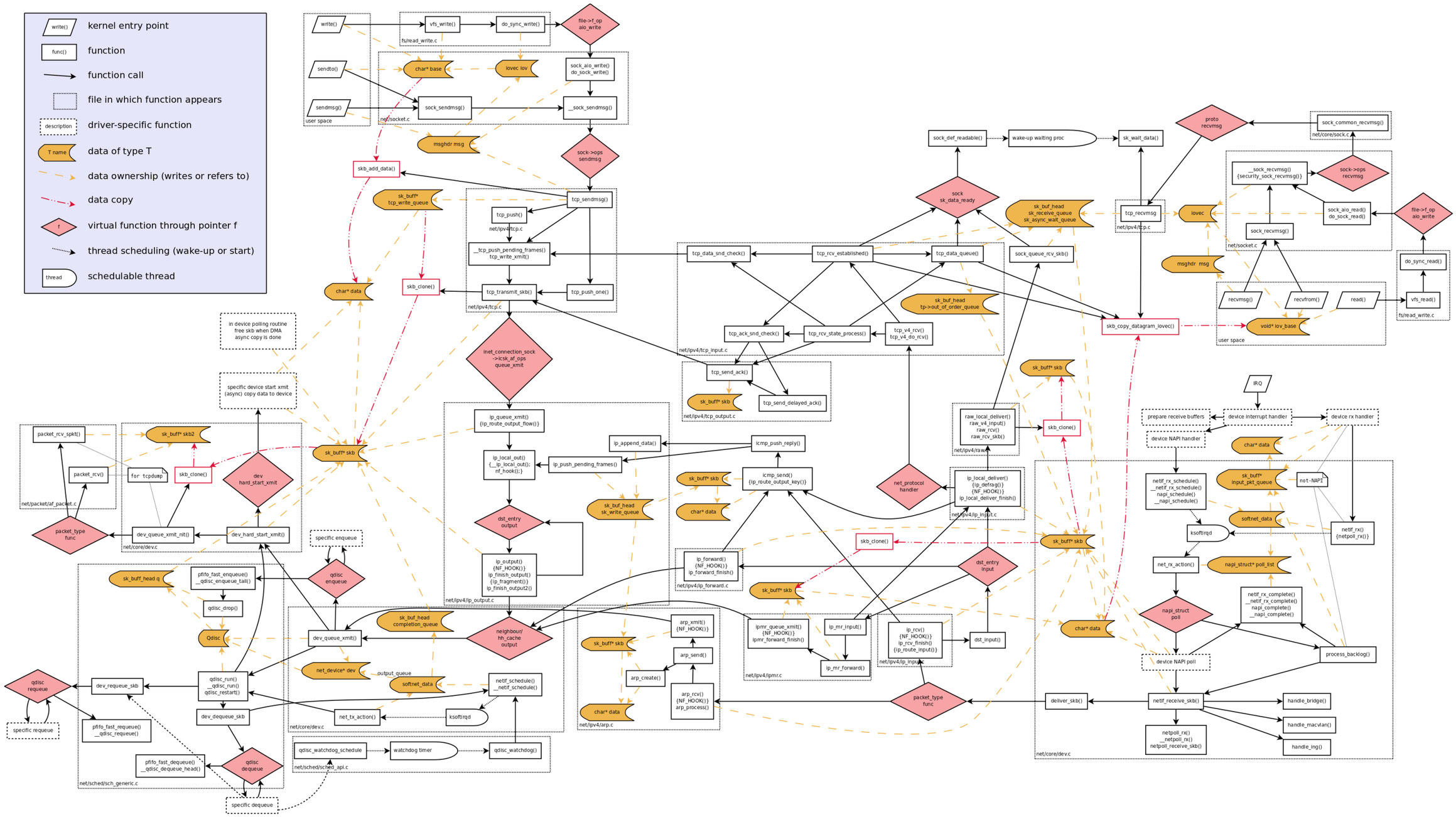
kernel entry point
function
function call
file in which function appears
driver-specific function
data of type T
data ownership (writes or refers to)
data copy
virtual function through pointer f
thread scheduling (wake-up or start)
schedulable thread
write()
vfs_write()
do_sync_write()
file->f_op aio_write
fs/read_write.c
sendto()
sendmsg()
user space
char* base
iovec iov
sock_aio_write() do_sock_write()
sock_sendmsg()
__sock_sendmsg()
net/socket.c
msghdr msg
sock->ops sendmsg
skb_add_data()
sk_buff* tcp_write_queue
tcp_sendmsg()
tcp_push()
net/ipv4/tcp.c
__tcp_push_pending_frames() tcp_write_xmit()
char* data
skb_clone()
tcp_transmit_skb()
tcp_push_one()
in device polling routine free skb when DMA async copy is done
inet_connection_sock ->icsk_af_ops queue_xmit
specific device start xmit (async) copy data to device
ip_queue_xmit() {ip_route_output_flow()}
ip_local_out() {__ip_local_out(); nf_hook();}
ip_push_pending_frames()
ip_append_data()
icmp_push_reply()
icmp_send() {ip_route_output_key()}
sk_buff* skb
char* data
sk_buff* skb2
packet_rcv_spkt()
packet_rcv()
for tcpdump
skb_clone()
dev hard_start_xmit
packet_type func
dev_queue_xmit_nit()
dev_hard_start_xmit()
net/packet/af_packet.c
net/core/dev.c
dst_entry output
sk_buff_head sk_write_queue
ip_output() {NF_HOOK()} ip_finish_output() {ip_fragment()} ip_finish_output2()
net/ipv4/ip_output.c
specific enqueue
qdisc enqueue
sk_buff_head q
pfifo_fast_enqueue() __qdisc_enqueue_tail()
qdisc_drop()
Qdisc
sk_buff_head completion_queue
dev_queue_xmit()
neighbour/ hh_cache output
net_device* dev
output_queue
softnet_data
netif_schedule() __netif_schedule()
net_tx_action()
ksoftirqd
net/core/dev.c
qdisc requeue
dev_requeue_skb
qdisc_run() __qdisc_run() qdisc_restart()
dev_dequeue_skb
specific requeue
pfifo_fast_requeue() __qdisc_requeue()
pfifo_fast_dequeue() __qdisc_dequeue_head()
qdisc dequeue
specific dequeue
net/sched/sch_generic.c
qdisc_watchdog_schedule
watchdog timer
qdisc_watchdog()
net/sched/sched_api.c
tcp_data_snd_check()
tcp_rcv_established()
tcp_data_queue()
tcp_ack_snd_check()
tcp_rcv_state_process()
tcp_v4_rcv() tcp_v4_do_rcv()
sk_buff_head tp->out_of_order_queue
tcp_send_ack()
tcp_send_delayed_ack()
net/ipv4/tcp_input.c
net/ipv4/tcp_output.c
sock_def_readable()
wake-up waiting proc
sk_wait_data()
sock sk_data_ready
sk_buff_head sk_receive_queue sk_async_wait_queue
sock_queue_rcv_skb()
tcp_recvmsg
net/ipv4/tcp.c
lovec
proto recvmsg
sock_common_recvmsg()
net/core/sock.c
__sock_recvmsg() {security_sock_recvmsg()}
sock->ops recvmsg
sock_aio_read() do_sock_read()
file->f_op aio_write
sock_recvmsg()
net/socket.c
msghdr msg
do_sync_read()
vfs_read()
fs/read_write.c
recvmsg()
recvfrom()
read()
void* iov_base
user space
skb_copy_datagram_iovec()
raw_local_deliver() raw_v4_input() raw_rcv() raw_rcv_skb()
net/ipv4/raw.c
net_protocol handler
ip_local_deliver() {ip_defrag()} {NF_HOOK()} ip_local_deliver_finish()
net/ipv4/ip_input.c
ip_forward() {NF_HOOK()} ip_forward_finish()
net/ipv4/ip_forward.c
dst_entry input
ipmr_queue_xmit() {NF_HOOK()} ipmr_forward_finish()
ip_mr_input()
ip_mr_forward()
net/ipv4/ipmr.c
ip_rcv() {NF_HOOK()} ip_rcv_finish() {ip_route_input()}
dst_input()
net/ipv4/ip_input.c
packet_type func
arp_xmit() {NF_HOOK()}
arp_send()
arp_create()
arp_rcv() {NF_HOOK()} arp_process()
net/ipv4/arp.c
IRQ
prepare receive buffers
device interrupt handler
device rx handler
device NAPI handler
netif_rx_schedule() __netif_rx_schedule() napi_schedule() __napi_schedule()
sk_buff* input_pkt_queue
not-NAPI
softnet_data
ksoftirqd
netif_rx() {netpoll_rx()}
net_rx_action()
napi_struct* poll_list
napi_struct poll
netif_rx_complete() __netif_rx_complete() napi_complete() __napi_complete()
device NAPI poll
process_backlog()
deliver_skb()
netif_receive_skb()
handle_bridge()
handle_macvlan()
handle_ing()
netpoll_rx() __netpoll_rx() netpoll_receive_skb()
net/core/dev.c