~~**CLAIMS**~~

~~WHAT IS CLAIMED IS:~~

CLAIMS

1.– A three-dimensional (3D) asynchronous network-on-chip (3D ANOC) architecture, comprising:

a 3D ultra-low power neuromorphic accelerator~~,~~ integrated into the 3D ANOC architecture, the 3D ultra-low power neuromorphic accelerator comprising:

~~a power manager;~~
a plurality of tiers;

a global power manager configured to control power for each of the plurality of tiers; and

a plurality of cores defined on each tier and coupled to the global power manager, each core comprising at least:

a local power manager configured to perform snoops on adjacent cores;

a processing element;

a non-volatile memory; and

a communications module.

2.– The 3D ~~ultra-low power neuromorphic accelerator~~ANOC architecture of claim 1, in which the ~~power manager comprises a global power manager, and each core comprises a local power manager.~~

~~3. The 3D ultra-low power neuromorphic accelerator of claim 2, in which the~~ local power manager comprises a power management integrated circuit (PMIC).

~~4. The 3D ultra-low power neuromorphic accelerator of claim 1, in which each of the plurality of cores is locally synchronous and globally asynchronous.~~

~~5. The 3D ultra-low power neuromorphic accelerator~~3. The 3D ANOC architecture of claim 1, in which the non-volatile memory comprises a hybrid distributed memory.

~~6~~4. The 3D ~~ultra-low power neuromorphic accelerator of~~ANOC architecture claim ~~5~~3, in which the hybrid distributed memory comprises magnetoresistive random access memory (MRAM), resistive random access memory (RRAM), phase-change memory (PCM), or a combination thereof.

~~7~~5. The 3D ~~ultra-low power neuromorphic accelerator~~ANOC architecture of claim 1, in which each of the plurality of tiers corresponds to a respective layer of a neural network of an edge computing platform.

~~8. The 3D ultra-low power neuromorphic accelerator of claim 1, in which the power manager determines a power state of each of the plurality of cores based on a probabilistic approximation.~~

~~9. The 3D ultra-low power neuromorphic accelerator~~6. The 3D ANOC architecture of claim 1, in which the communications module communicates with other ones of the plurality of cores on at least one of a same tier, a different tier, or a combination thereof.

~~10.~~7. The 3D ~~ultra-low power neuromorphic accelerator of claim 1, integrated into a 3D asynchronous network-on-chip (3D-~~ANOC~~)~~ architecture~~.~~

~~11. The 3D ultra-low power neuromorphic accelerator~~ of claim 1, in which the tiers are of a same processing node.

~~12. The~~8. A 3D ultra-low power neuromorphic accelerator ~~of claim 1, in which the tiers are of different processing nodes.~~

~~13. A method of fabricating a three-dimensional (3D) ultra-low power neuromorphic accelerator~~, comprising:

~~fabricating a first tier of the 3D ultra-low power neuromorphic accelerator including at least a memory unit;~~

~~fabricating a second tier of the 3D ultra-low power neuromorphic accelerator including at least a communications module;~~

~~fabricating a third tier of the 3D ultra-low power neuromorphic accelerator including at least a processing element;~~

~~stacking the second tier on the first tier and the third tier on the second tier to form the ultra-low power neuromorphic accelerator; and~~

~~fabricating a power manager to adaptively scale a voltage of the ultra-low power neuromorphic accelerator.~~

~~14. The method of claim 13, further comprising fabricating communications channels between at least the first tier, the second tier, and the third tier.~~

~~15. The method of claim 13, further comprising fabricating the processing element, the memory unit, and the communications module on each of the first tier, the second tier, and the third tier.~~

~~16. The method of claim 13, in which fabricating the power manager comprises:~~

~~fabricating a local power manager on each of the first tier, the second tier, and the third tier; and~~

~~configuring a global power manager to coordinate with the local power manager on each of the first tier, the second tier, and the third tier to adaptively scale the voltage of the ultra-low power neuromorphic accelerator.~~

~~17. The method of claim 16, further comprising:~~

~~determining a power state of one of the plurality of cores based on the power state of an adjacent one of the plurality of cores.~~

~~18. The method of claim 13, in which the third tier of the 3D ultra-low power neuromorphic accelerator is fabricated using a process node different from a process node used to fabricate the first tier and/or the second tier of the 3D ultra-low power neuromorphic accelerator.~~

~~19. A three-dimensional (3D) ultra-low power neuromorphic accelerator, comprising:~~

~~a communications and power management tier comprising a plurality of communications modules and a plurality of local power managers;~~

~~a memory tier comprising a plurality of memory units; and~~

~~a logic computation tier comprising a plurality of neural processing units, the logic computation tier coupled to the communications and power management tier and the memory tier, each neural processing unit coupled to one of the plurality of memory units and one of the plurality of the local power managers.~~

~~20. The 3D ultra-low power neuromorphic accelerator of claim 19, further comprises~~ a plurality of tiers;

a global power manager configured to ~~coordinate the plurality of the local power managers to adaptively scale a voltage of the ultra-low power neuromorphic accelerator.~~control power for each of the plurality of tiers; and

a plurality of cores defined on each tier and coupled to the global power manager, each core comprising at least:

a local power manager configured to perform snoops on adjacent cores;

a processing element;

a non-volatile memory; and

a communications module, in which each of the plurality of cores is locally synchronous and globally asynchronous.

~~21. The~~9. A 3D ultra-low power neuromorphic accelerator, comprising:

a plurality of ~~claim 19, in which in which~~tiers;

a global power manager configured to control power for each of the plurality of ~~memory units comprise a hybrid distributed~~tiers; and

a plurality of cores defined on each tier and coupled to the global power manager, each core comprising at least:

a local power manager configured to perform snoops on adjacent cores;

a processing element;

a non-volatile memory~~, comprising magnetoresistive random access memory (MRAM), resistive random access memory (RRAM), phase-change memory (PCM), or a combination thereof. ~~; and

a communications module, in which the global power manager determines a power state of each of the plurality of cores based on a probabilistic approximation.

~~22. The~~10. A 3D ultra-low power neuromorphic accelerator ~~of claim 19, in which each of the plurality of logic computation tiers corresponds to a respective layer of a neural network of an edge computing platform.~~comprising:

a power manager;

a plurality of tiers; and

a plurality of cores defined on each tier and coupled to the power manager, each core comprising at least:

a processing element;

a non-volatile memory; and

a communications module, in which the tiers are of different process nodes.