Table of Contents

Table of Figures

UDP Process Timing Test Information

The tests used netperf to netserver UDP_RR data transfer to provide dataflow for collecting network process timing data. Timing test data was collected on the side running netperf.

The timing data was collected by logging the current processor clock cycles of a moderate number of locations using Linux kernel get_cycles() function then calculating time and averaging over 4096 packets, then printing results and writing results to the kernel log. After the test the relevant kernel log data is saved and put into a spread sheet to calculate the average over the whole test. Many tests with measurements at different kernel locations with a complete kernel build and reboot were used to collect the full timing data set.

System Information for Tests

Linux Information

The Linux release kernel 2.6.36.2 was used as the base kernel with and flush kernel change patch. An IXGBE device driver patch to change the IXGBE device driver to the Intel low latency IXGBE test driver version LL84P (Low Latency ixgbe-2.0.84.9P+LL-tx+tx-Fdir) for the base test kernel. Each test used the base kernel with a test kernel patch to create the test kernel for each test.

For the test case of low latency polling disabled which is the base interrupt based test, the **Low Latency Receive Flush**, **Low Latency TX Flow Change** and **Low Latency TCP Receive Flush** are disabled in the Linux kernel configuration.

Network Interface Information

NIANTIC interface with interrupts affinitized on both sides.

Test for: Westmere

netperf Side

3.3 GHz Westmere 6 core, CPU Family 6, Model 44-X5680, Step 2 (B1 206C2) with 2 CPUs socketed, 12288KB cache, QPI 6.4 GT/s rate, Hyperthreading enabled.

1333 MHz memory bus, 12 GB memory (6 GB using 3 X 2GB memory modules per socket)

CPU Socket 0 core was used for test.

netserver Side: GC-NH5 test system

2.932 GHz Nehalem 4 core, CPU Family 6, Model 26 Nehalem B0/B1, Step 2 (106A2), with only 1 CPU socketed. 8192 KB cache, QPI 5.866 GT/s rate, Hyperthreading enabled.

1067 MHz memory bus, 6 GB memory (6 GB using 3 X 2GB memory modules for socket 0).

Test for: DELL Westmere

System is a DELL PowerEdge 710 production server.

netperf Side

2.93 GHz Westmere 6 core, CPU Family 6, Model 44-X5670, Step 2 with 2 CPUs socketed, 12288KB cache, Hyperthreading enabled.

12 GB memory installed.

No CPU Socket core was set for test.

netserver Side: GC-NH5 test system

2.932 GHz Nehalem 4 core, CPU Family 6, Model 26 Nehalem B0/B1, Step 2 (106A2), with only 1 CPU socketed. 8192 KB cache, QPI 5.866 GT/s rate, Hyperthreading enabled.

1067 MHz memory bus, 6 GB memory (6 GB using 3 X 2GB memory modules for socket 0).

Test for: Single Core DELL Westmere

System is a DELL PowerEdge 710 production server.

System is same as DELL Westmere, but with SMP disabled in kernel build for single core, non-SMP timing data. This test provides for comparing non-SMP network data timing with SMP data timing.

netperf Side

2.93 GHz Westmere 6 core, CPU Family 6, Model 44-X5670, Step 2 with 2 CPUs socketed, 12288KB cache, single core non-SMP kernel build.

12 GB memory installed.

No CPU Socket core was set for test.

netserver Side: GC-NH5 test system

2.932 GHz Nehalem 4 core, CPU Family 6, Model 26 Nehalem B0/B1, Step 2 (106A2), with only 1 CPU socketed. 8192 KB cache, QPI 5.866 GT/s rate, Hyperthreading enabled.

1067 MHz memory bus, 6 GB memory (6 GB using 3 X 2GB memory modules for socket 0).

Test for: 2S Nehalem

2S is for two CPU socketed in system. Data, data for same system with one CPU pulled out was collected (but not included in the diagrams as of this time).

netperf Side

3.07 GHz Nehalem 4 core, CPU Family 6, Model 26, Step 2 with 2 CPUs socketed, 8192 KB cache, Hyperthreading enabled.

Netperf was not affinitized to any CPU Socket.

netserver Side: GC-NH5 test system

2.932 GHz Nehalem 4 core, CPU Family 6, Model 26 Nehalem B0/B1, Step 2 (106A2), with only 1 CPU socketed. 8192 KB cache, QPI 5.866 GT/s rate, Hyperthreading enabled.

1067 MHz memory bus, 6 GB memory (6 GB using 3 X 2GB memory modules for socket 0).

UDP Process Timing Range across Multiple Processors

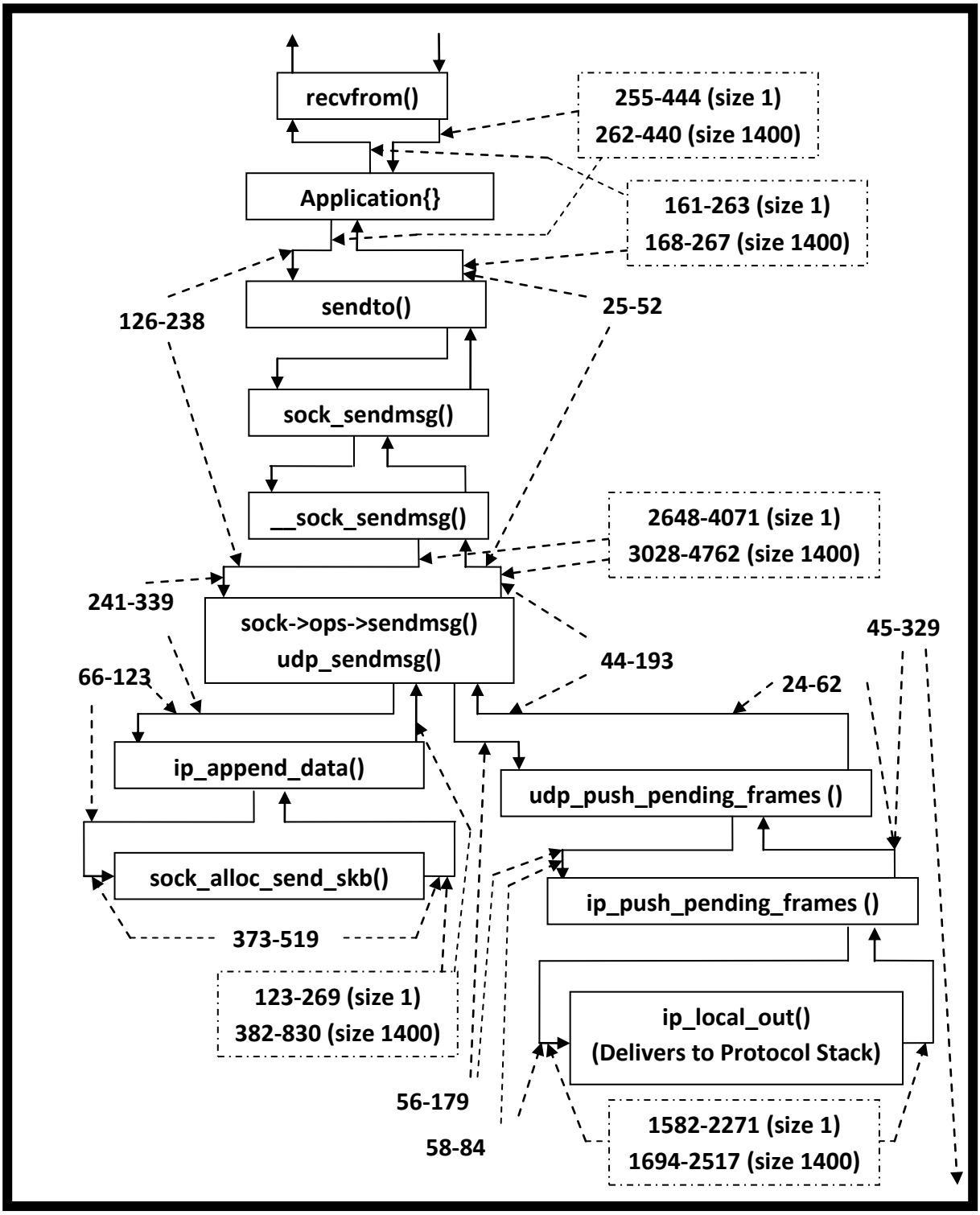


Figure 1: Multiple Systems SMP Socket & UDP TX Processing above IP Stack

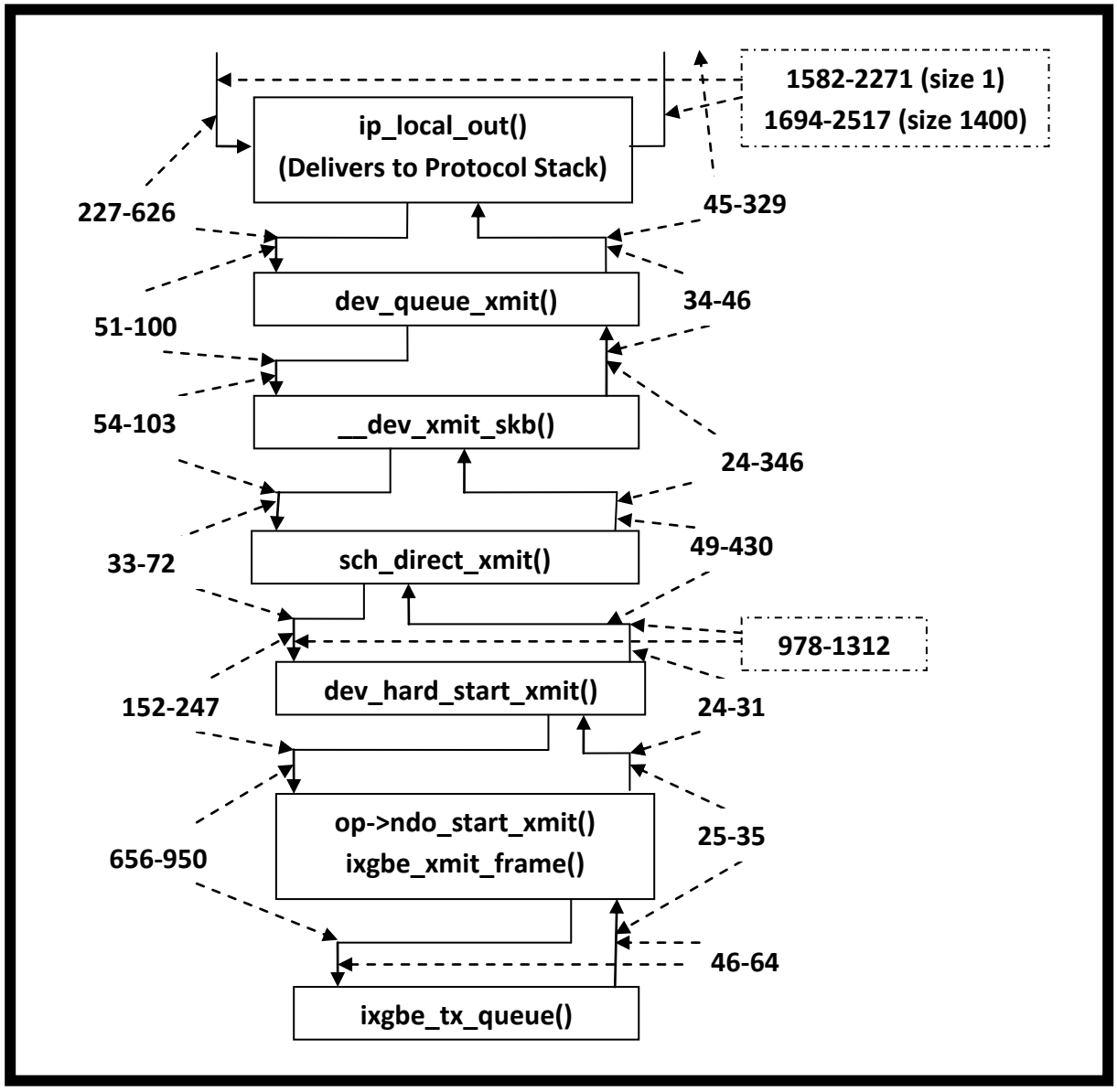


Figure 2: Multiple Systems SMP UDP TX Processing Below IP Stack

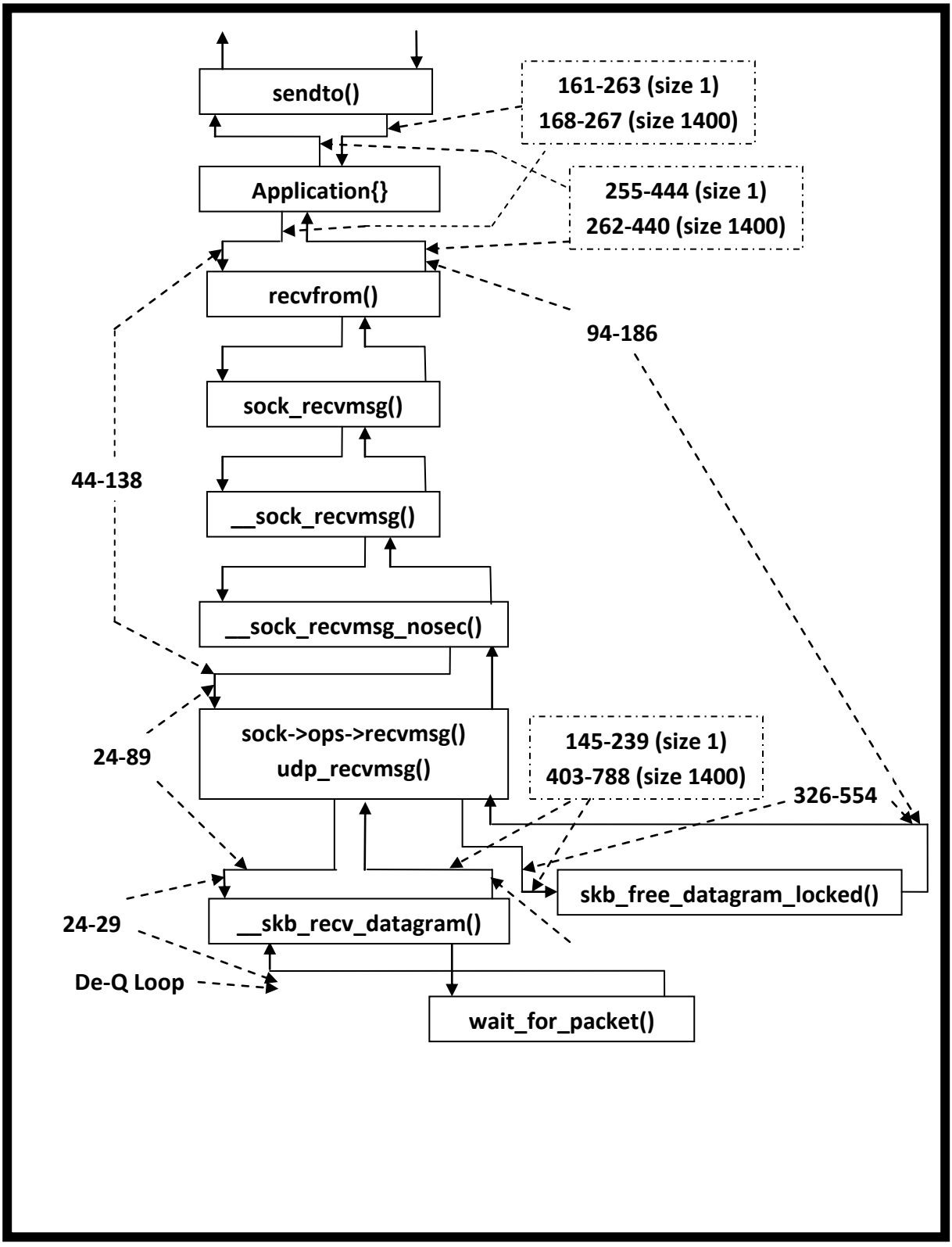


Figure 3: Multiple Systems SMP Socket & UDP RX Processing above IP Stack

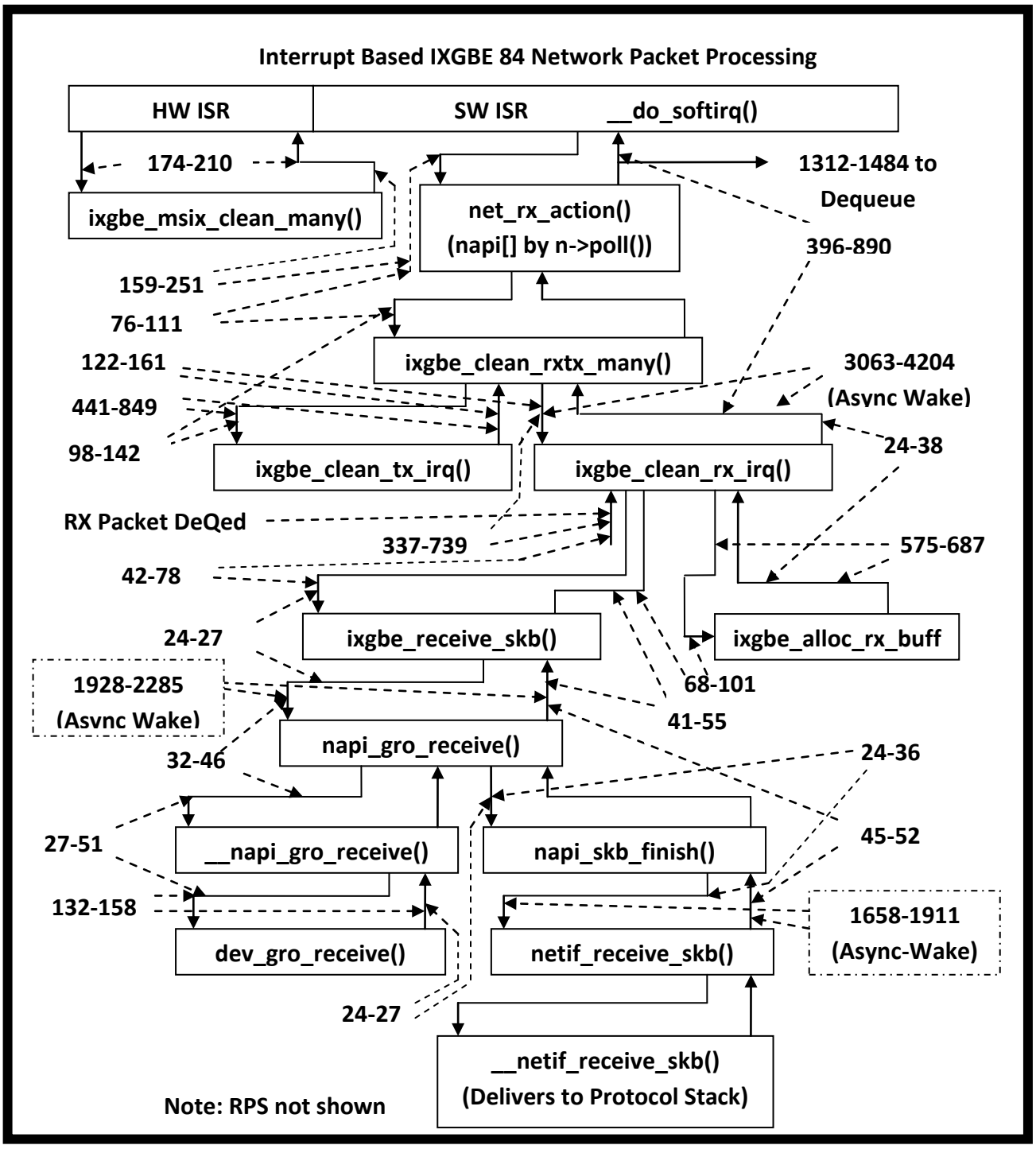


Figure 4: Multiple Systems SMP IXGBE Interrupt Driven UDP RX Processing Below IP Stack

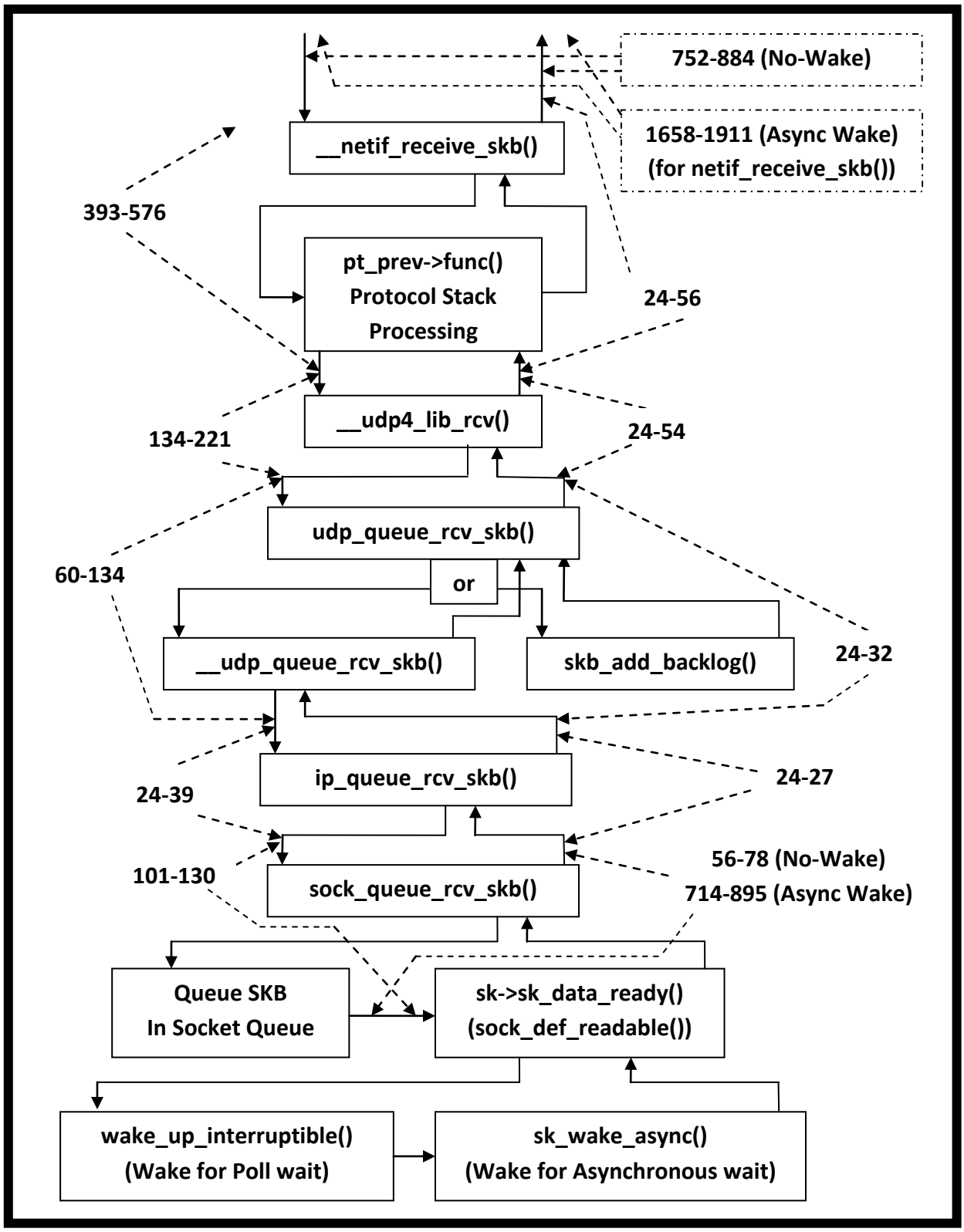


Figure 5: Multiple Systems SMP RX Protocol Stack Processing of UDP above IP Stack

UDP Process Timing Range for Westmere Processor

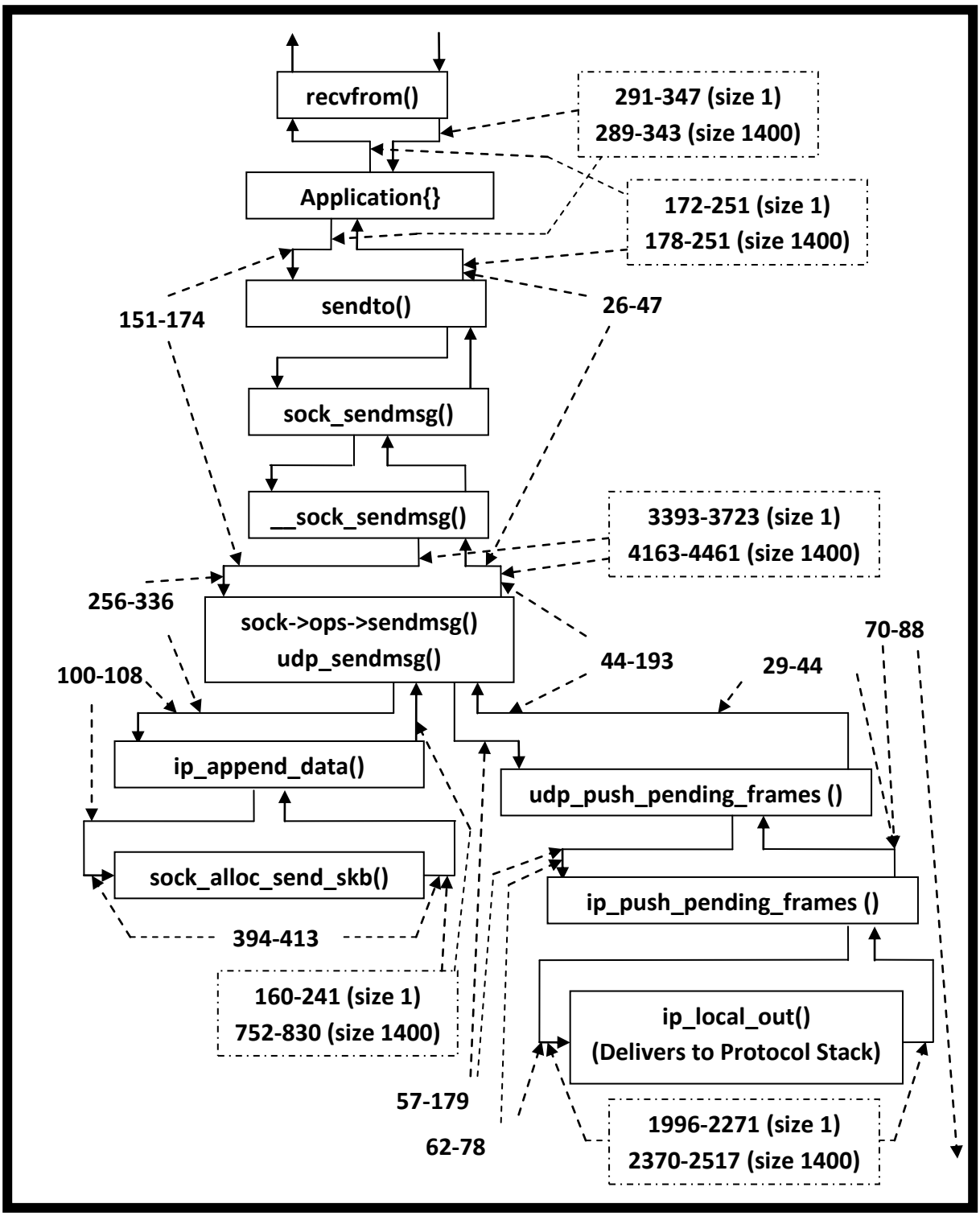


Figure 6: Westmere SMP Socket & UDP TX Processing above IP Stack

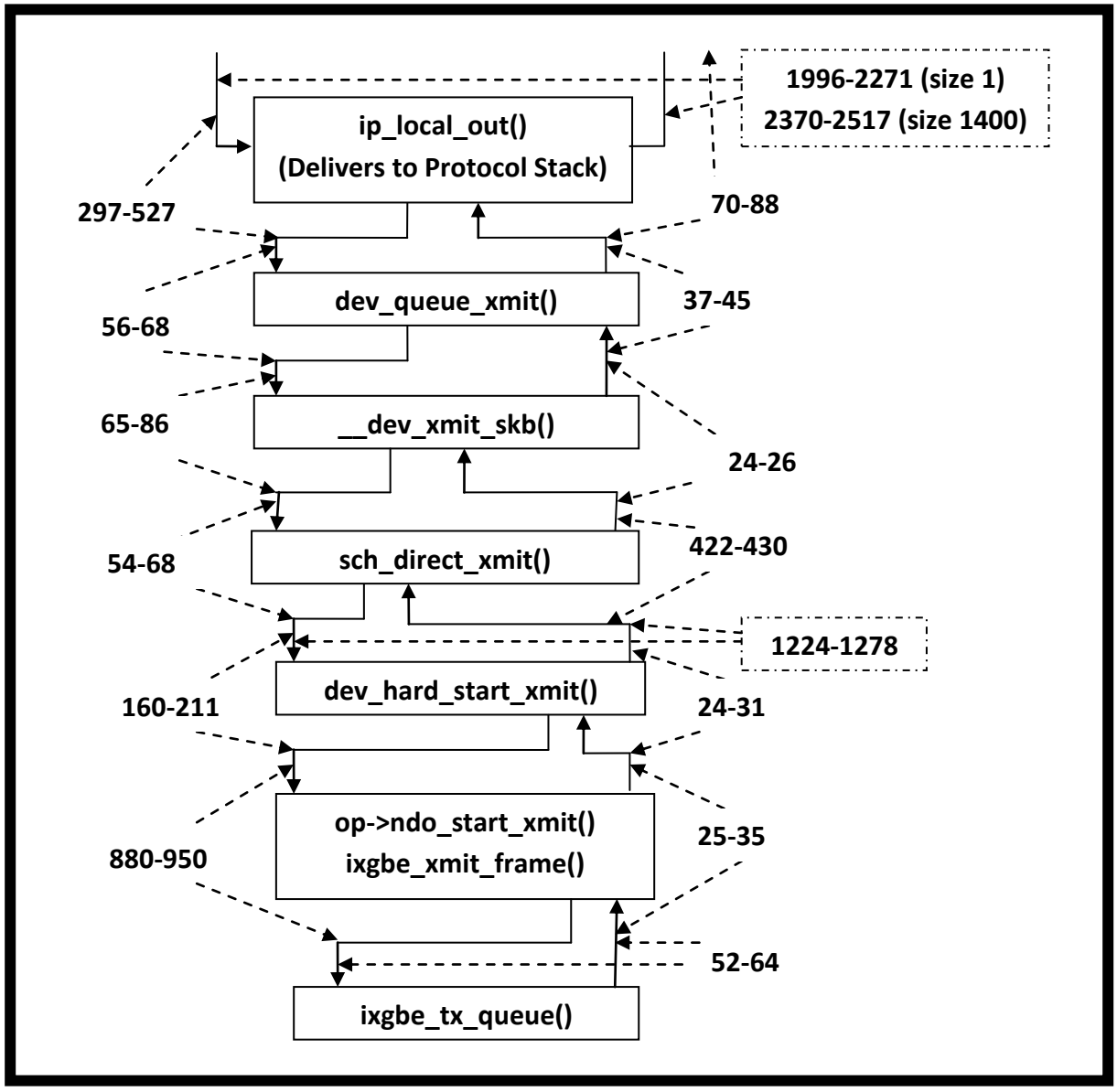


Figure 7: Westmere SMP UDP TX Processing Below IP Stack

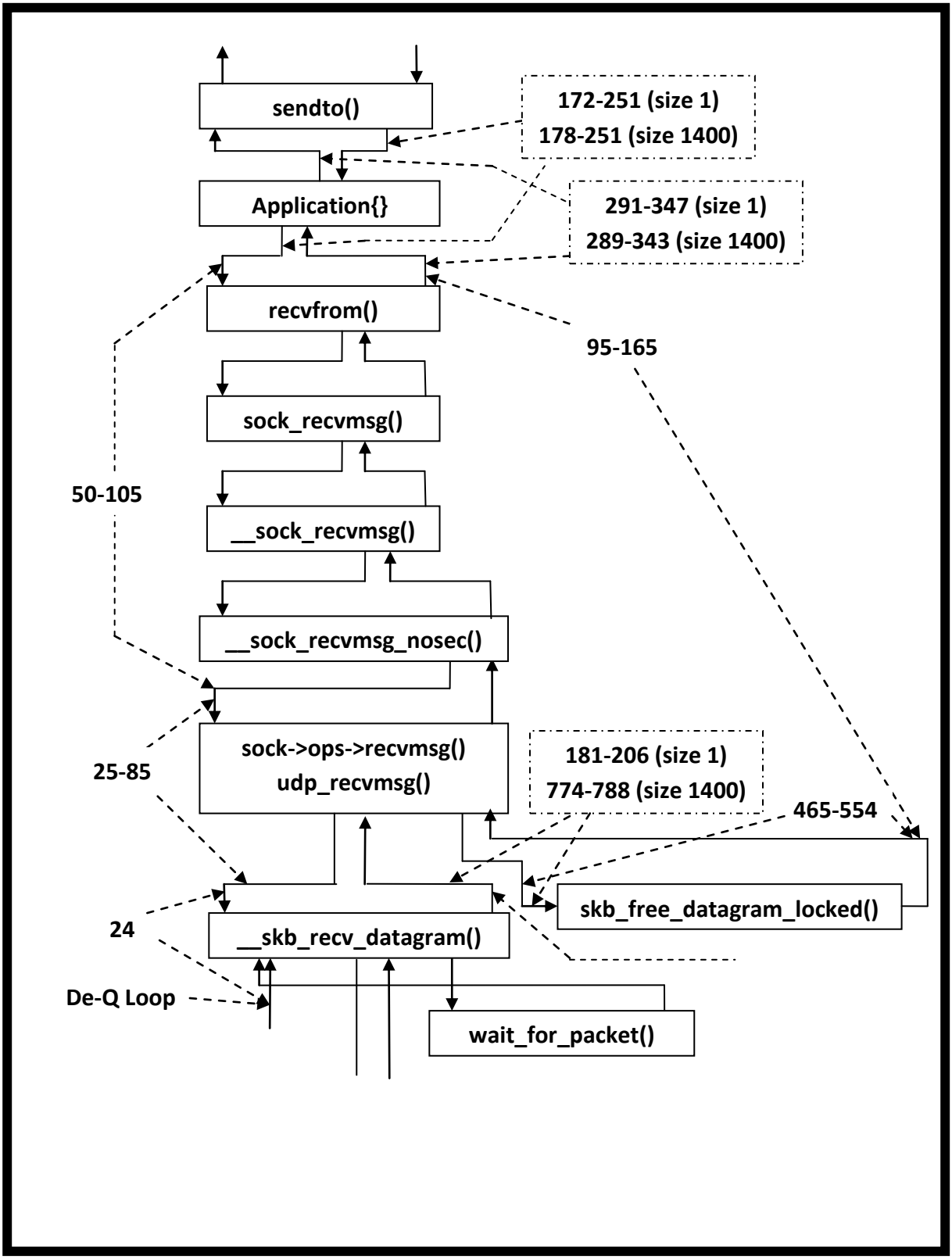


Figure 8: Westmere SMP Socket & UDP RX Processing above IP Stack

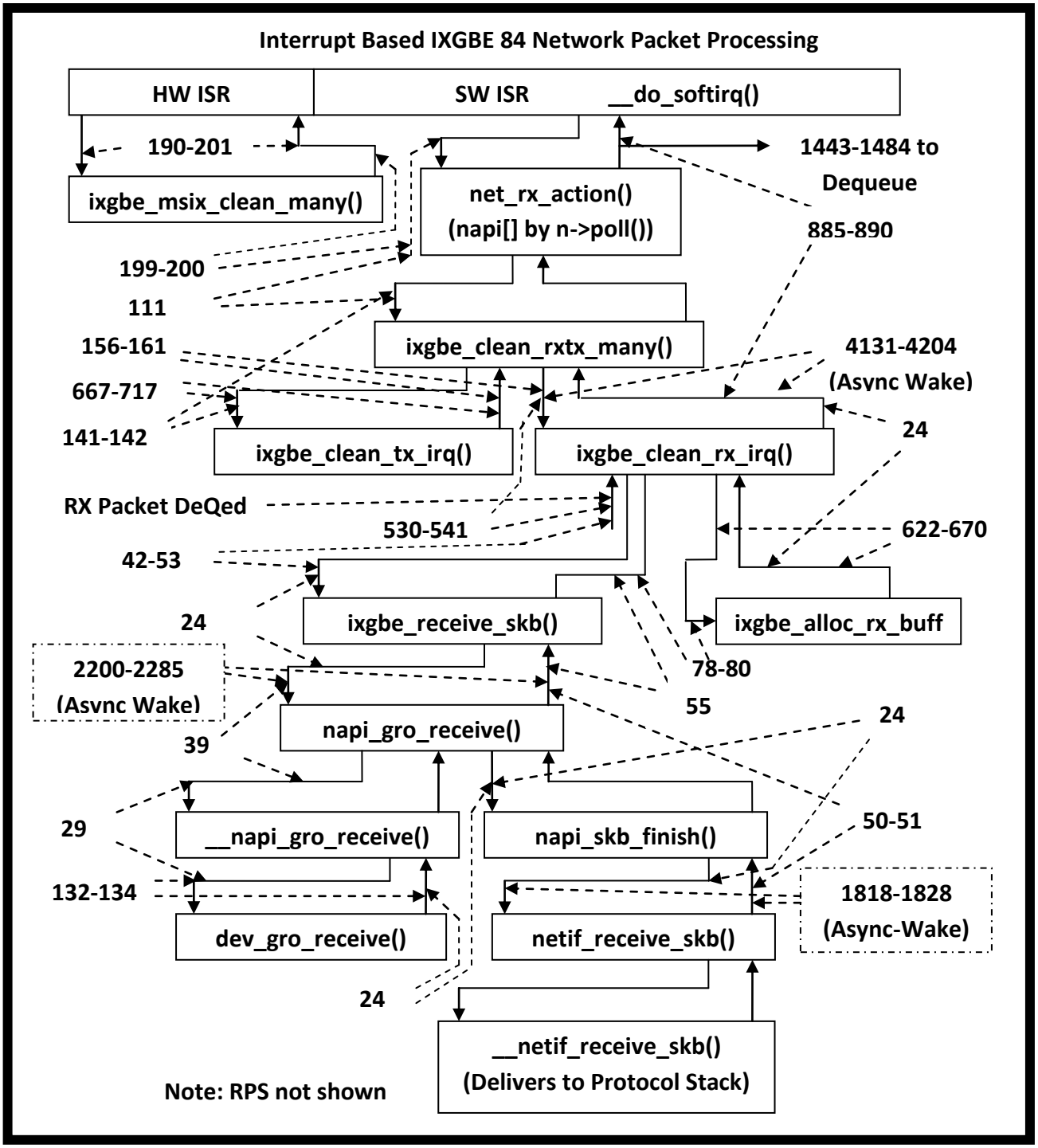


Figure 9: Westmere SMP IXGBE Interrupt Driven UDP RX Processing Below IP Stack

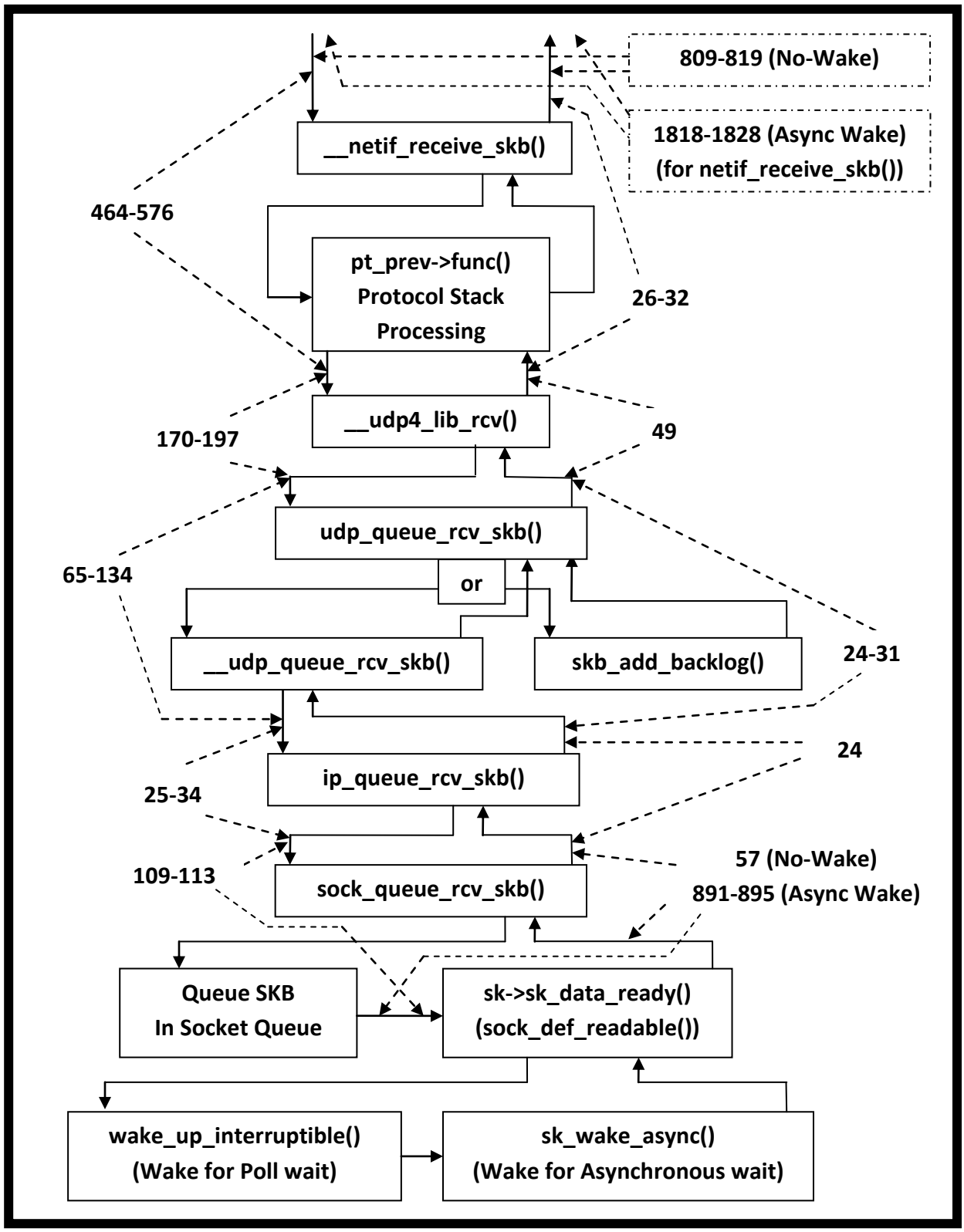


Figure 10: Westmere SMP RX Protocol Stack Processing of UDP above IP Stack

UDP Process Timing for Non-SMP Single Core Westmere CPU

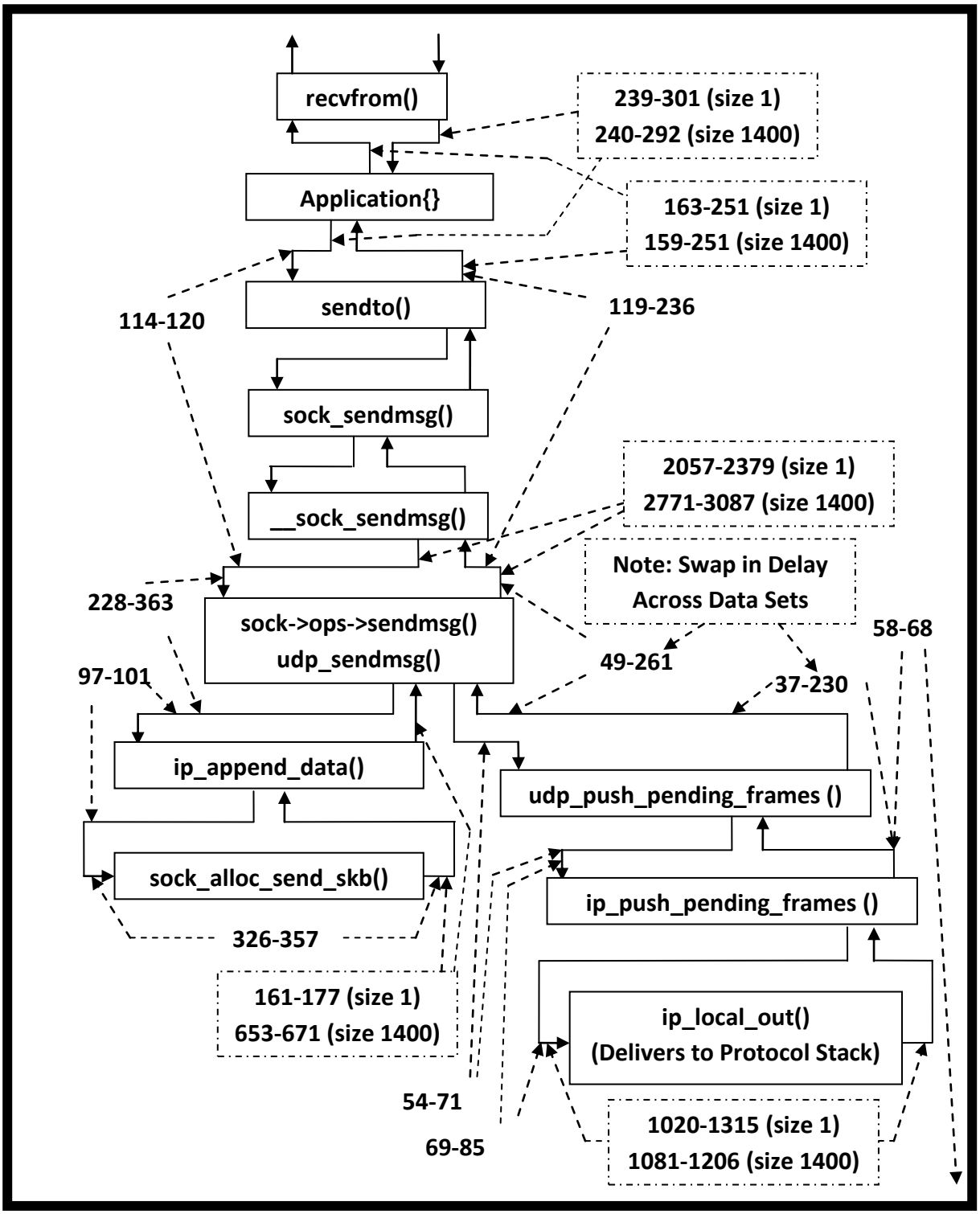


Figure 11: Single Core Dell Westmere Socket & UDP TX Processing above IP Stack

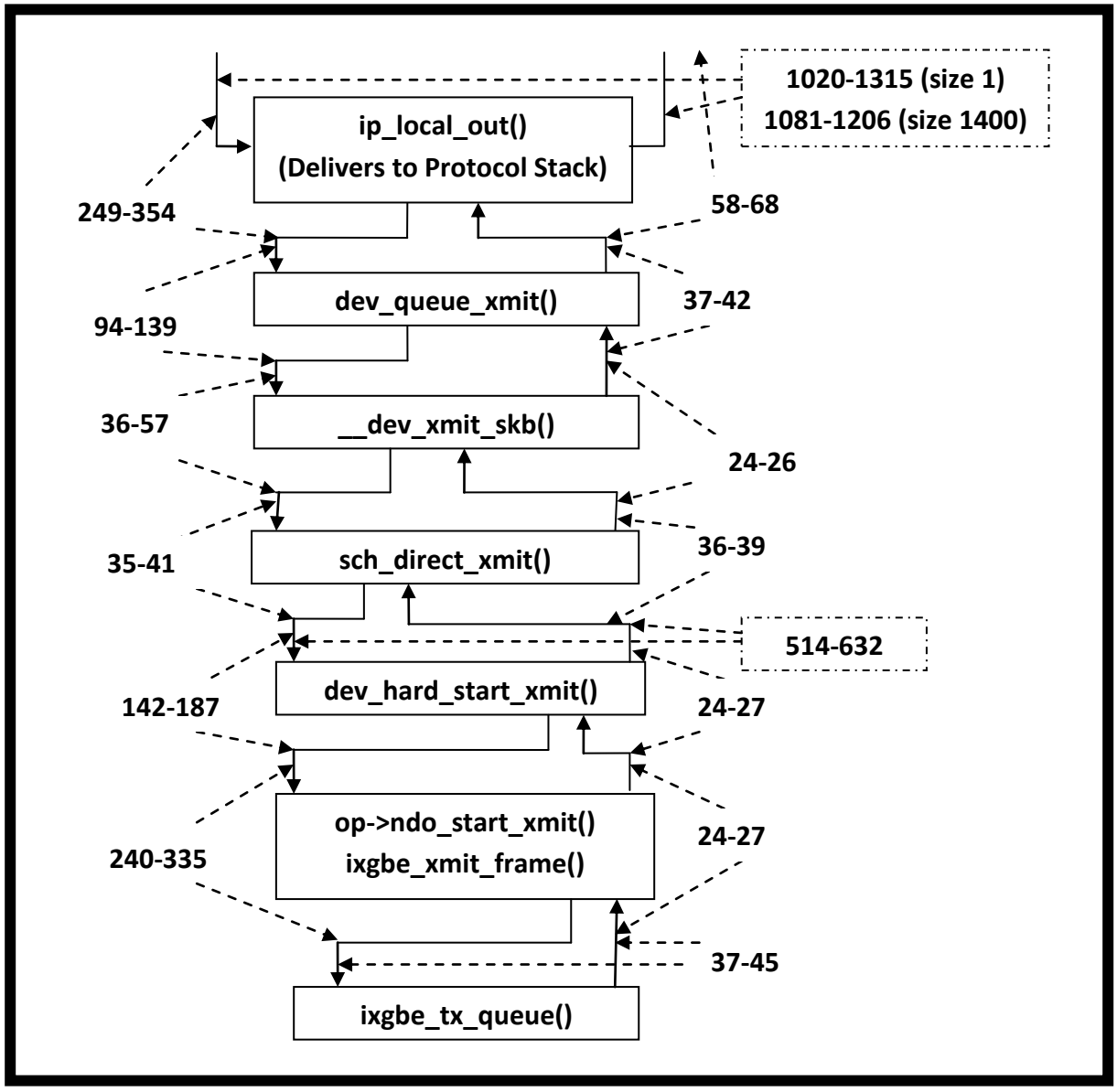


Figure 12: Single Core Dell Westmere UDP TX Processing Below IP Stack

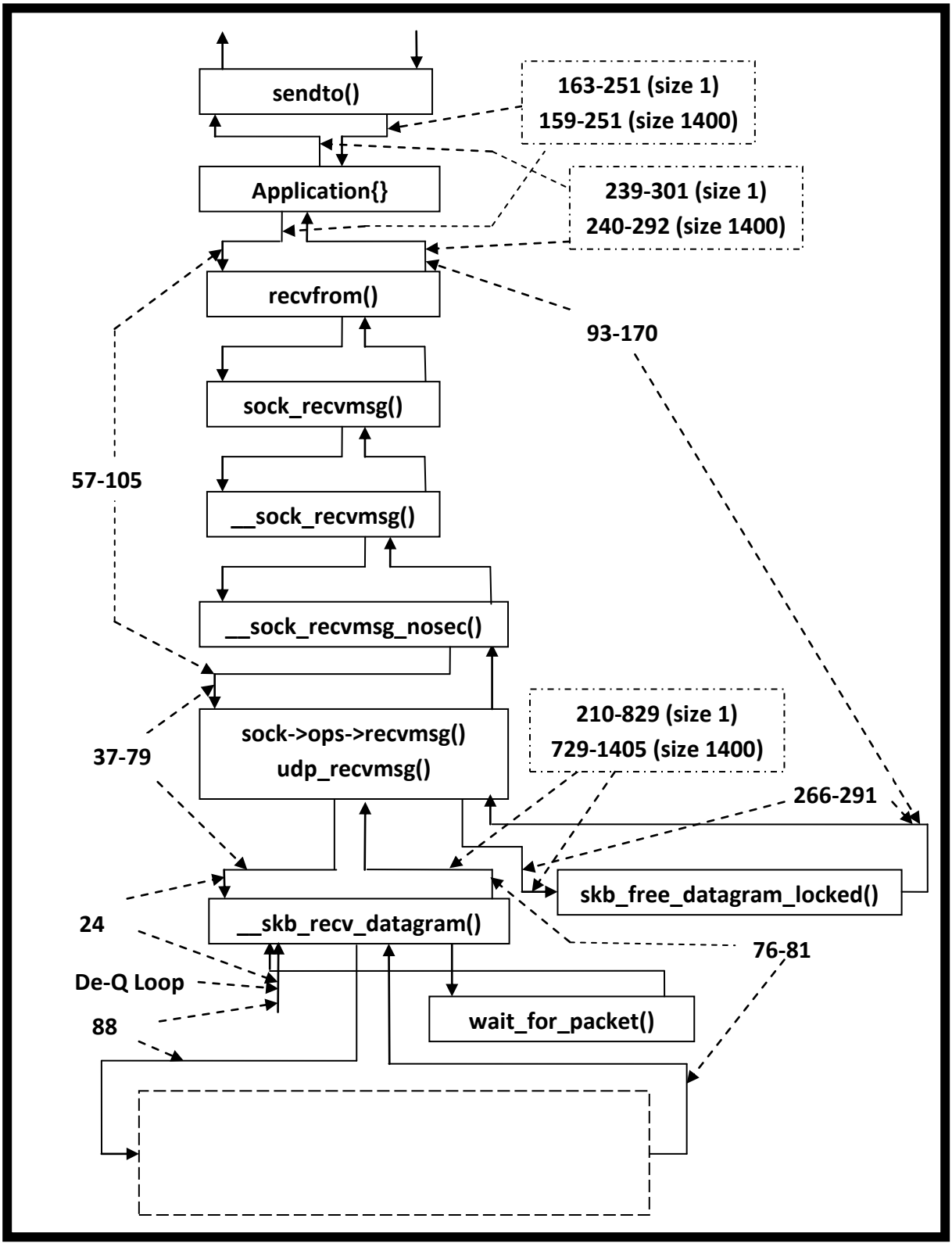


Figure 13: Single Core Dell Westmere Socket & UDP RX Processing above IP Stack

Note: Non-SMP Interrupt Processing different interrupt and NAPI subroutines then SMP processing.

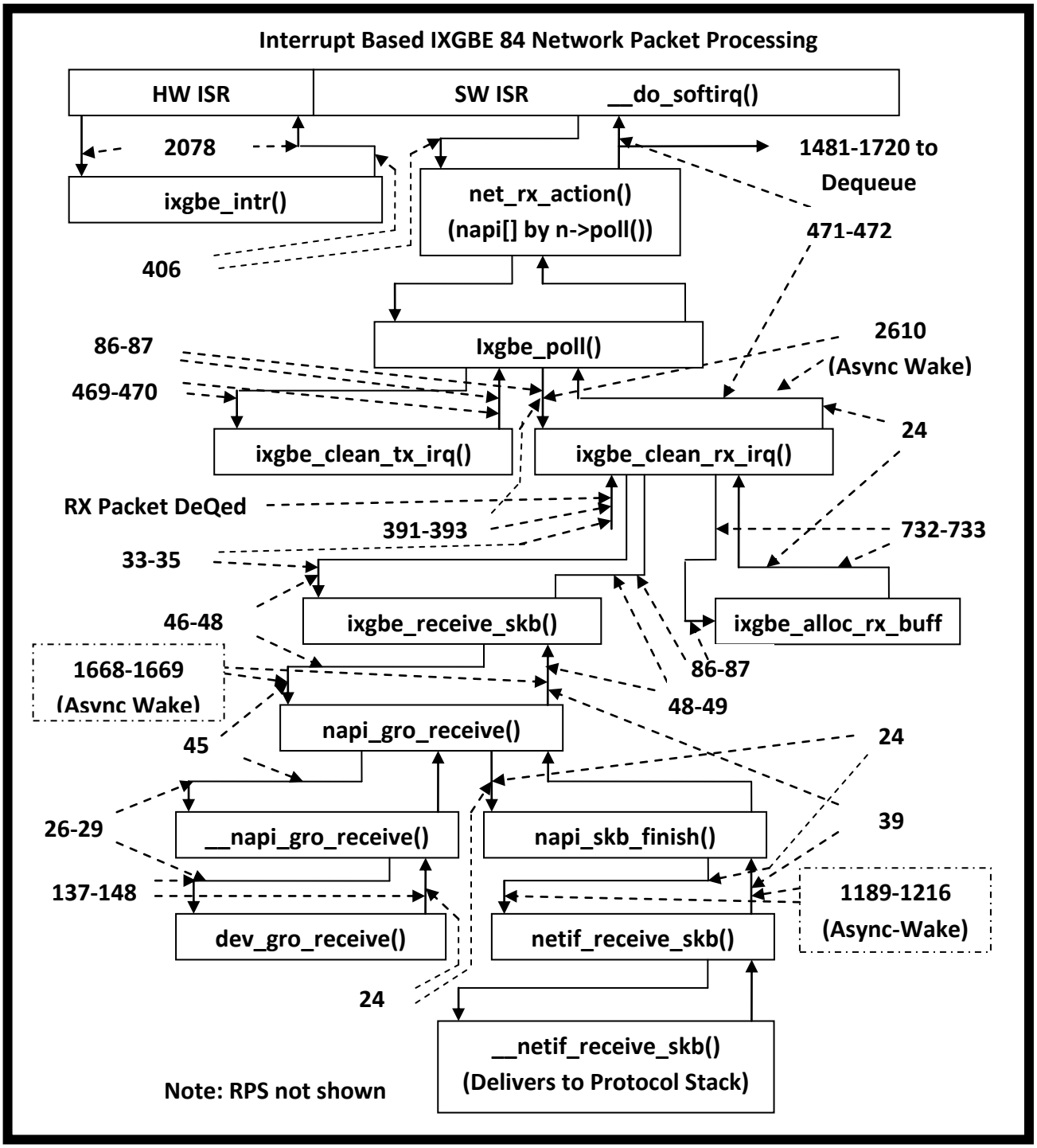


Figure 14: Single Core Dell Westmere IXGBE Interrupt Driven UDP RX Processing Below IP Stack

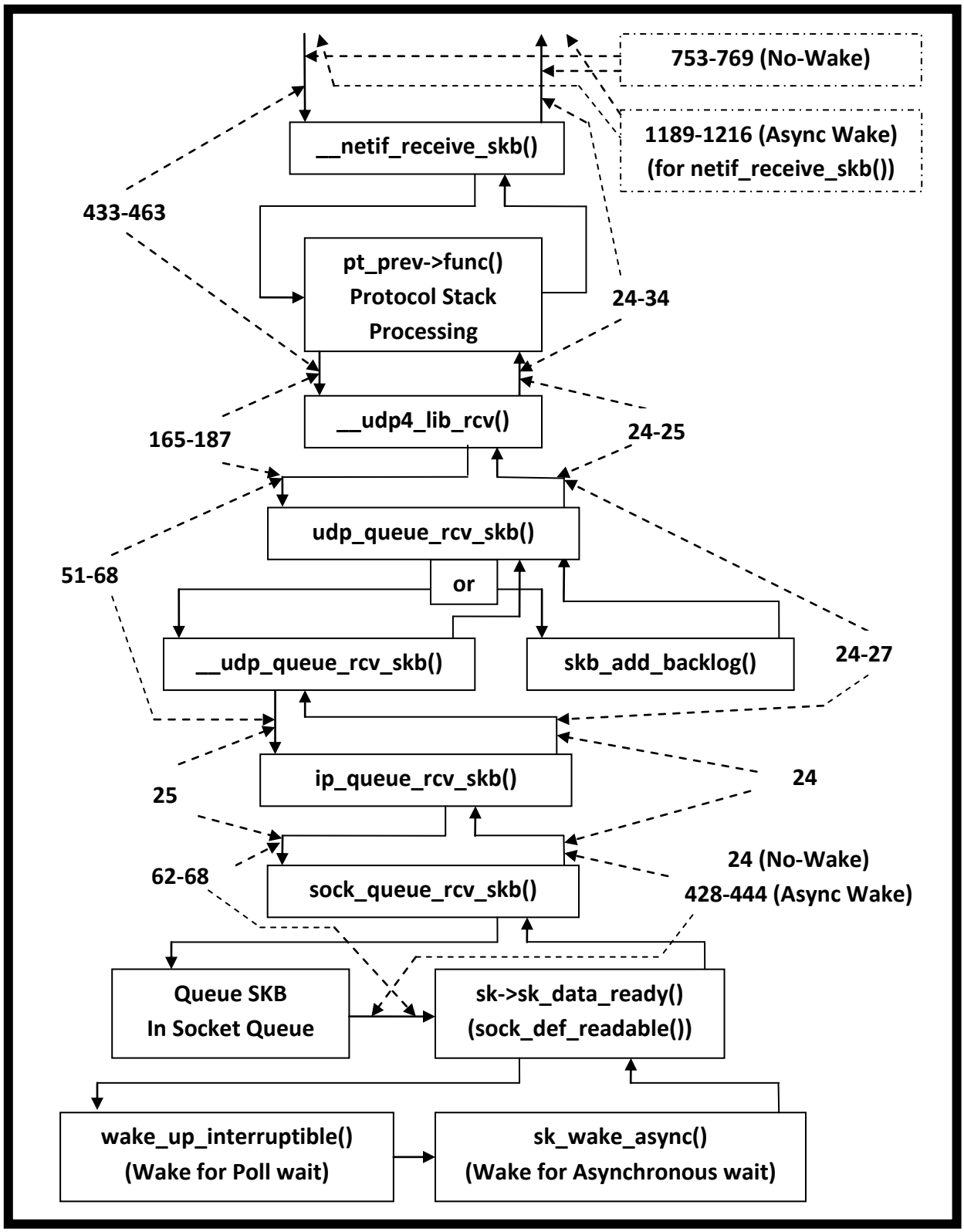
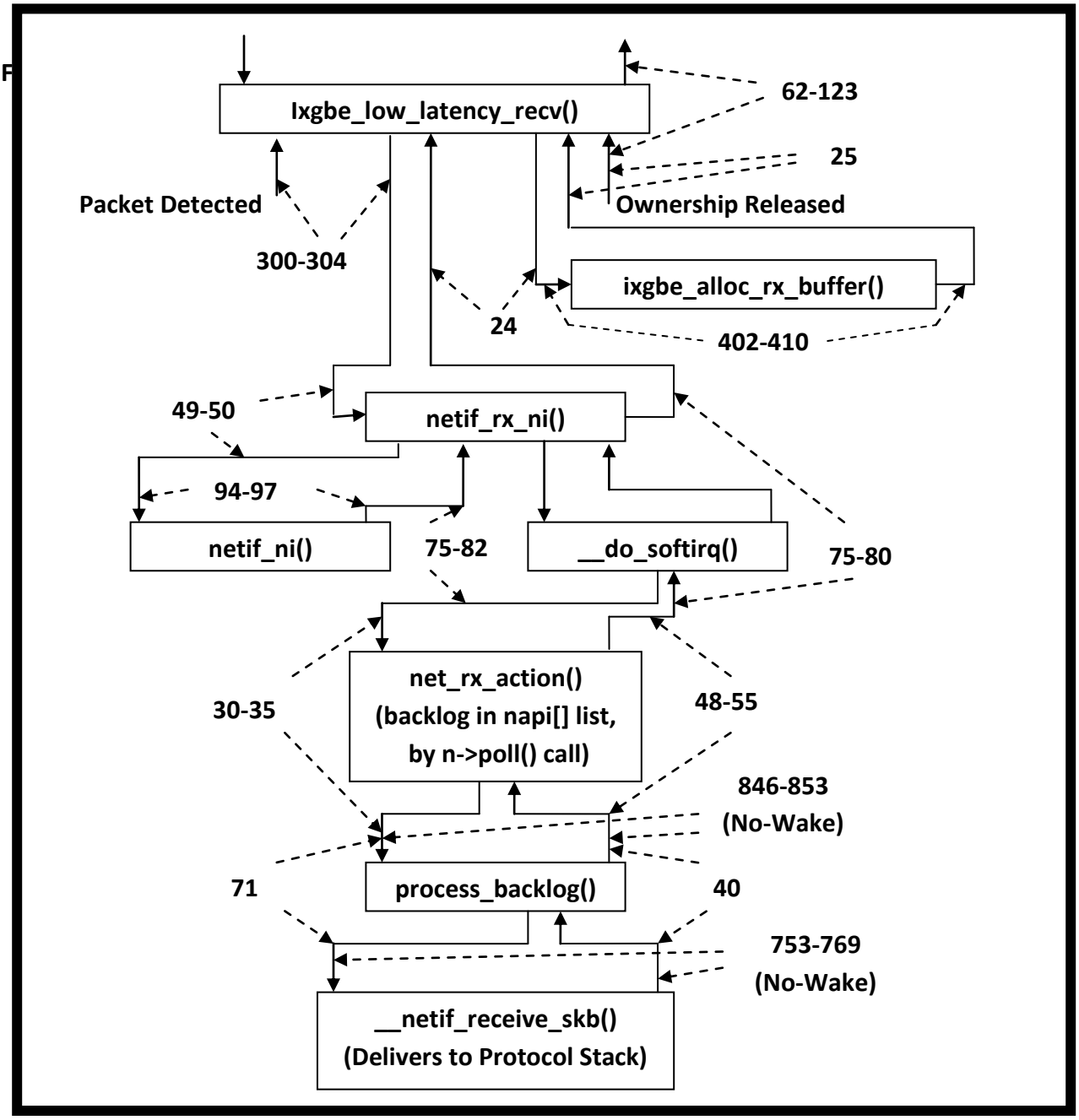


Figure 15: Single Core Dell Westmere RX Protocol Stack Processing of UDP above IP Stack


F
Ixgbe_low_latency_recv()
62-123
25
Packet Detected
Ownership Released
300-304
ixgbe_alloc_rx_buffer()
24
402-410
49-50
netif_rx_ni()
94-97
netif_ni()
75-82
__do_softirq()
75-80
net_rx_action()
(backlog in napi[] list,
by n->poll() call)
30-35
48-55
846-853
(No-Wake)
71
process_backlog()
40
753-769
(No-Wake)
__netif_receive_skb()
(Delivers to Protocol Stack)

UDP Process Timing Range for Dell Production Westmere

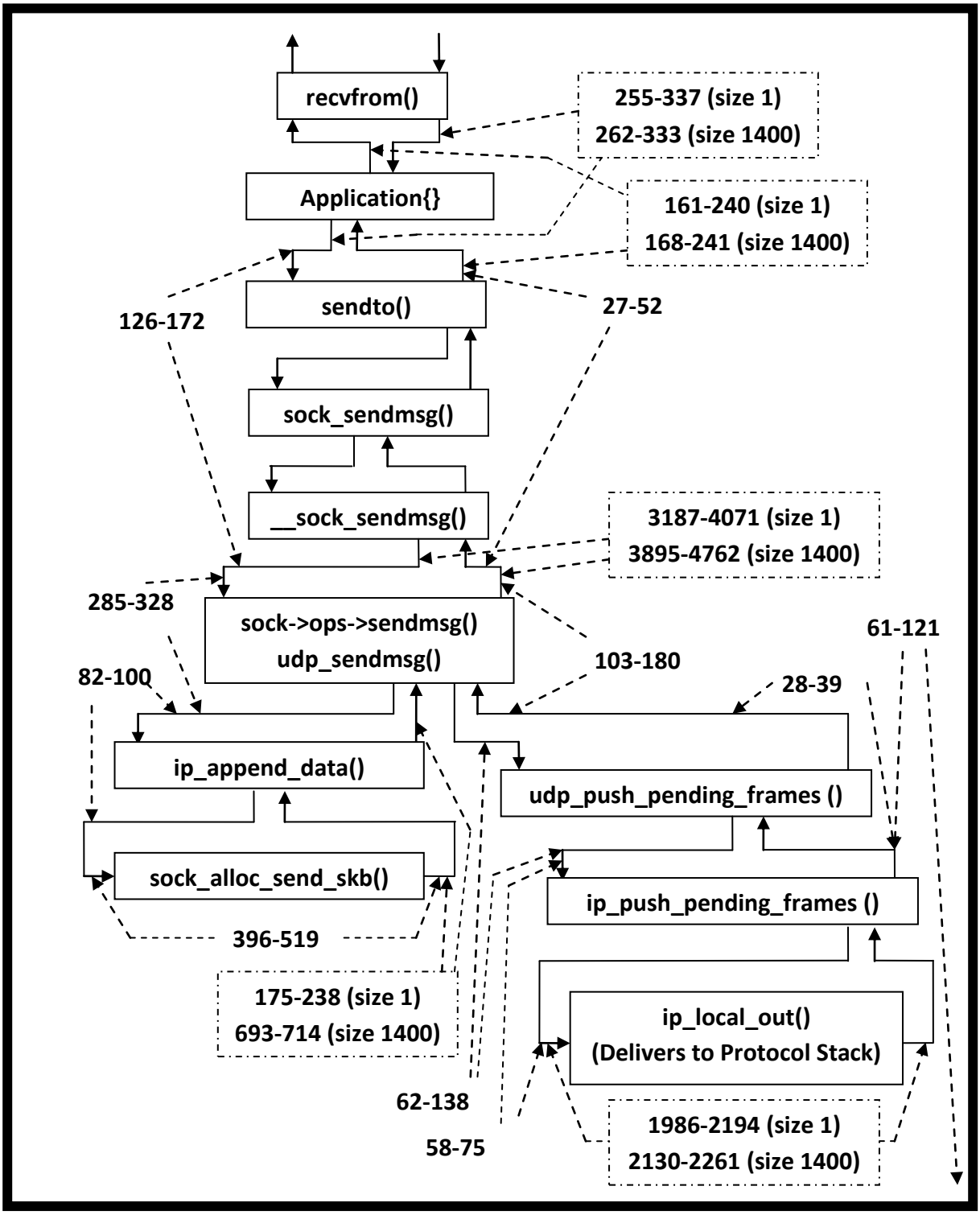


Figure 17: Dell Westmere SMP Socket & UDP TX Processing above IP Stack

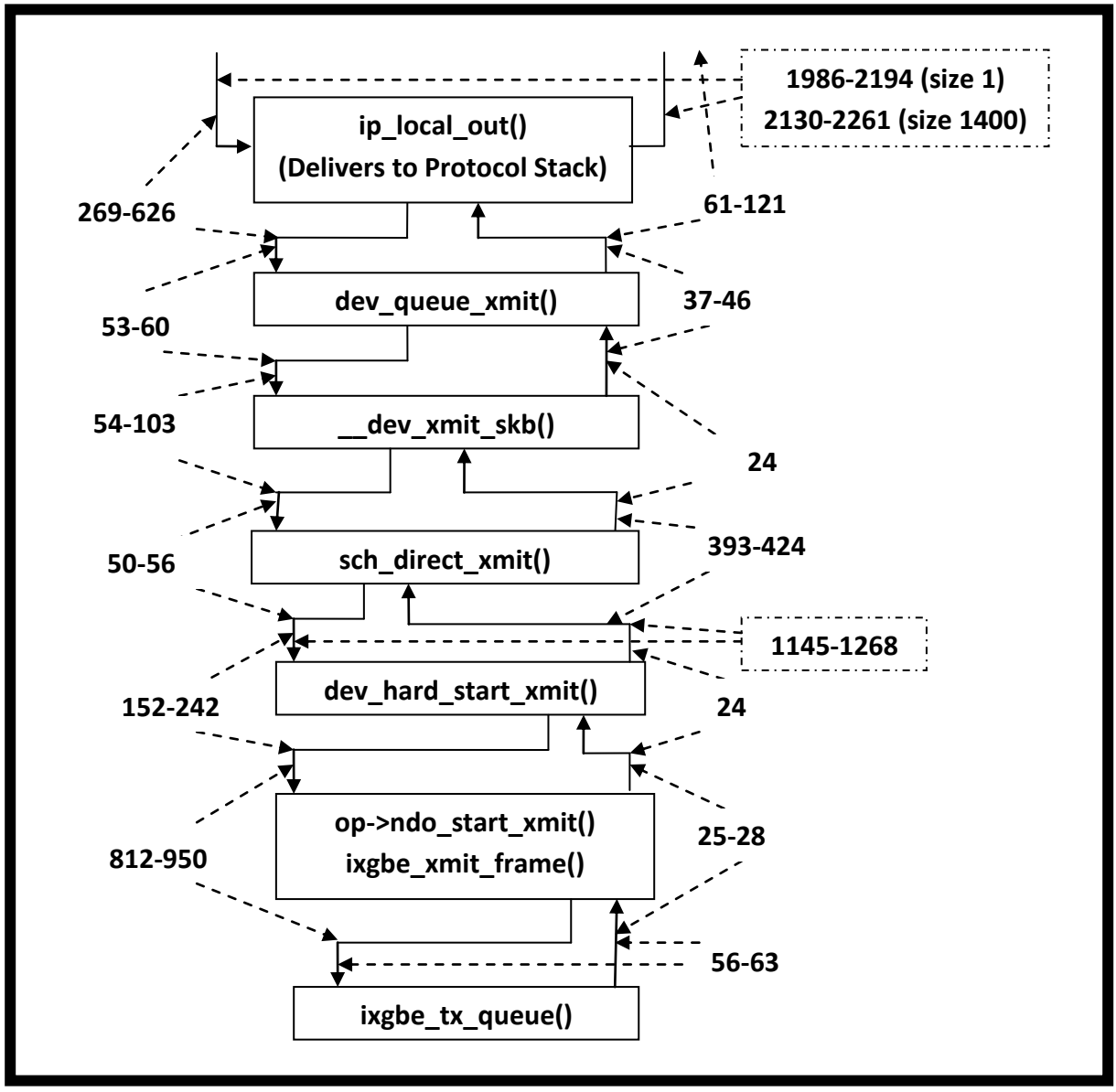


Figure 18: Dell Westmere SMP UDP TX Processing Below IP Stack

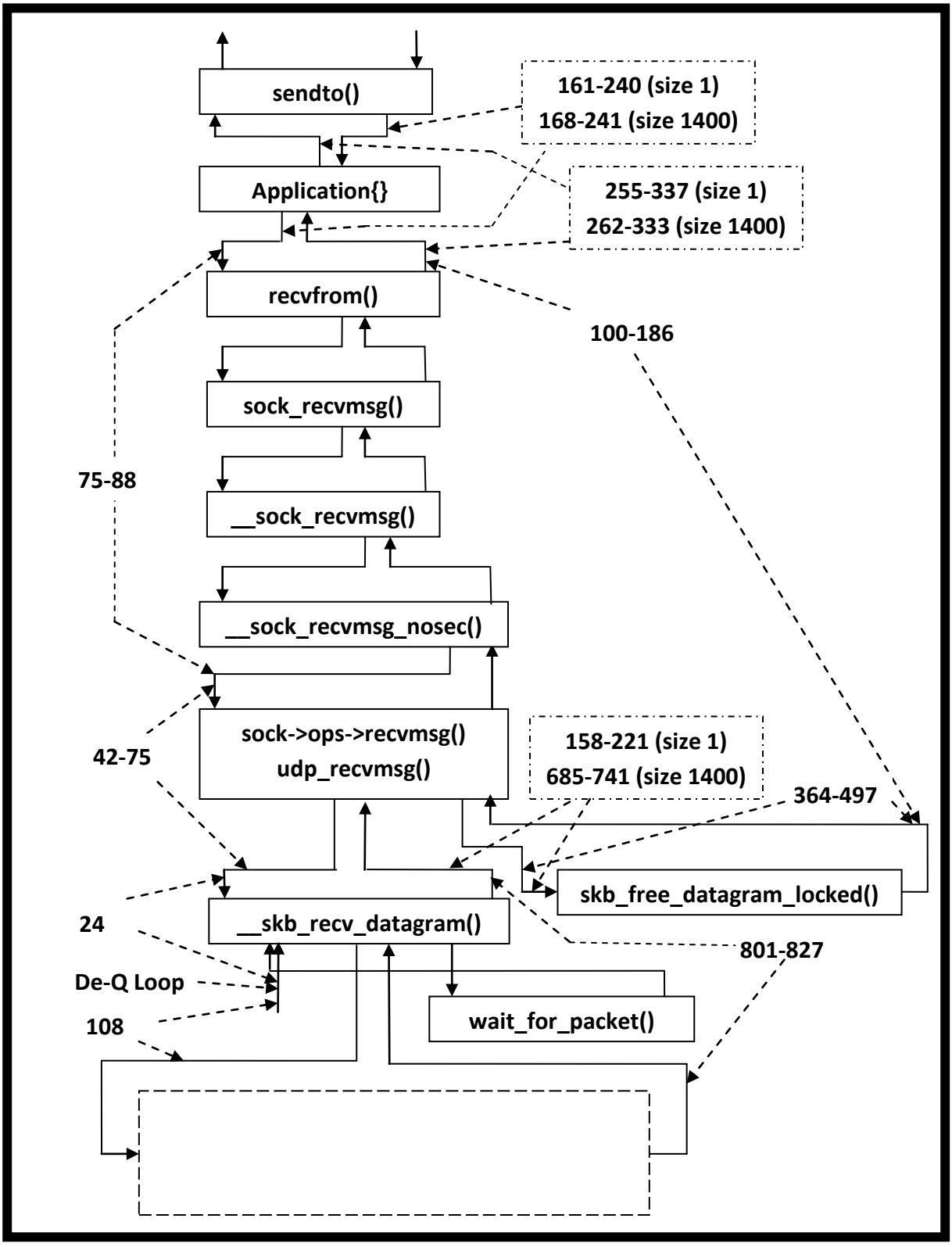


Figure 19: Dell Westmere SMP Socket & UDP RX Processing above IP Stack

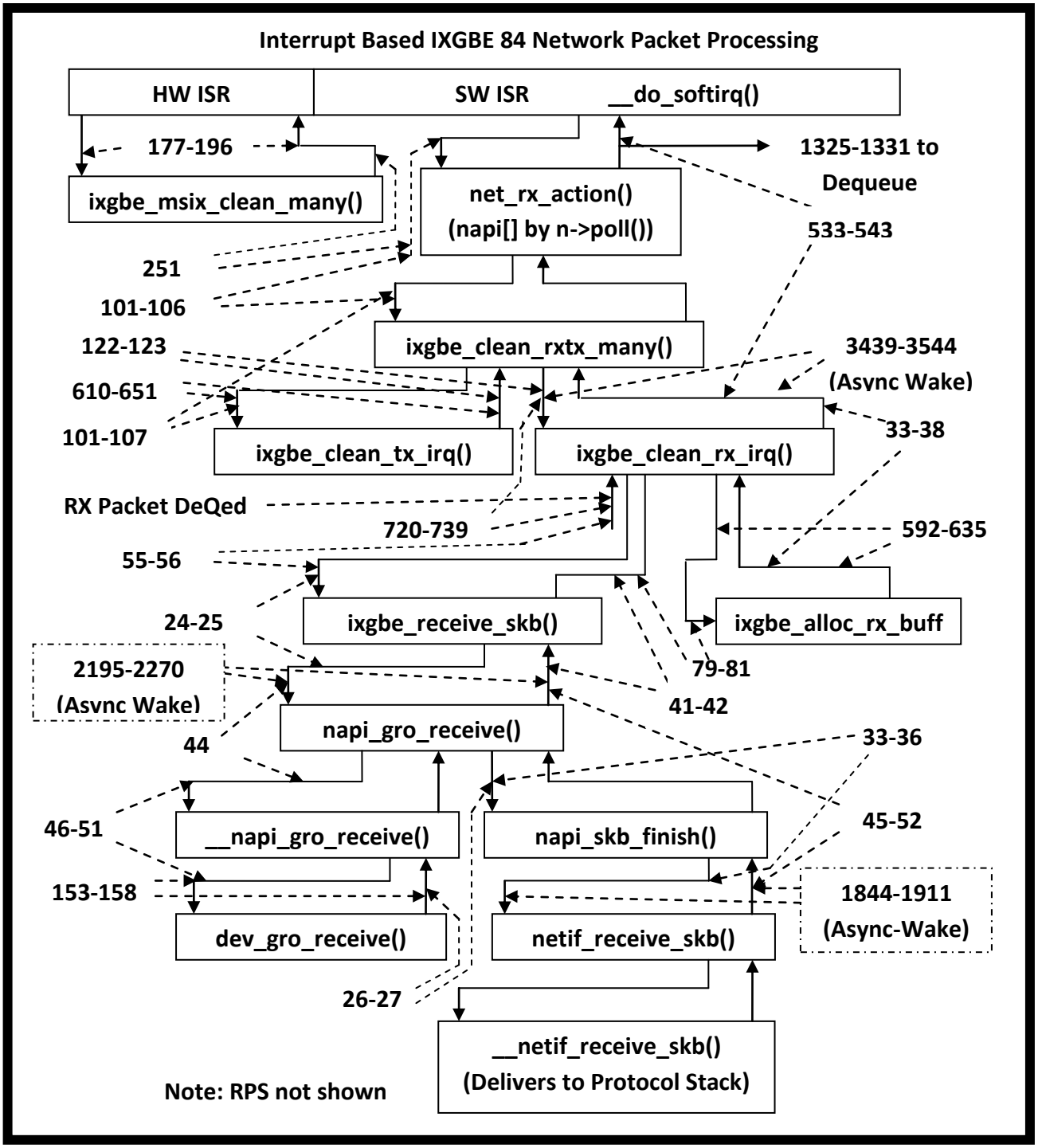


Figure 20: Dell Westmere SMP IXGBE Interrupt Driven UDP RX Processing Below IP Stack

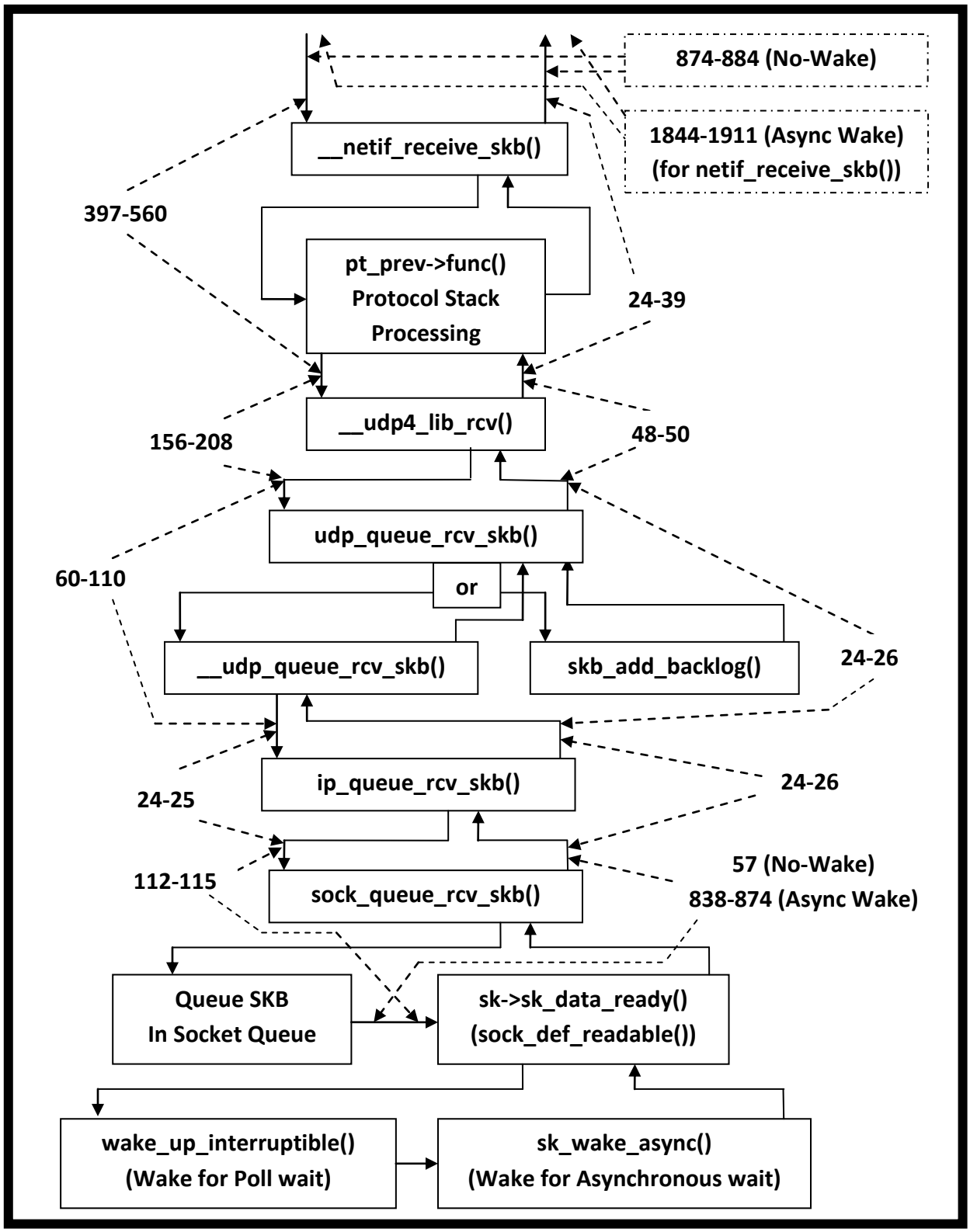


Figure 21: Dell Westmere SMP RX Protocol Stack Processing of UDP above IP Stack

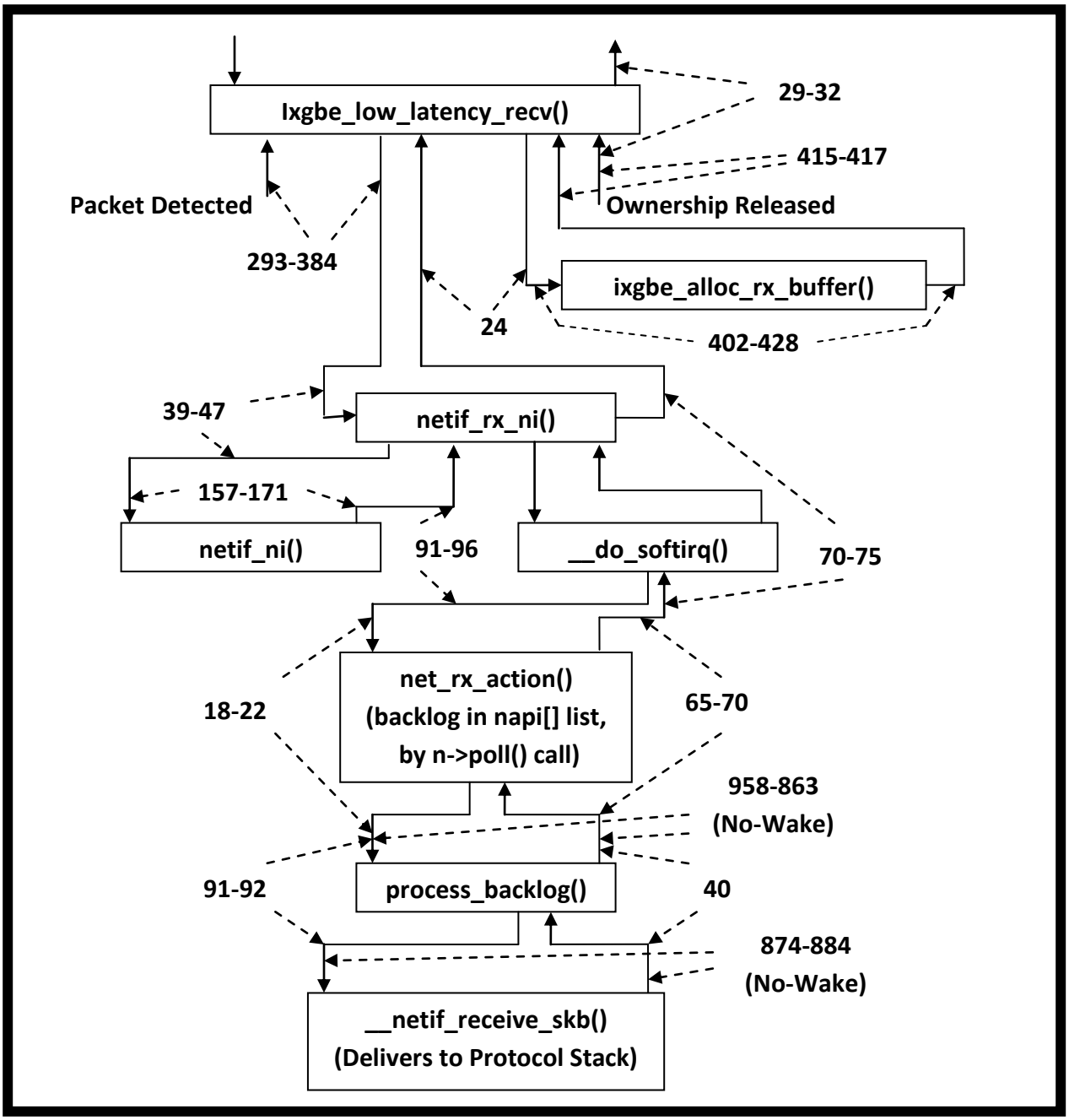


Figure 22: Dell Westmere SMP ixgbe_low_latency_recv() Driver RX Flush Processing

UDP Process Timing Range for Nehalem

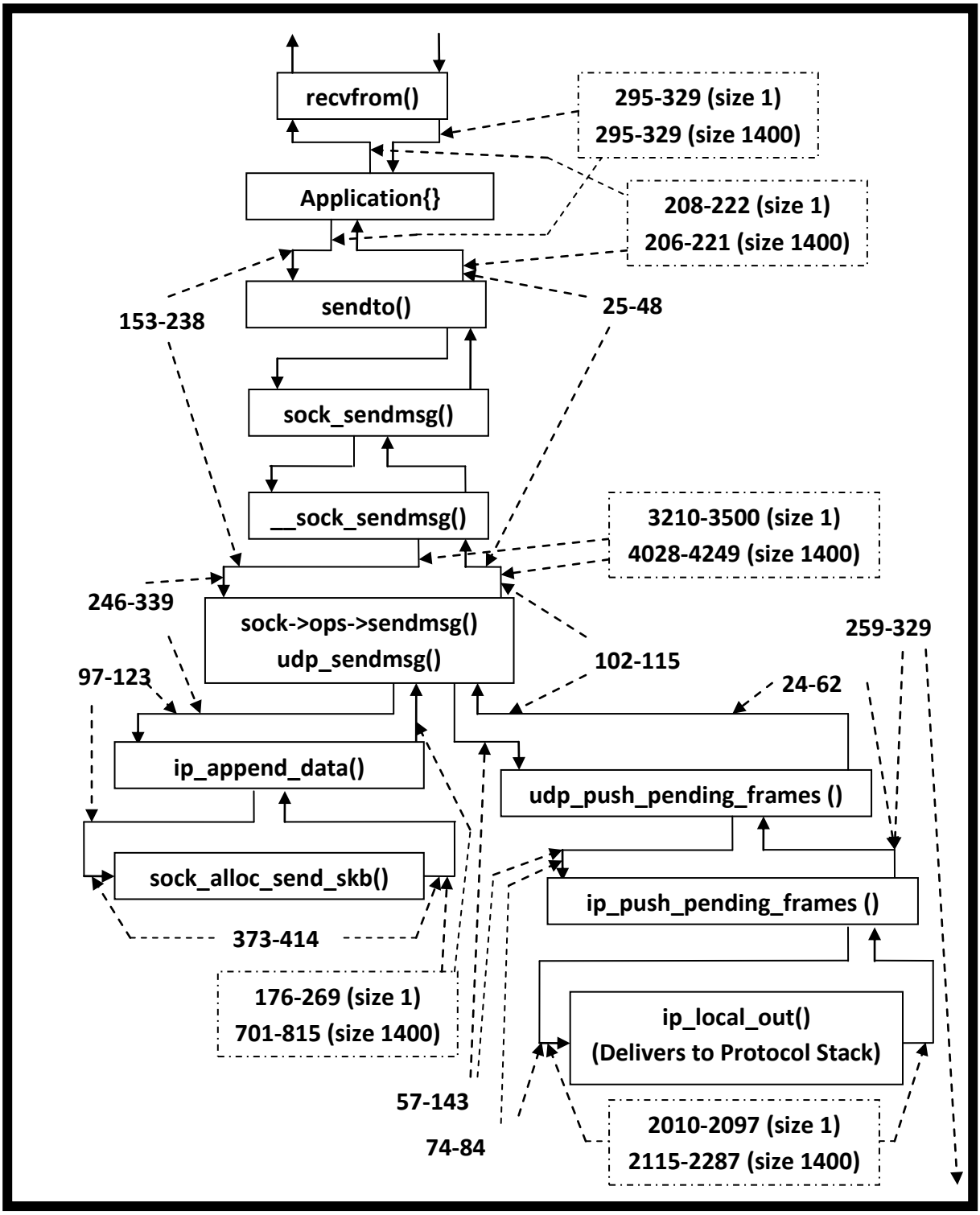


Figure 23: 2S Nehalem SMP Socket & UDP TX Processing above IP Stack

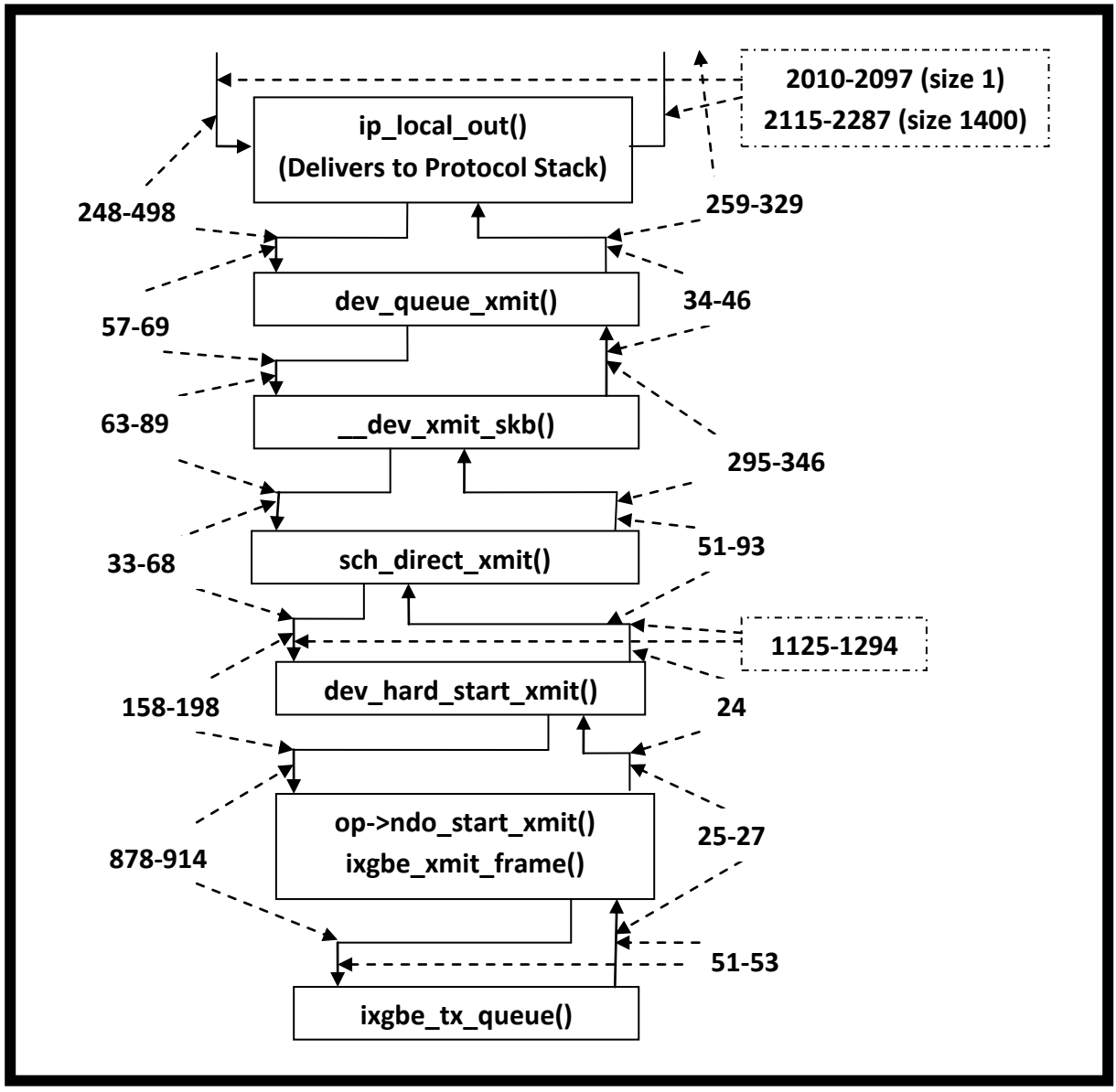


Figure 24: 2S Nehalem SMP UDP TX Processing Below IP Stack

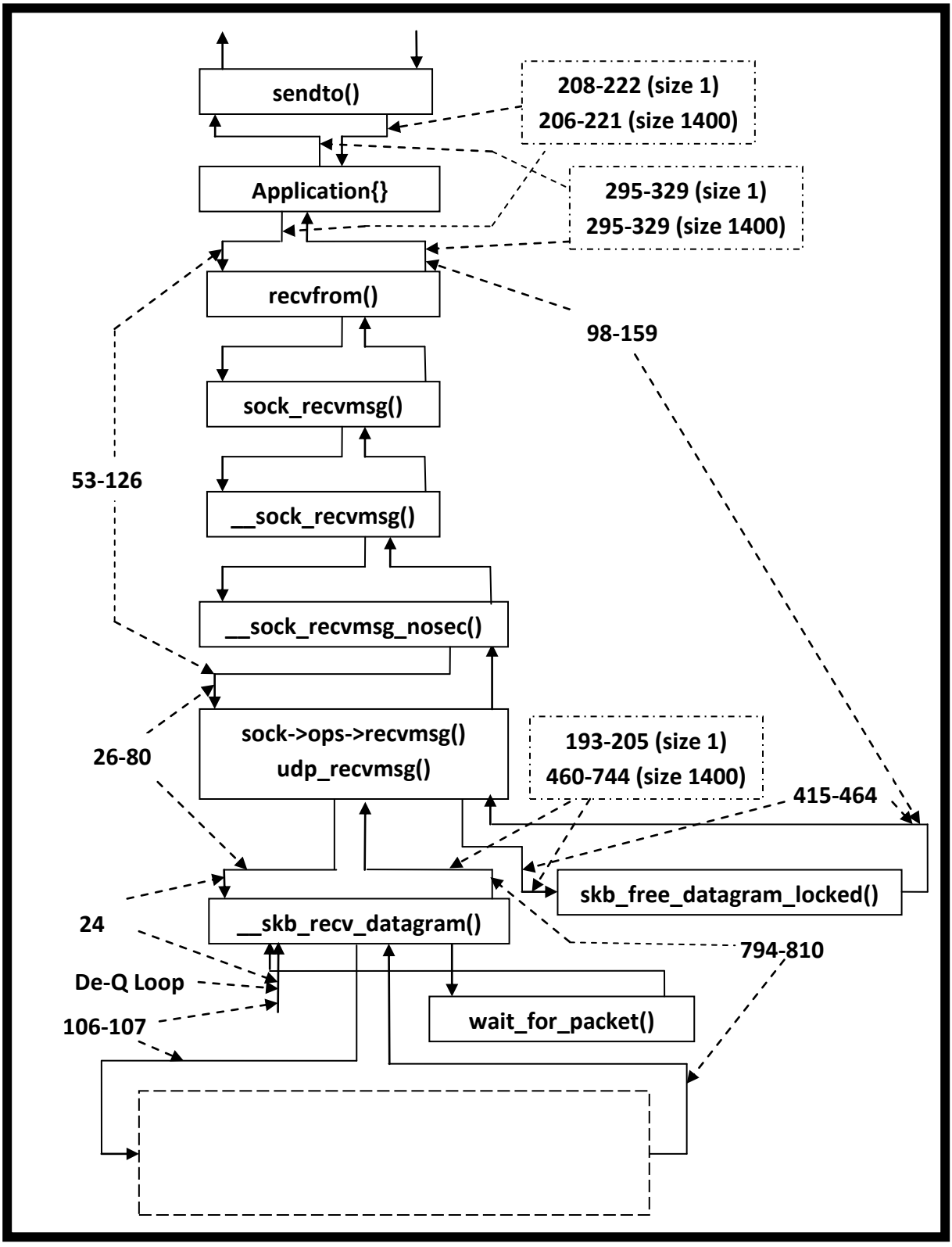


Figure 25: 2S Nehalem SMP Socket & UDP RX Processing above IP Stack

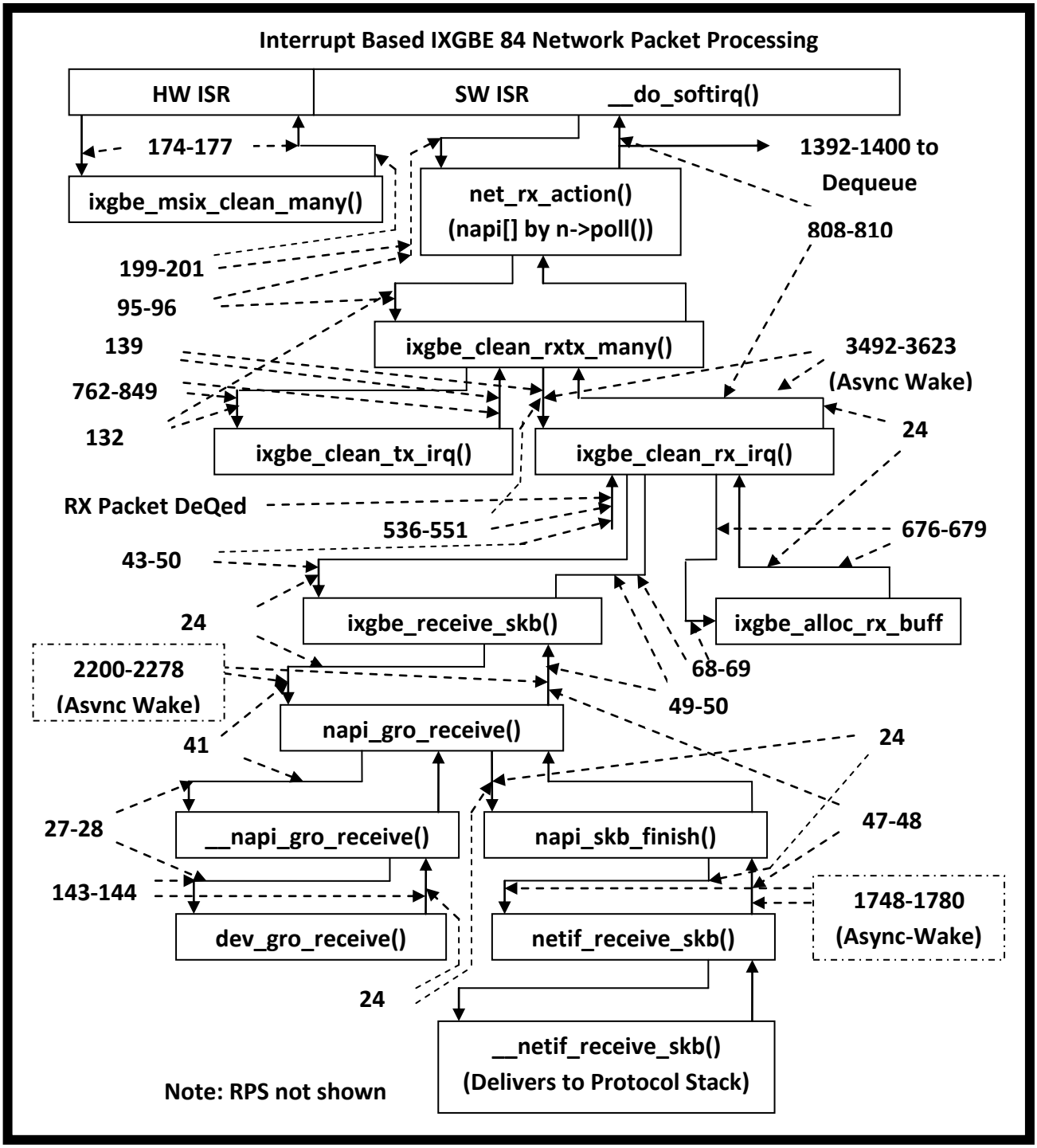


Figure 26: 2S Nehalem SMP IXGBE Interrupt Driven UDP RX Processing Below IP Stack

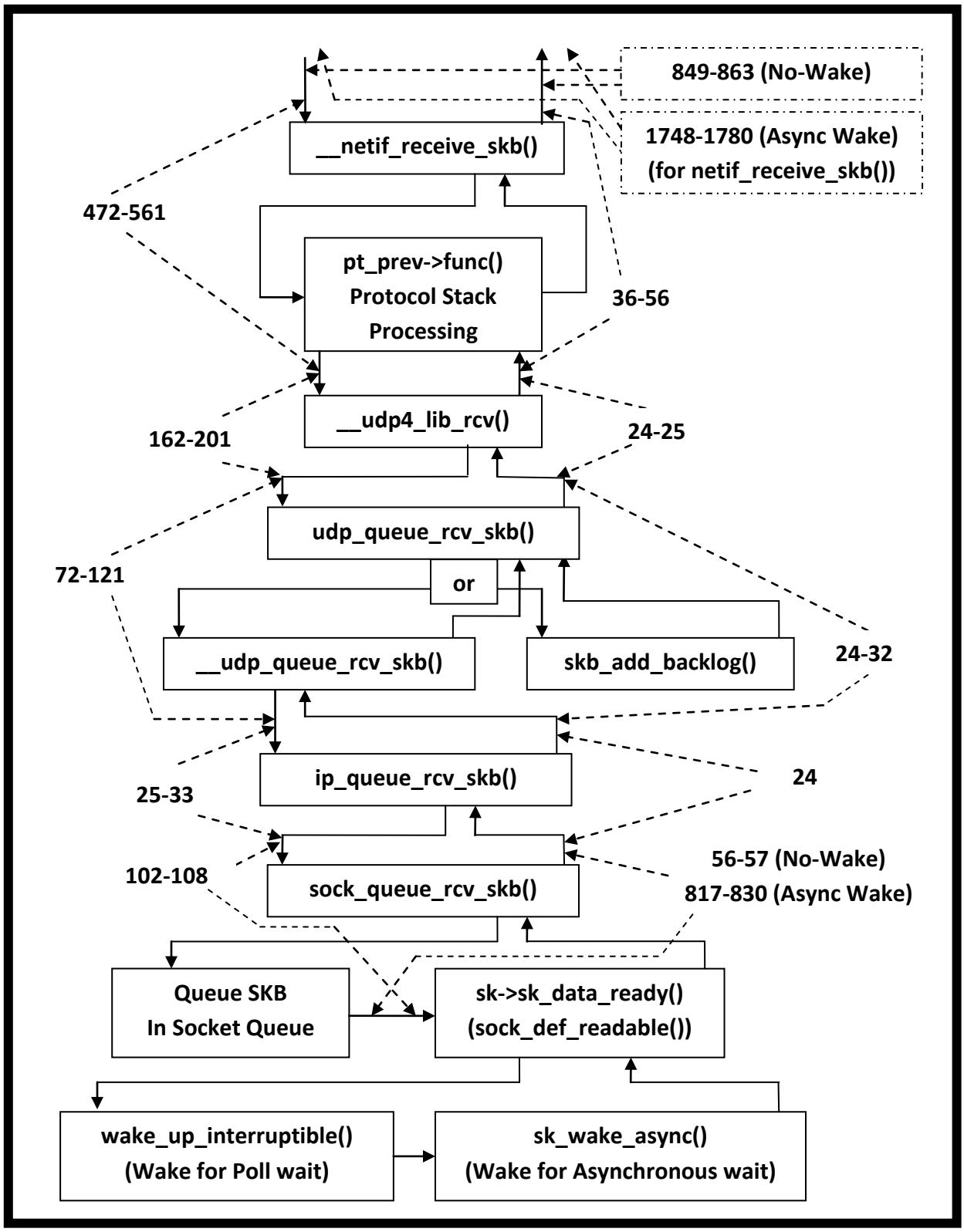


Figure 27: 2S Nehalem SMP RX Protocol Stack Processing of UDP above IP Stack

UDP Process Timing Range for Nehalem with 1 Socketed CPU

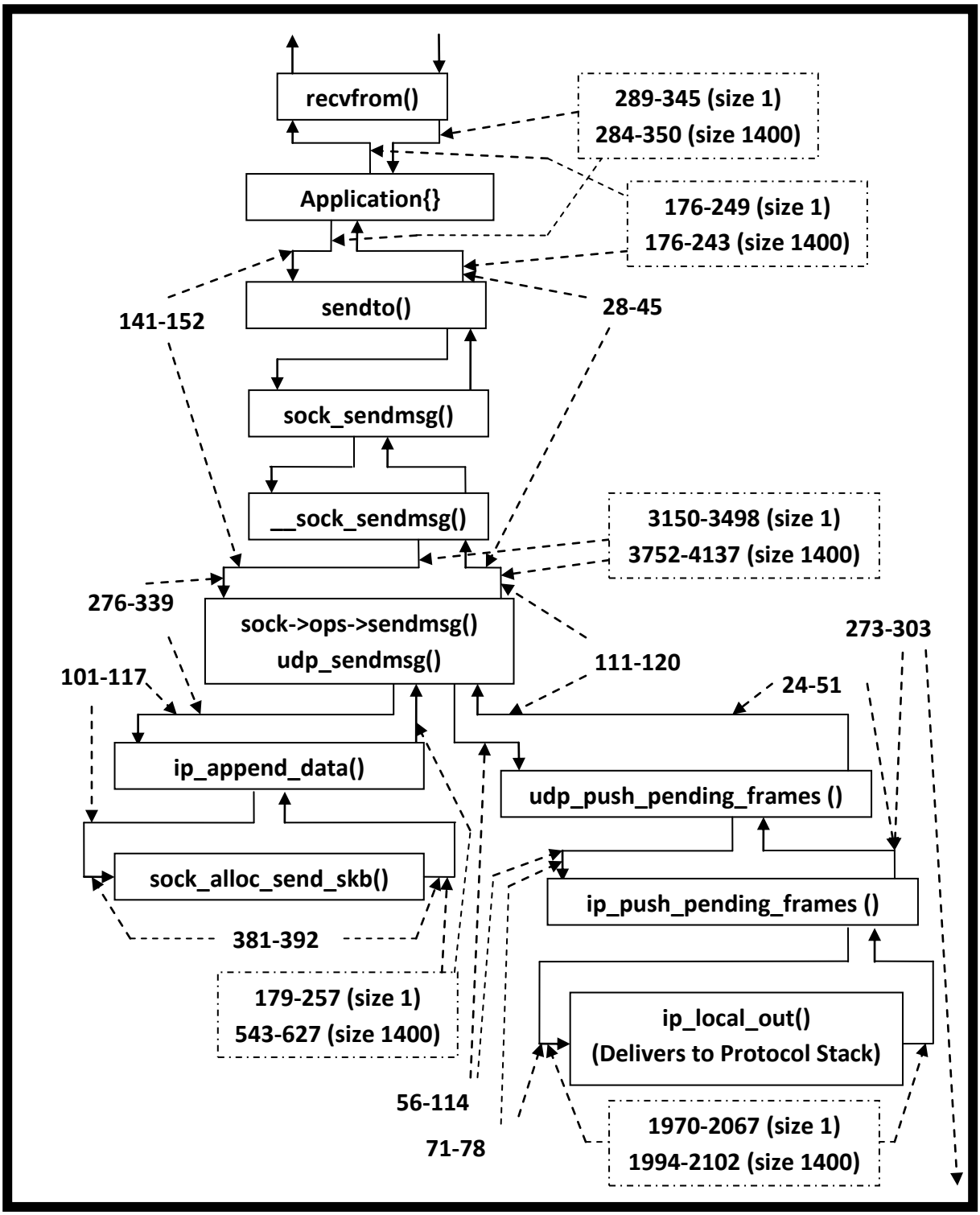


Figure 28: Nehalem, 1 CPU, SMP Socket & UDP TX Processing above IP Stack

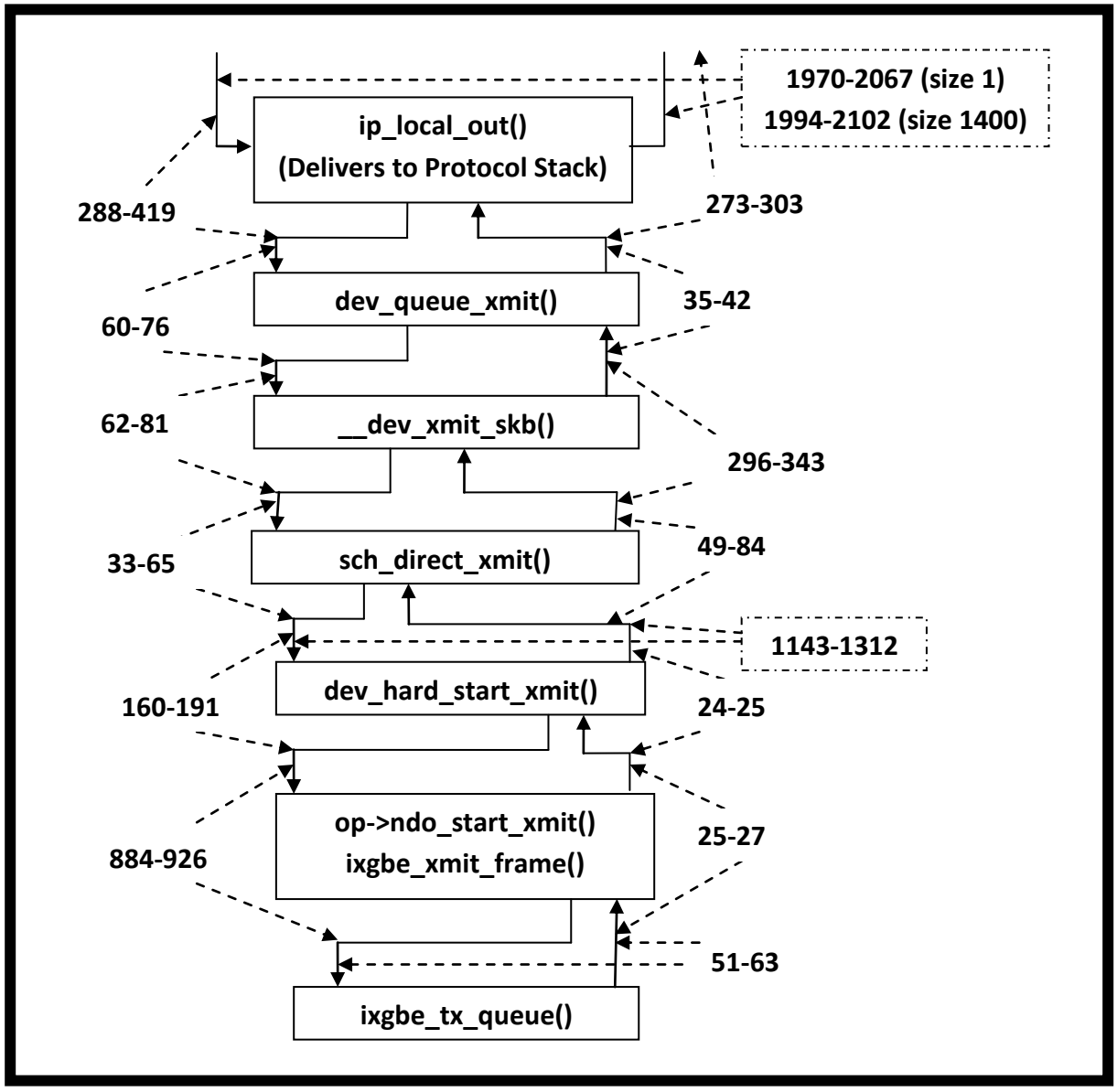


Figure 29: Nehalem, 1 CPU, SMP UDP TX Processing Below IP Stack

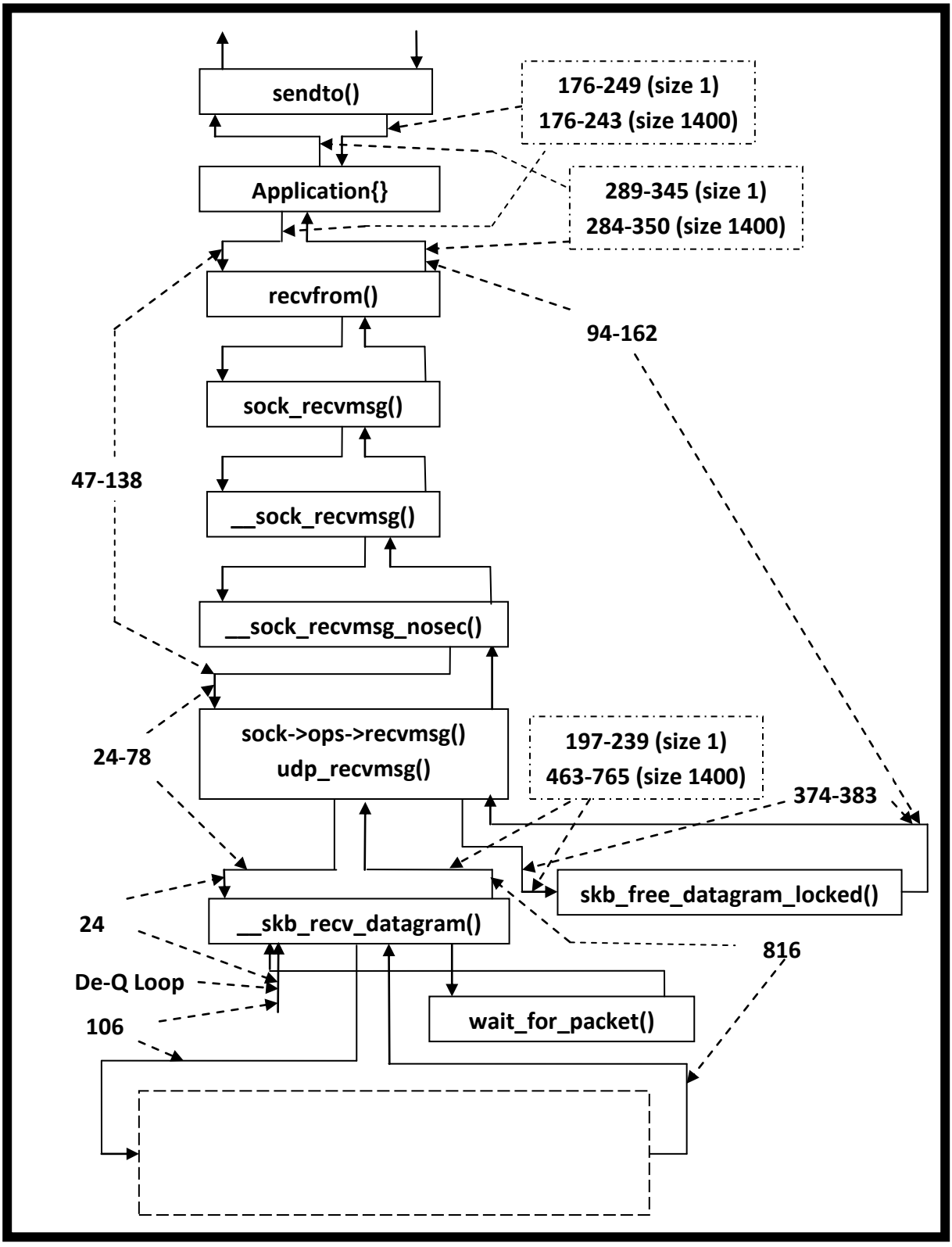


Figure 30: Nehalem, 1 CPU, SMP Socket & UDP RX Processing above IP Stack

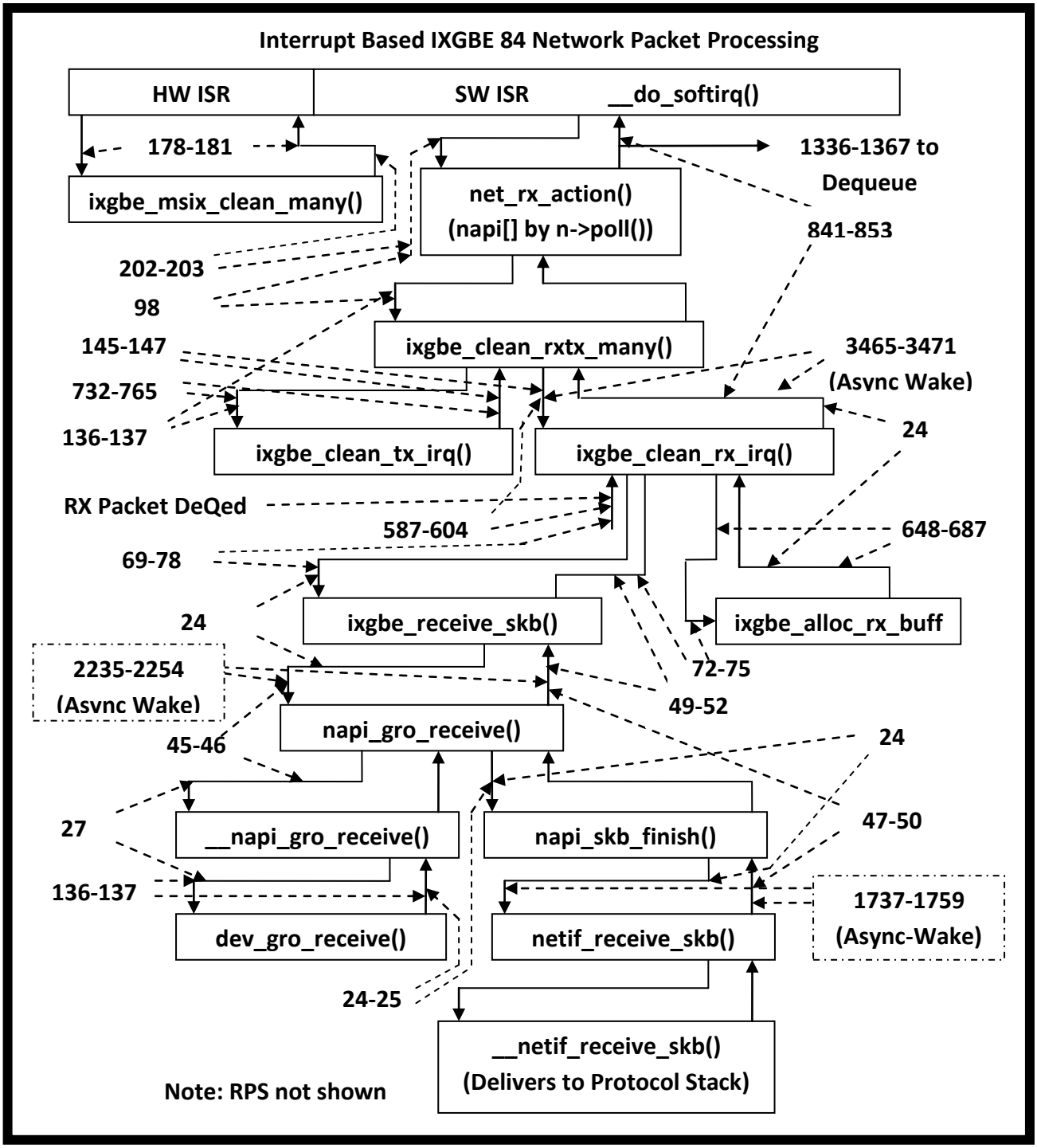


Figure 31: Nehalem, 1 CPU, SMP IXGBE Interrupt Driven UDP RX Processing Below IP Stack

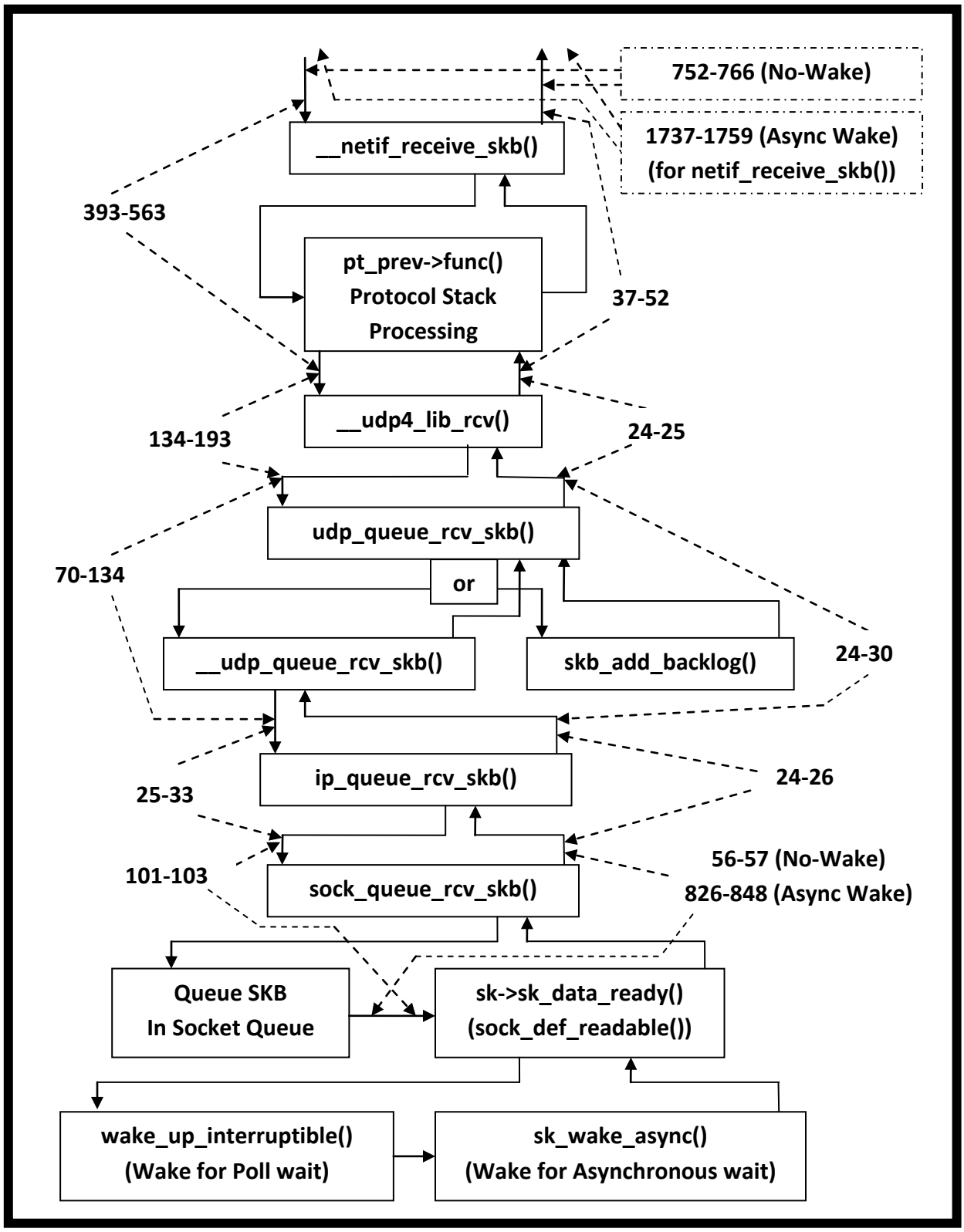


Figure 32: Nehalem, 1 CPU, SMP RX Protocol Stack Processing of UDP above IP Stack