

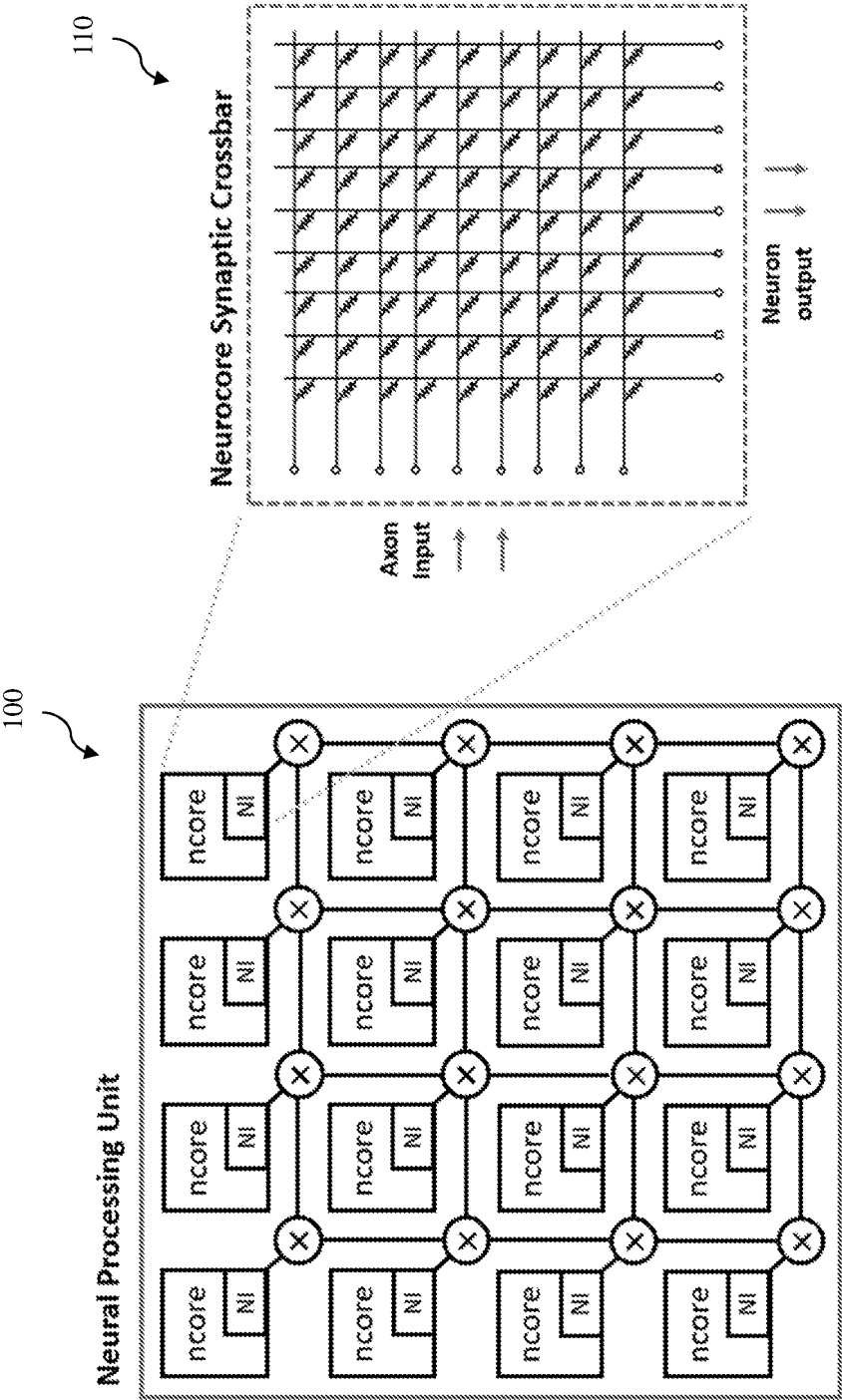


FIG. 1

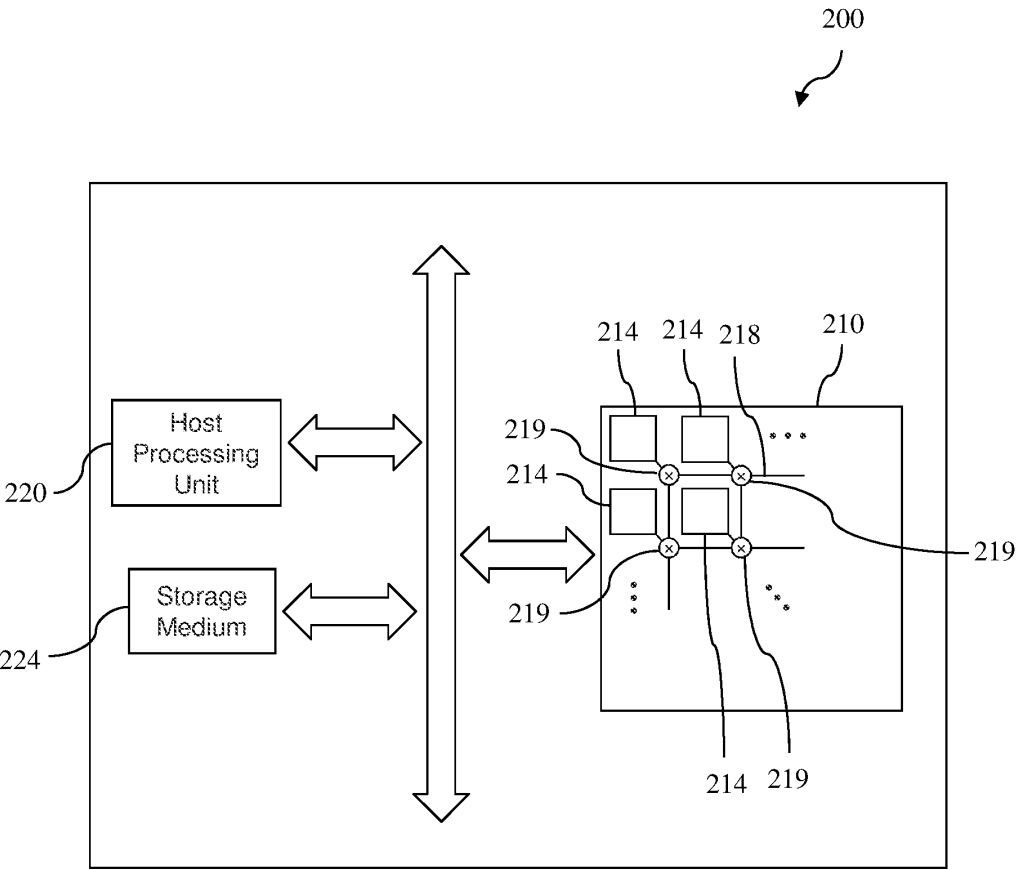
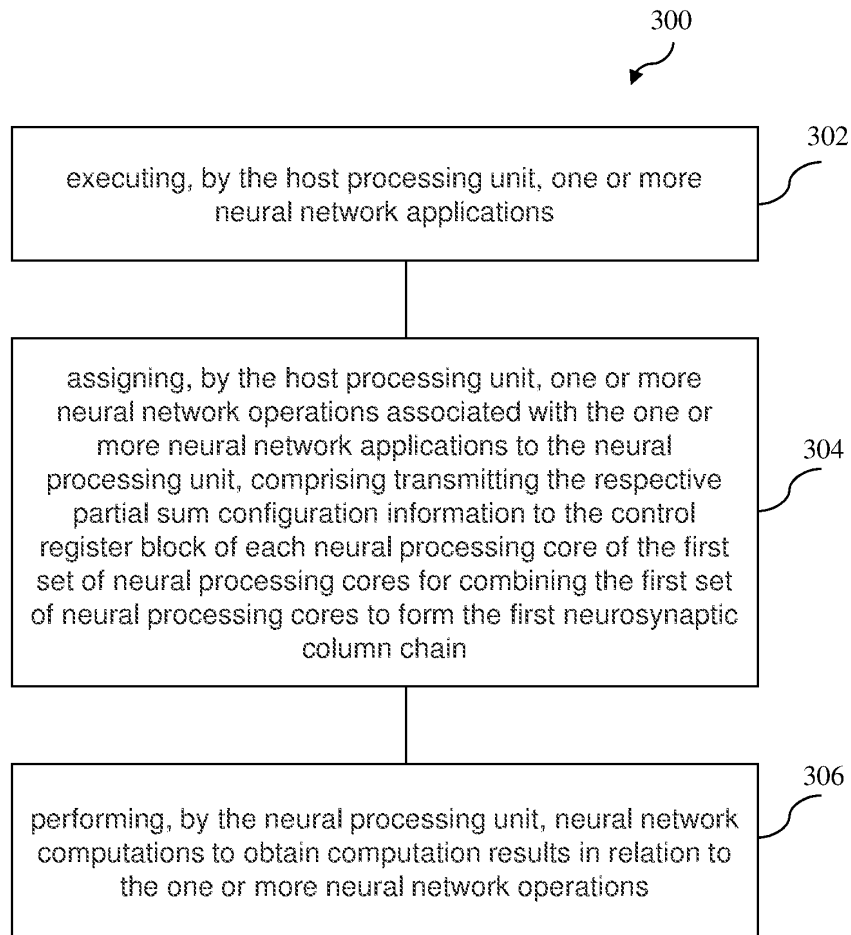
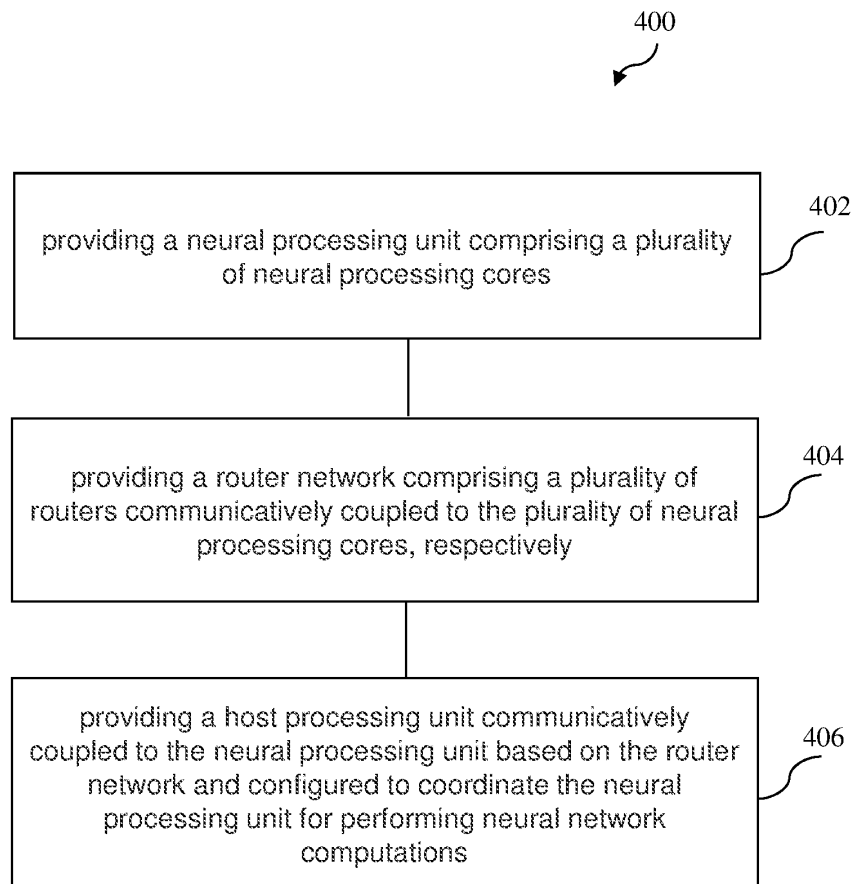


FIG. 2

3/25

**FIG. 3**

4/25

**FIG. 4A**

5/25

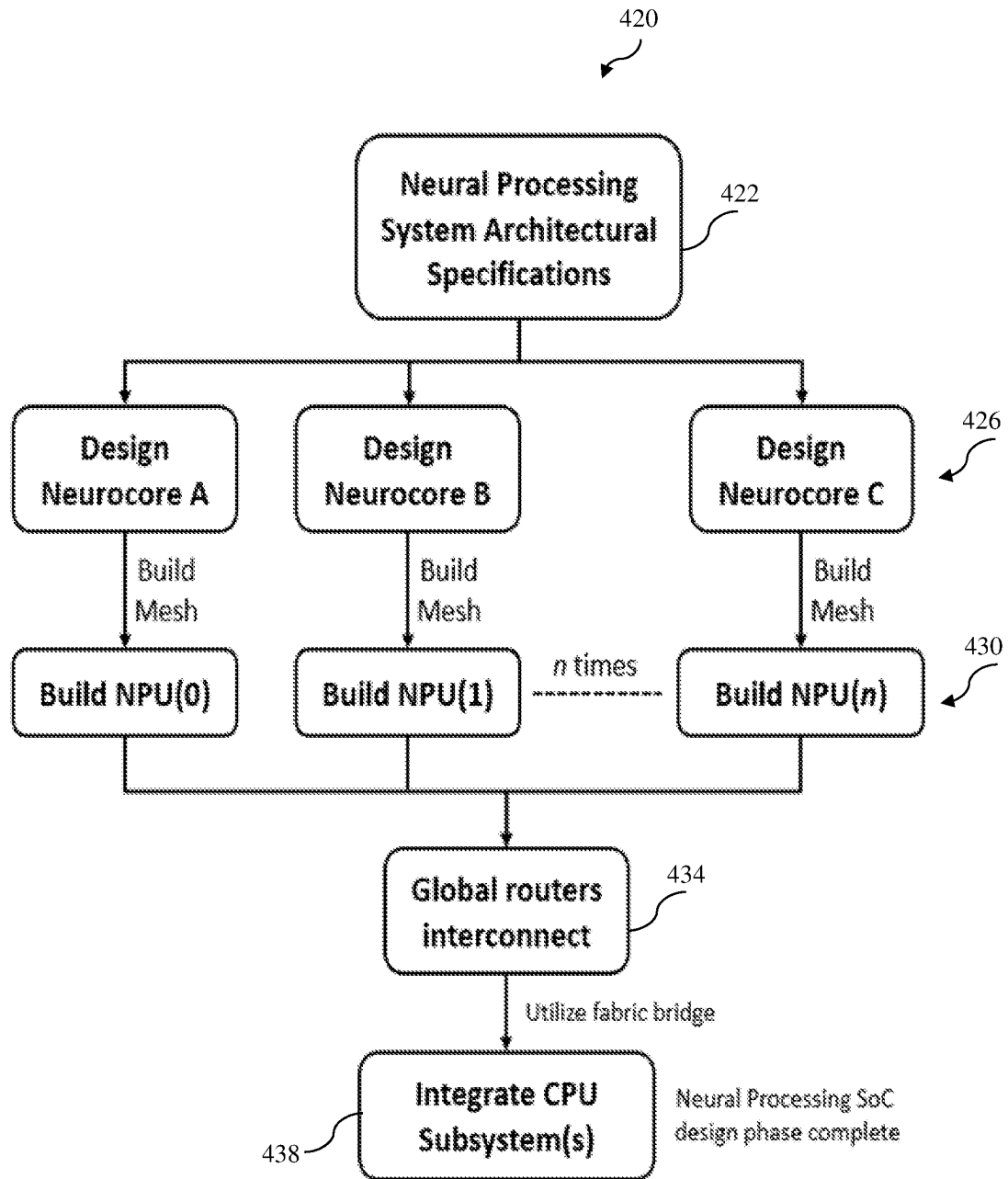


FIG. 4B

6/25

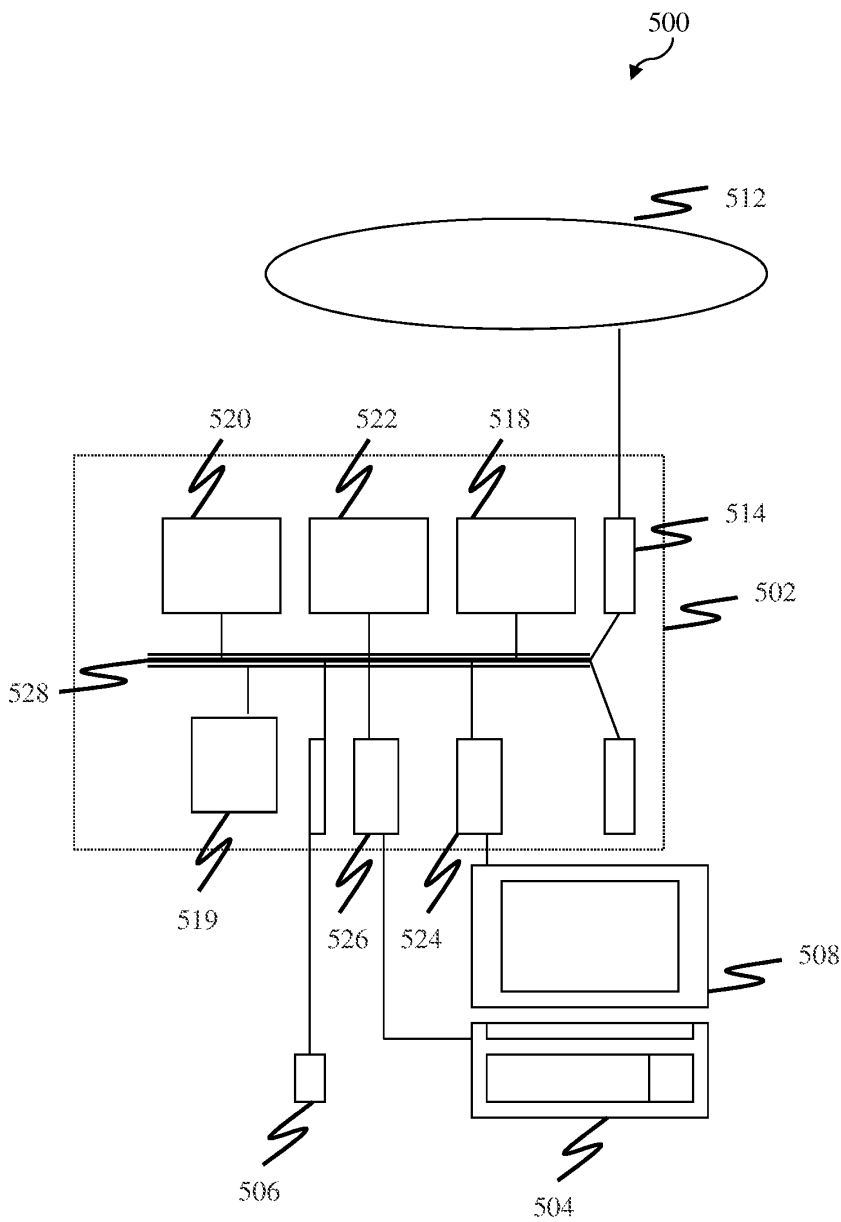


FIG. 5

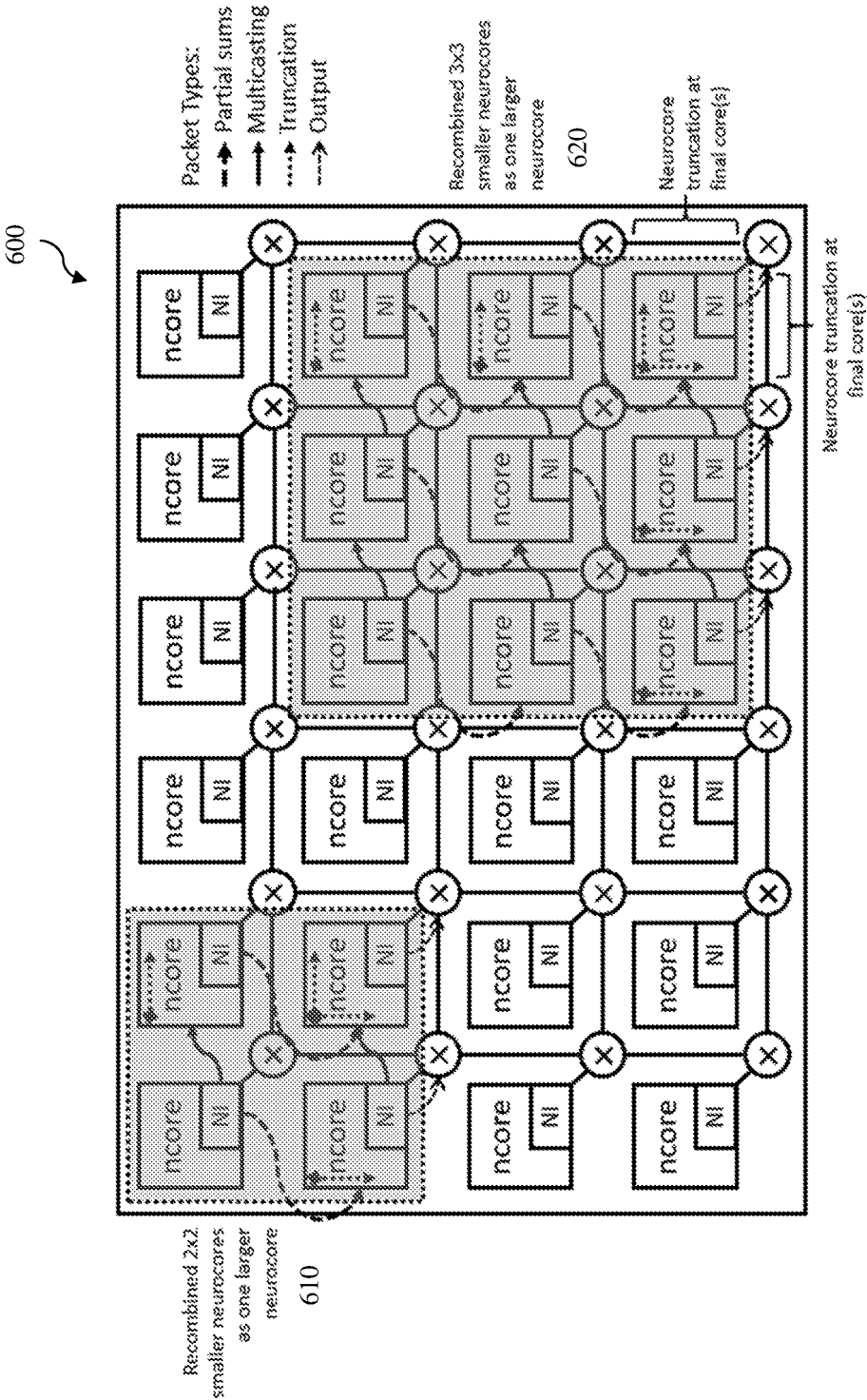


FIG. 6

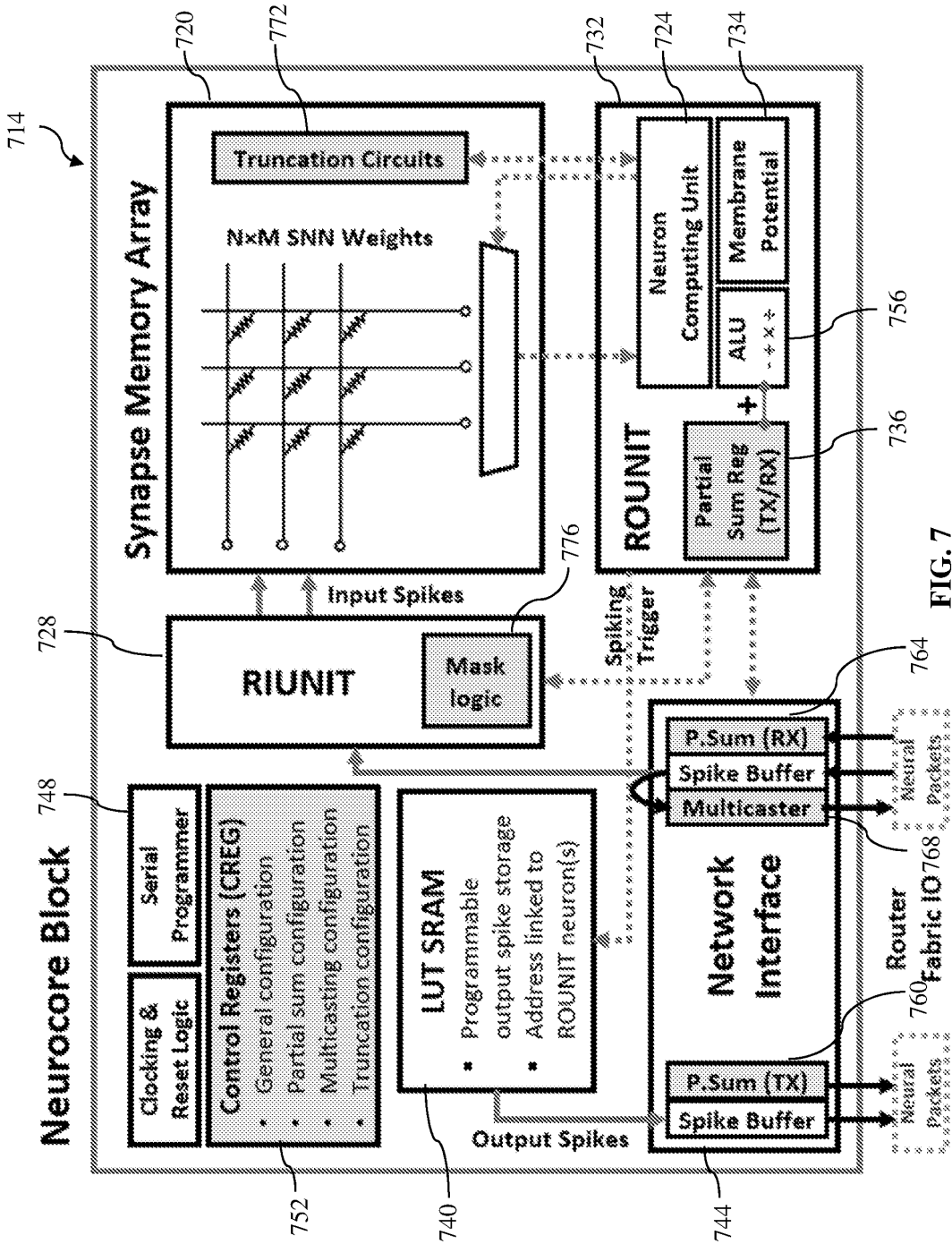


FIG. 7

Bits	TX Direction
00b	North neurocore
01b	East neurocore
10b	South neurocore
11b	West neurocore

FIG. 8A

Bits	Mode
00b	Single core mode
01b	Receiver (RX)
10b	Transmitter (TX)
11b	Transceiver (TX-RX)

FIG. 8B

10/25

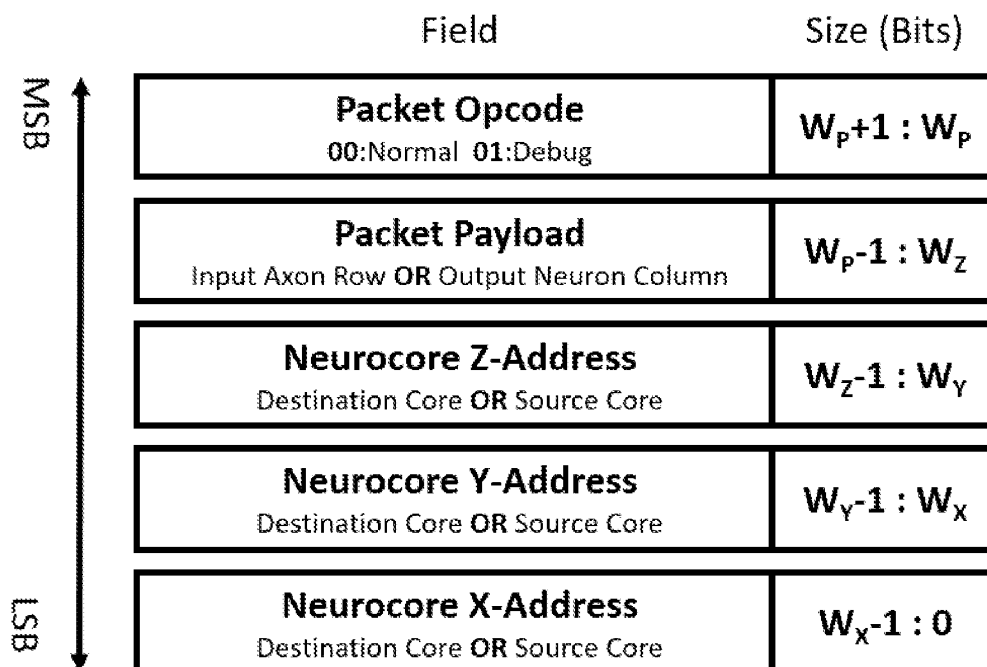


FIG. 9A

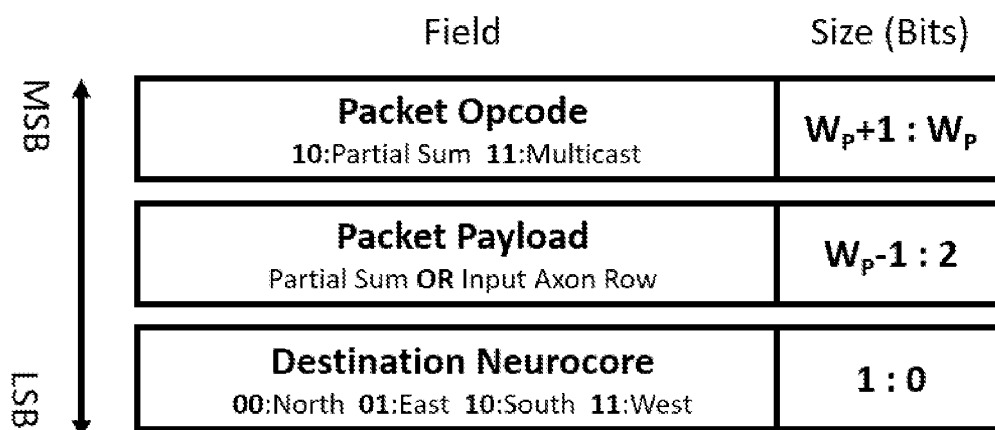


FIG. 9B

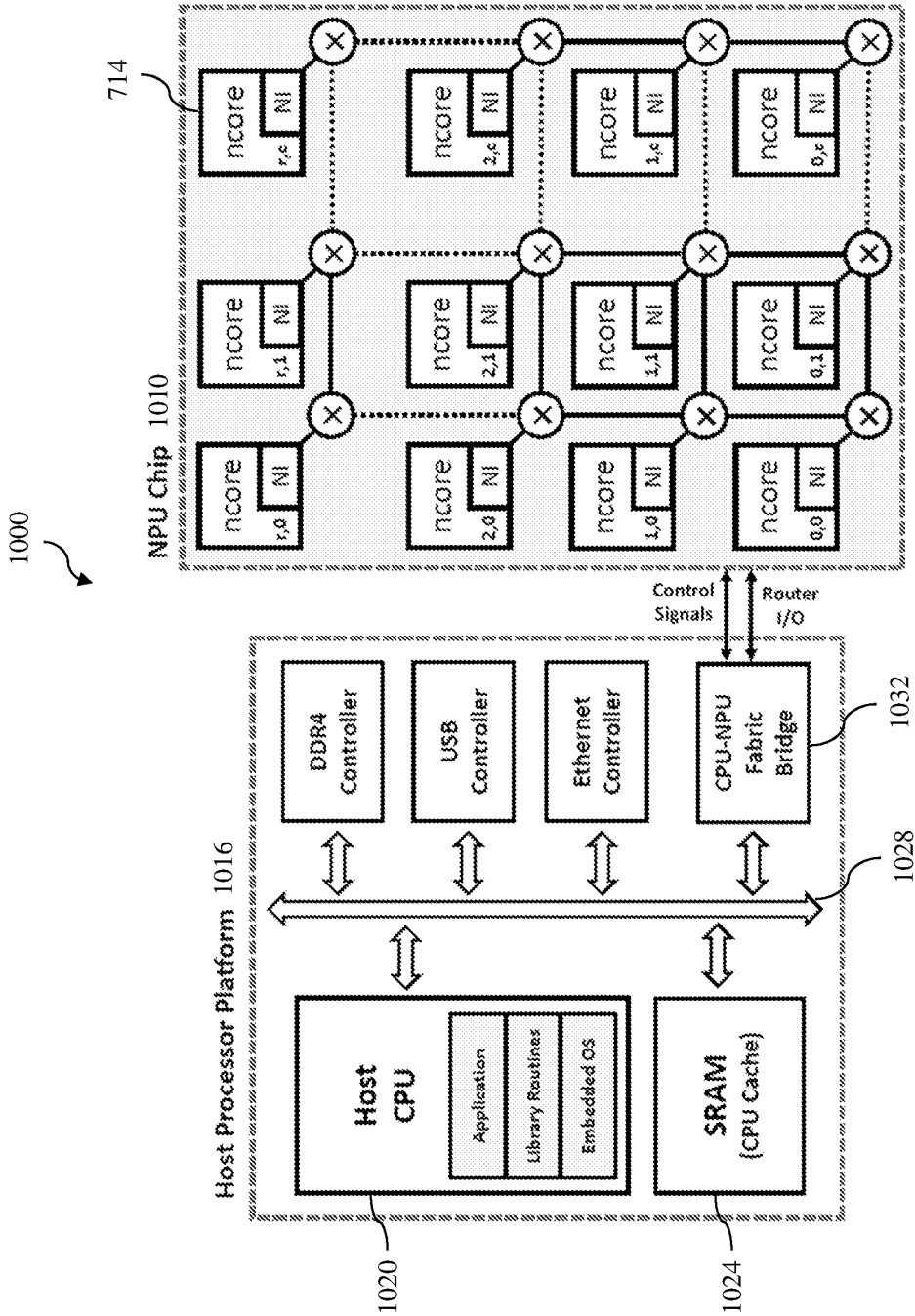


FIG. 10

12/25

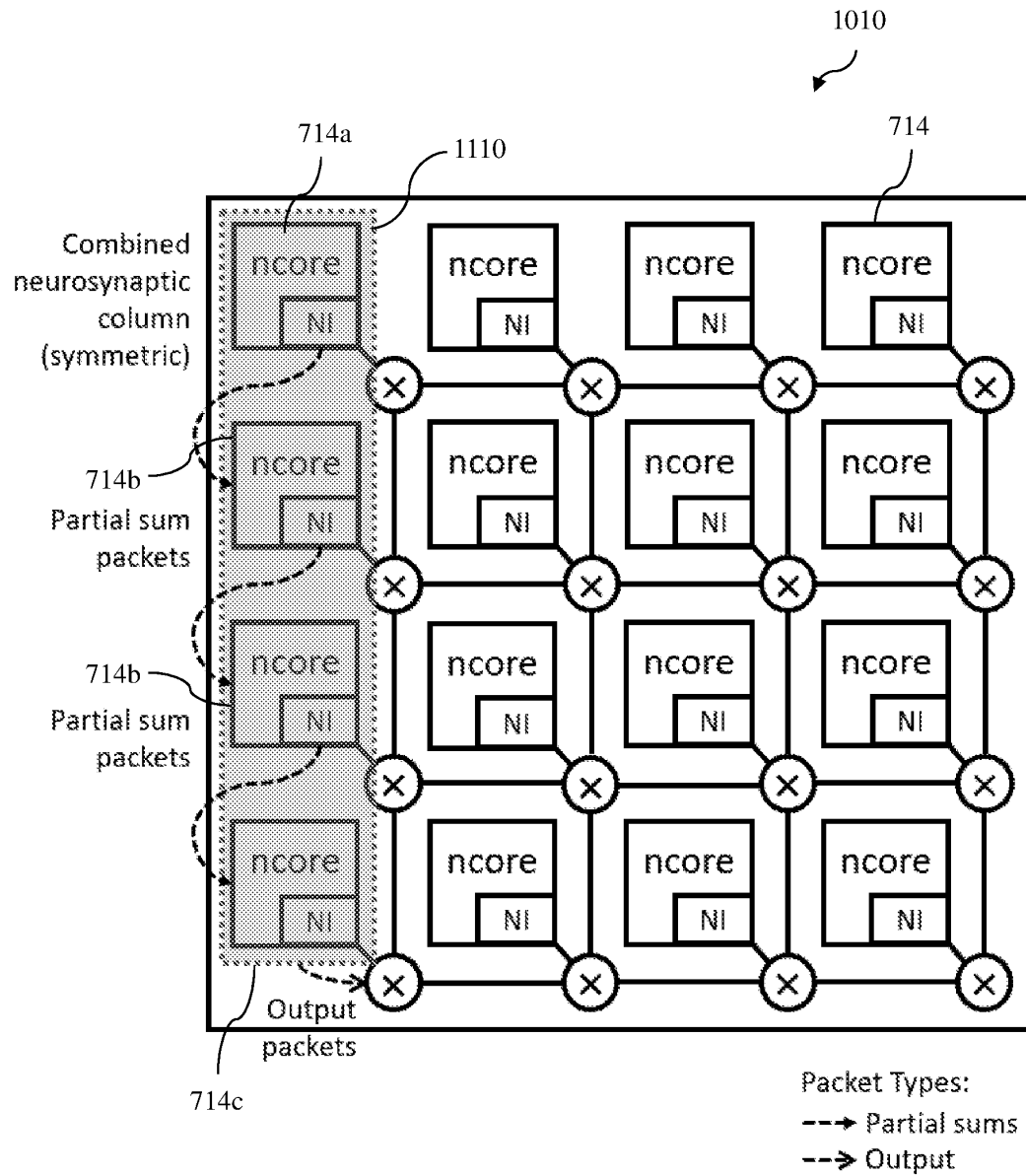


FIG. 11A

1110

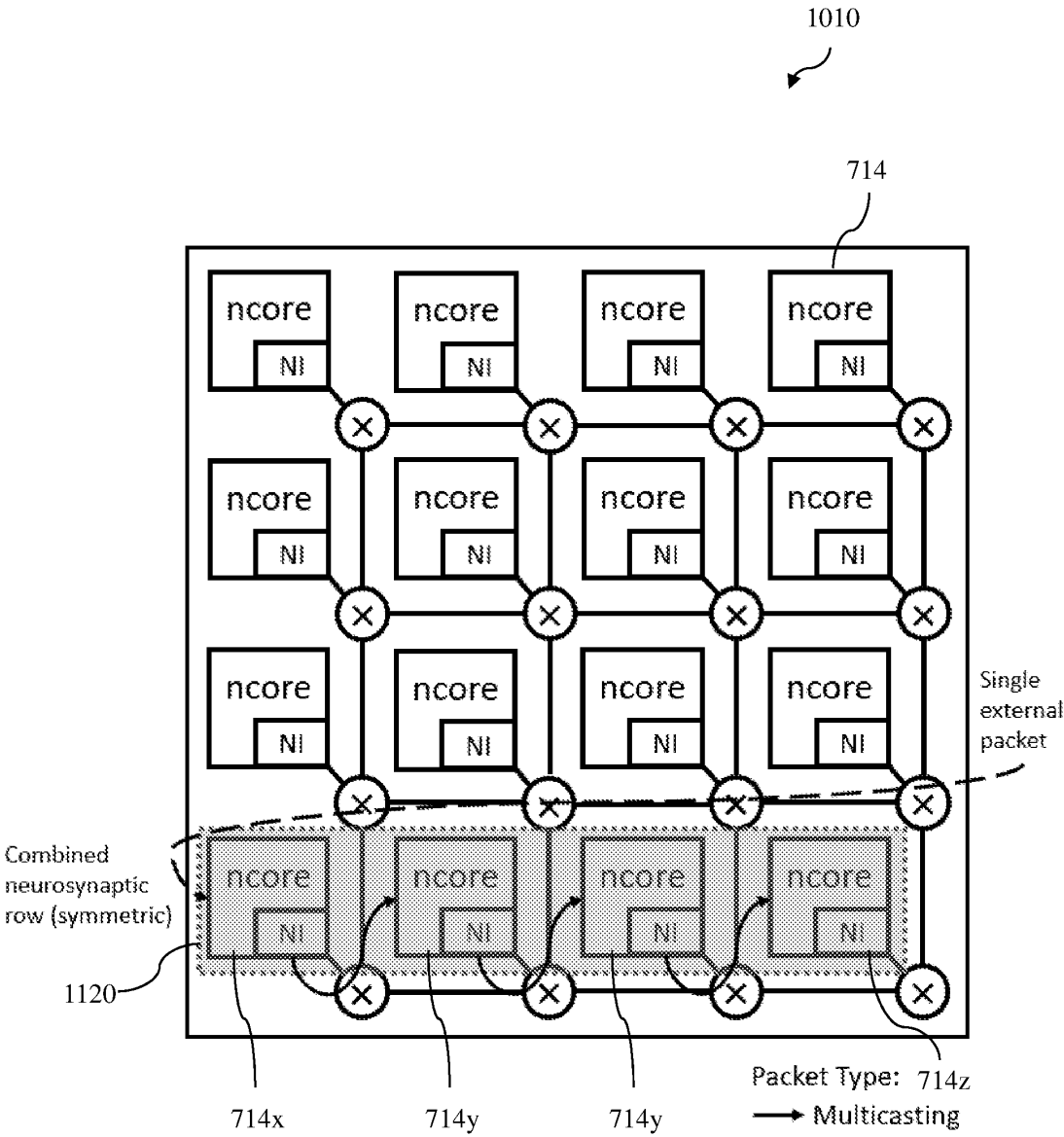


FIG. 12A

15/25

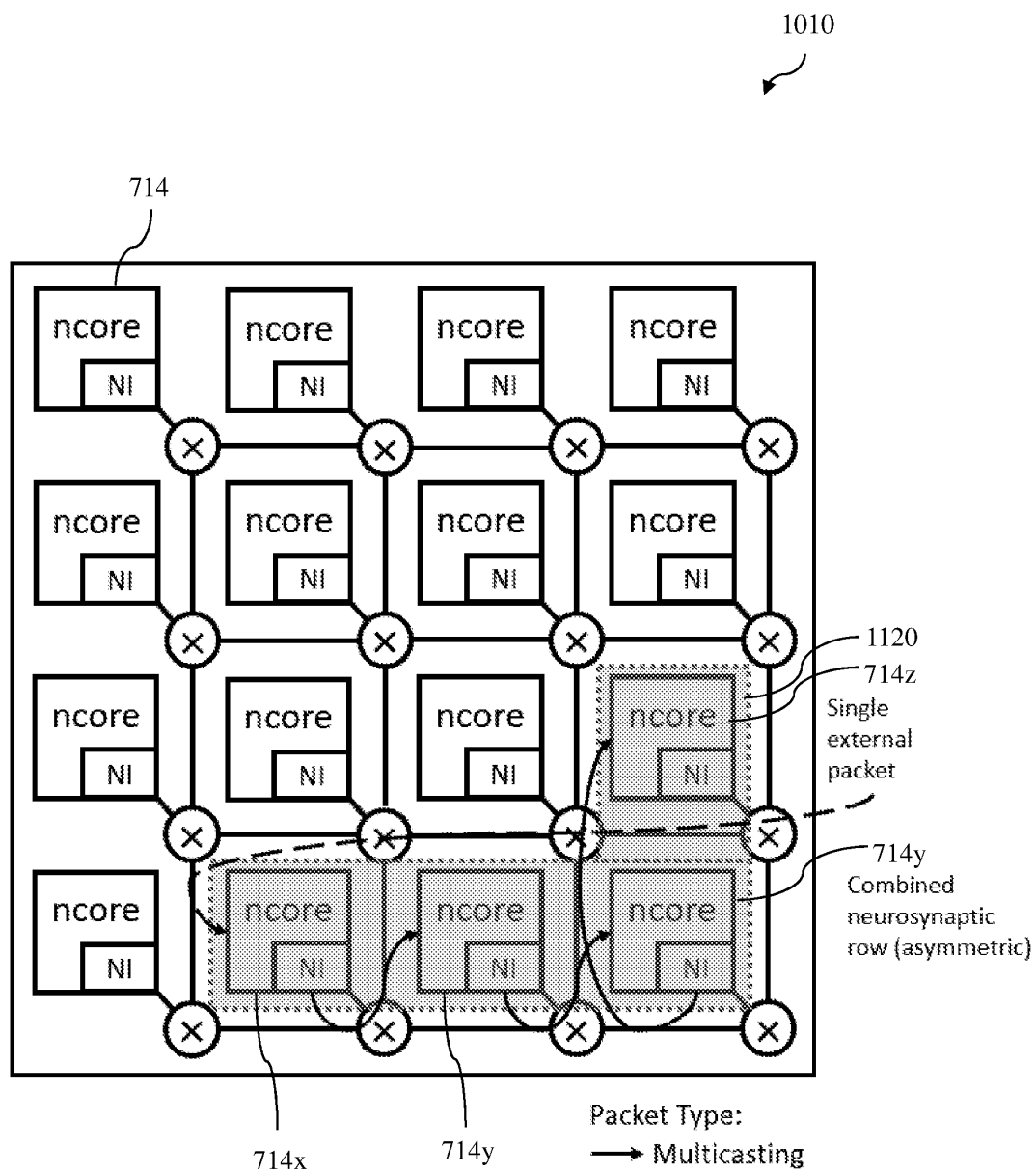


FIG. 12B

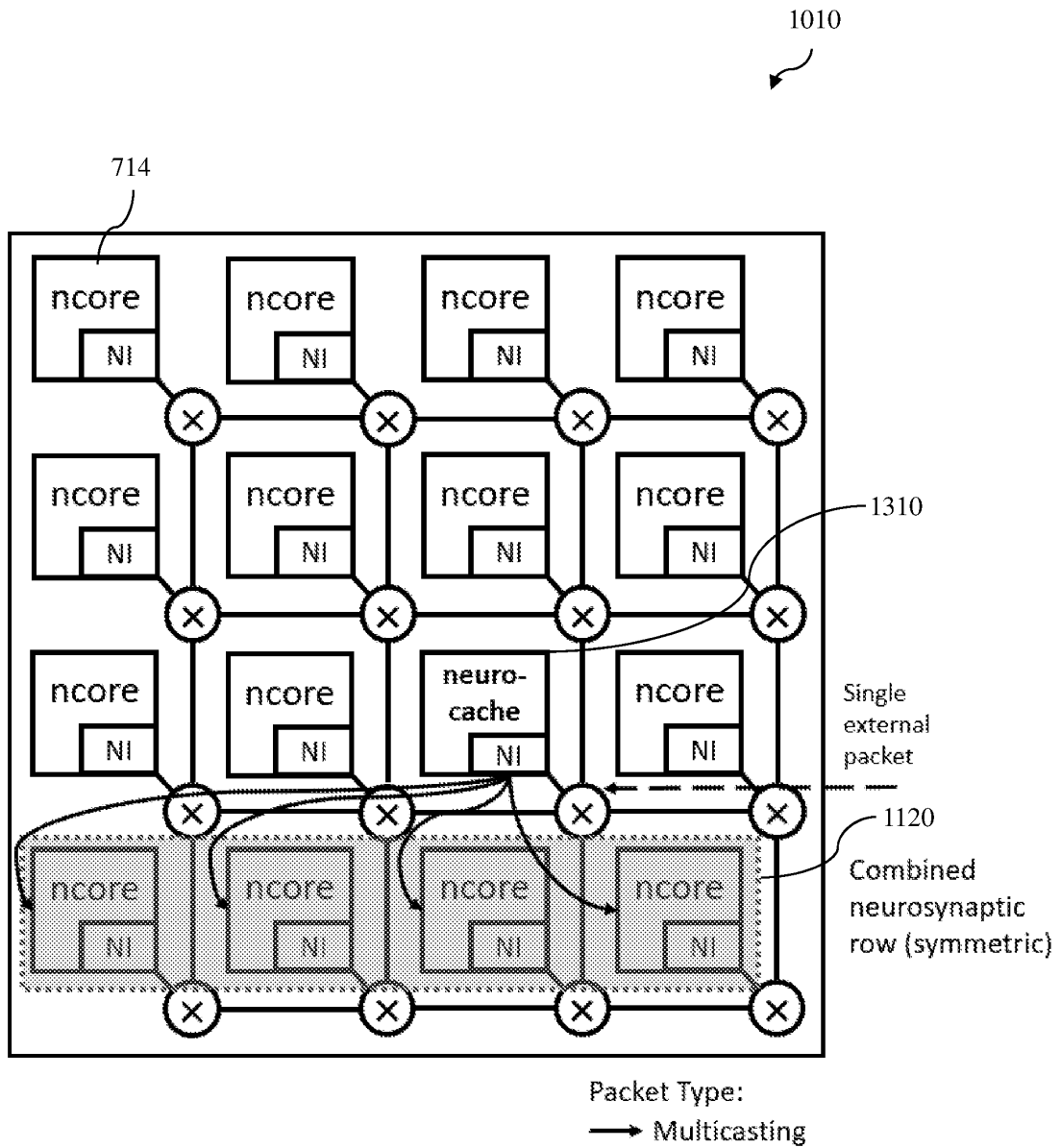


FIG. 13A

17/25

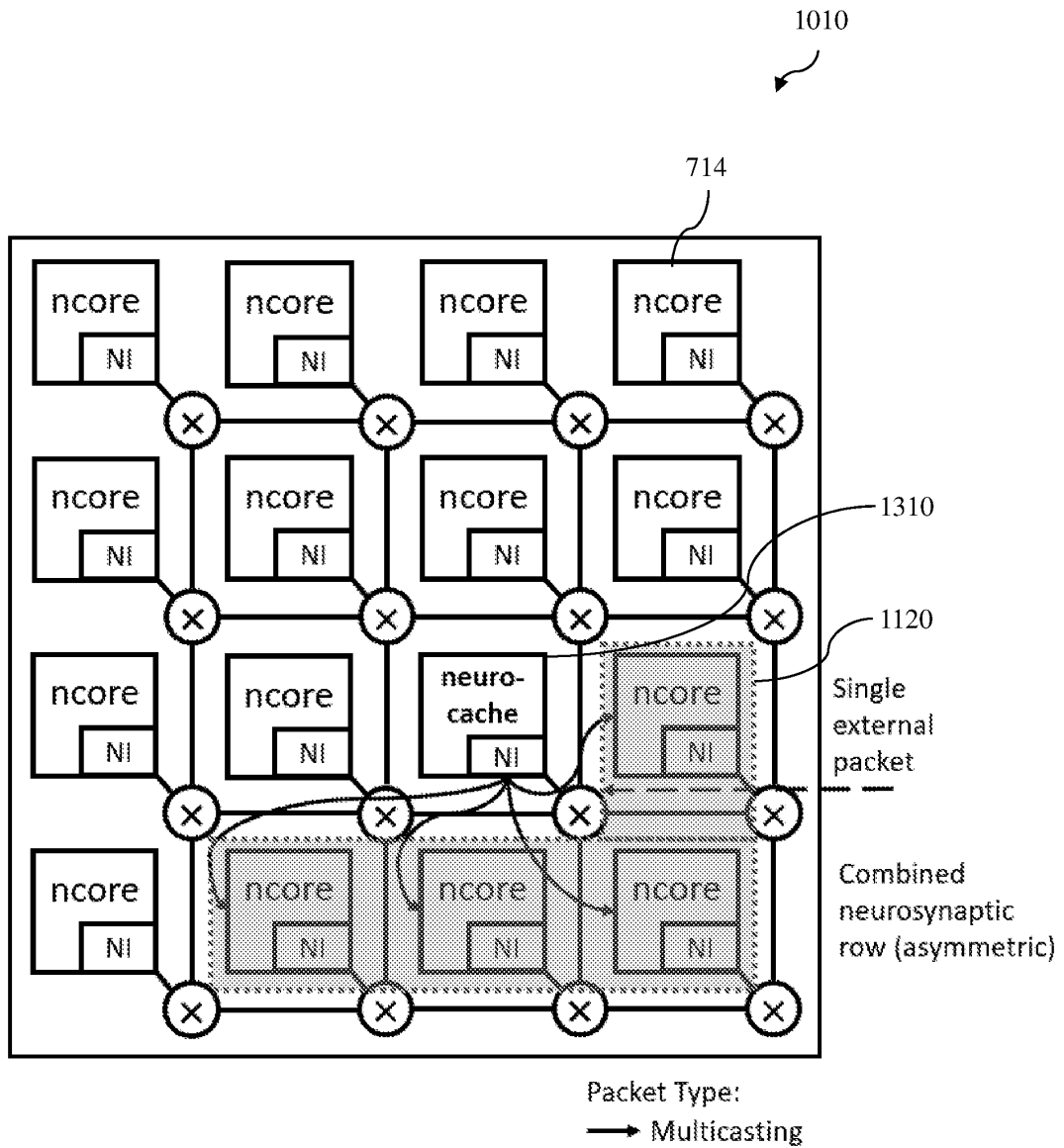


FIG. 13B

18/25

1400

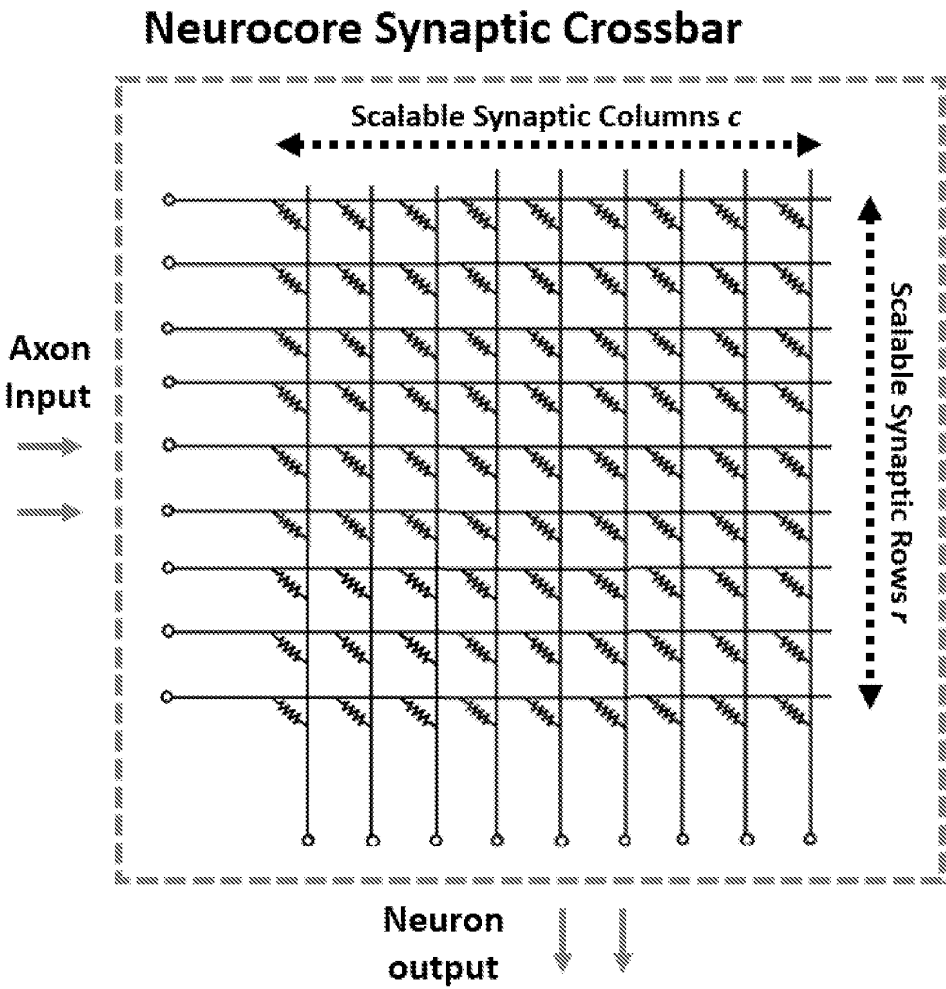


FIG. 14

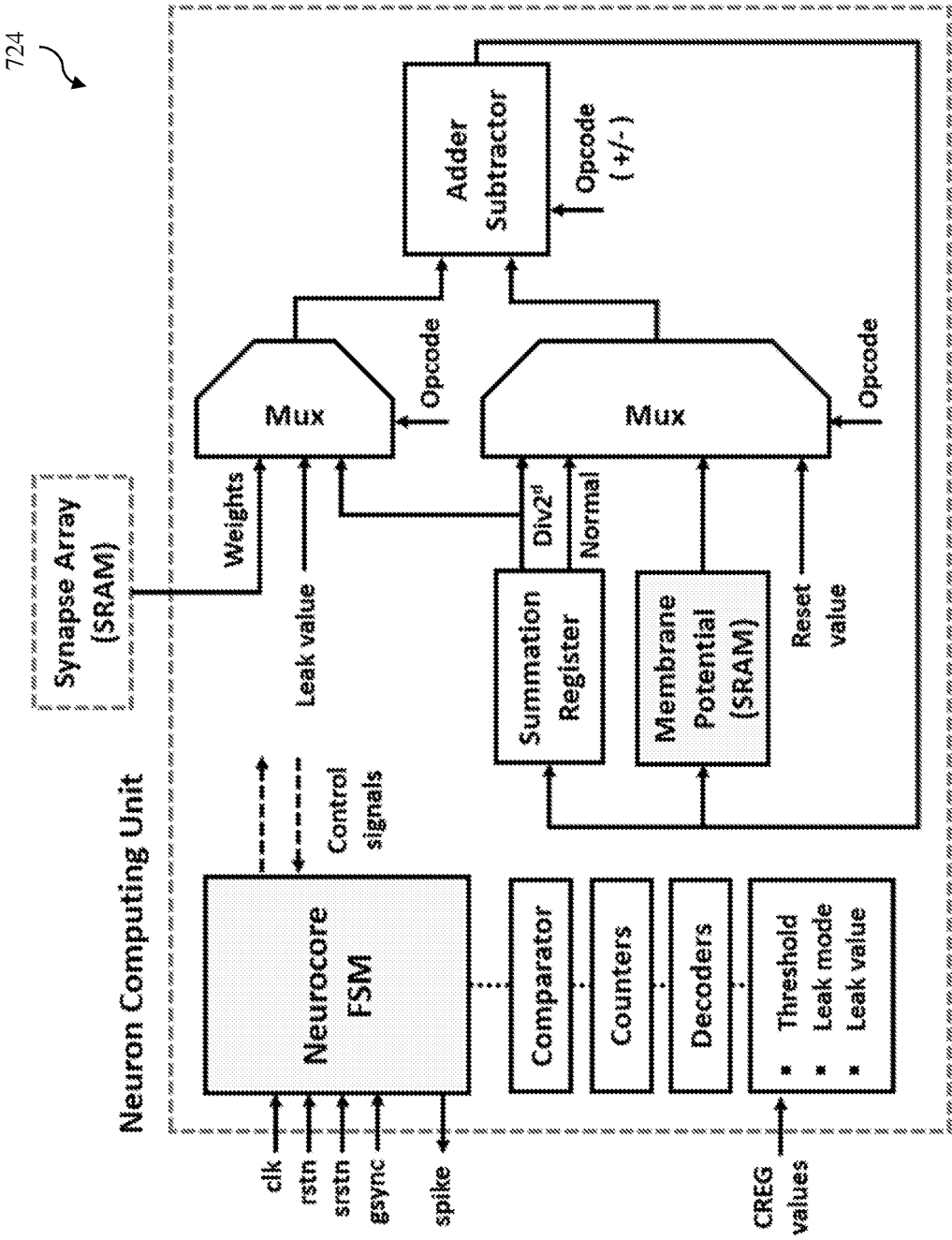


FIG. 15

20/25

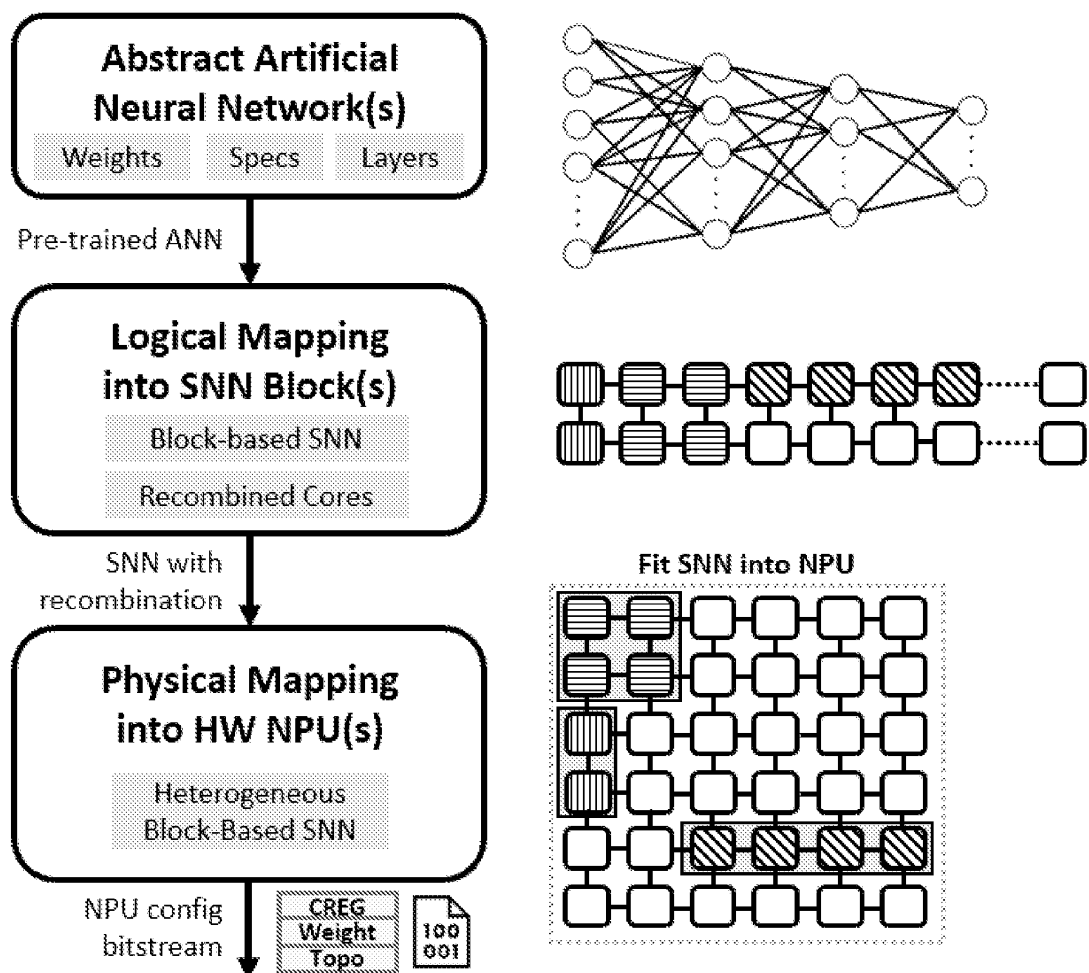


FIG. 16

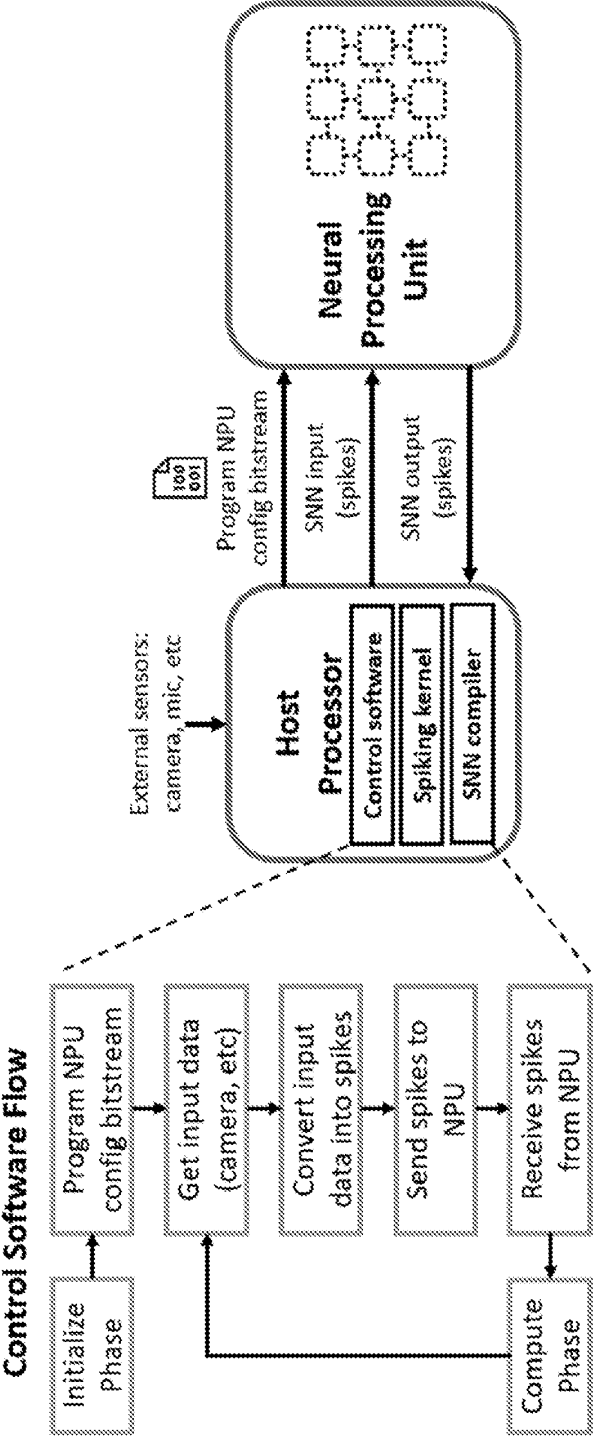


FIG. 17

Spiking CNN Parameters				# of Cores with internal synaptic array configuration as per each sub-column below				
i	Input Size	Input Kernel	N_o	256^2	512^2	1024^2	2048^2	2560^2
0	$64 \times 64 \times 3$	$5 \times 5 \times 3$	16	128	48	20	8	8
1	$32 \times 32 \times 16$	$3 \times 3 \times 16$	32	128	64	24	12	8
2	$16 \times 16 \times 32$	$3 \times 3 \times 32$	64	192	32	16	6	8
3	$8 \times 8 \times 64$	$3 \times 3 \times 64$	128	64	48	8	4	4
4	$4 \times 4 \times 128$	$3 \times 3 \times 128$	256	24	16	12	2	2
5	$2 \times 2 \times 256$	$3 \times 3 \times 256$	512	20	6	4	3	1
Total # of cores required*				428	166	64	27	23
Equivalent # of 256^2 cores*				428	664	1024	1728	2300
Total # of neurons*				109568	84992	65536	55296	58880
Total # of synapses*				28,049,408	43,515,904	67,108,864	113,246,208	150,732,800
Total # of NPU chips†				6.6875	10.375	16	27	35.9375

* Excludes the first layer (layer 0) in the neural network, this is typically handled by the host processor.
† A single NPU chip in this example has an equivalent chip area of 64 neurocores, each configured with a 256^2 synaptic array.
‡ i denotes i -th layer of the spiking CNN, N_o denotes the number of output feature maps

FIG. 18

Input size	Network Topology	Accuracy	Required core size (if there is no core recombination)
28x28	784 - 1000 - 10	97.73%	1024x1024
14x14	196 - 256 - 10	98.15%	256x256
8x8	64 - 128 - 10	91.2%	128x128

FIG. 19

Physical Core Size	Recombination	Accuracy	# of cores*	Network Topology
128×128	No	91.2%	2	64 - 128 - 10
128×128	Yes	98.15%	6	196 - 256 - 10
256×256	Doesn't Matter	98.15%	8*	196 - 256 - 10
1024×1024	Doesn't Matter	98.15%	128*	196 - 256 - 10

* Number of cores are in 128×128 core physical area equivalent, hence a single 1024×1024 core is equivalent to 64 128×128 cores.

FIG. 20

Neuromorphic architecture	Core size	Core recombination?	Accuracy	# of cores
TrueNorth	256 ²	No	89.06%	16210
This invention	256 ²	Yes	89.13%	4991
This invention	512 ²	Yes	89.13%	2359

FIG. 21