



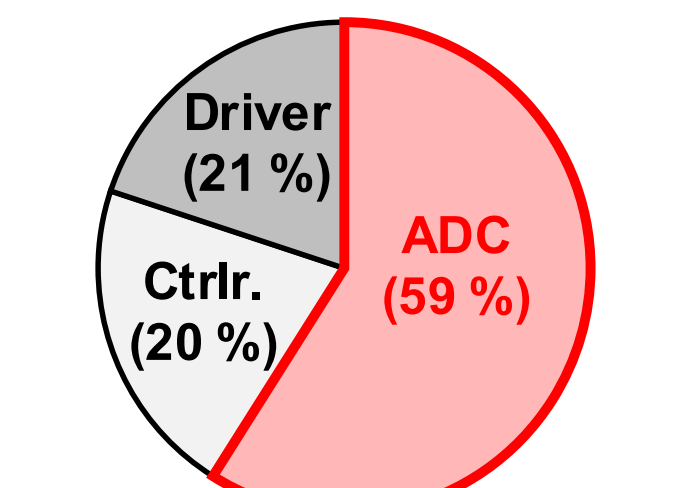
A Reconfigurable 1T1C eDRAM-based Spiking Neural Network Computing-In-Memory Processor for High System-Level Efficiency

Seryeong Kim, Soyeon Kim, Soyeon Um, Zhiyong Li, Sangyeob Kim, Wooyoung Jo, and Hoi-Jun Yoo

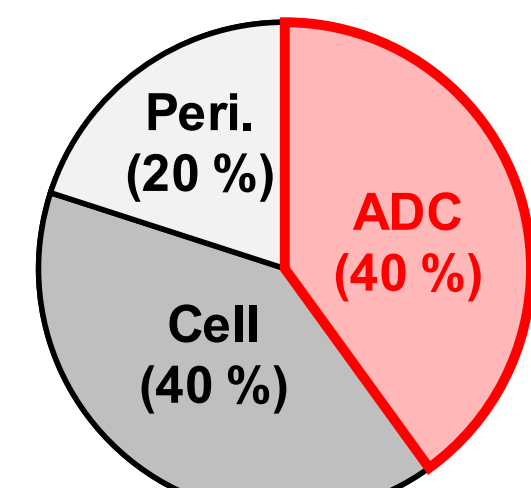


Introduction

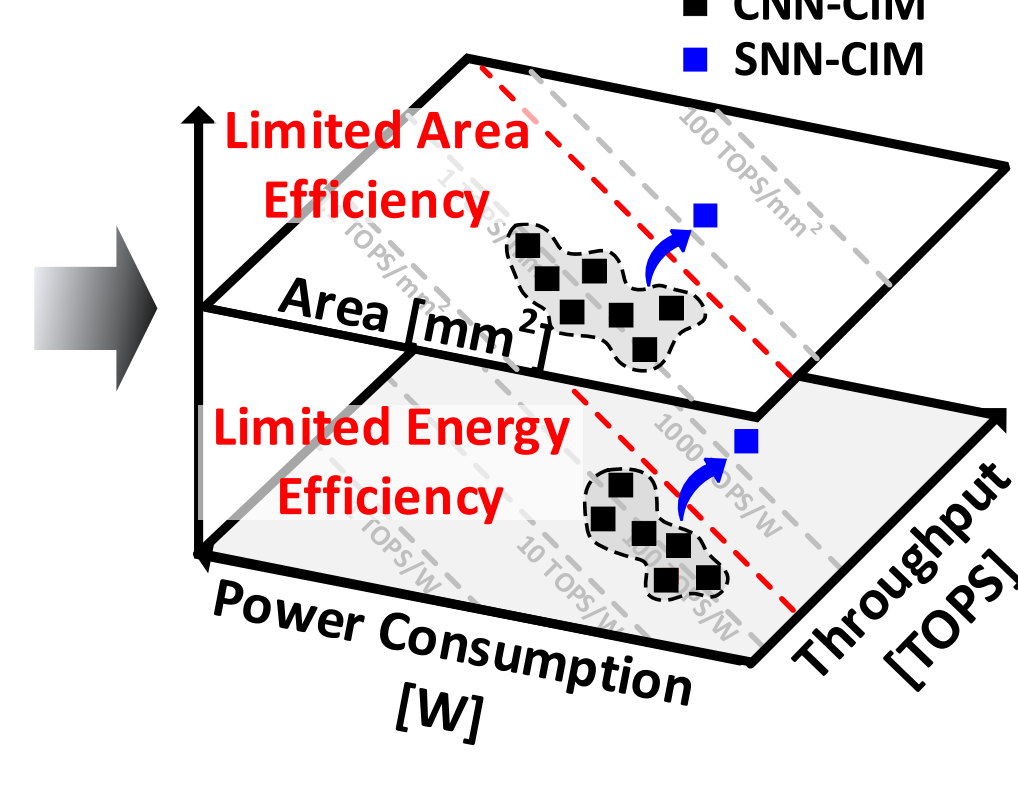
Problems of Computing-in-Memory



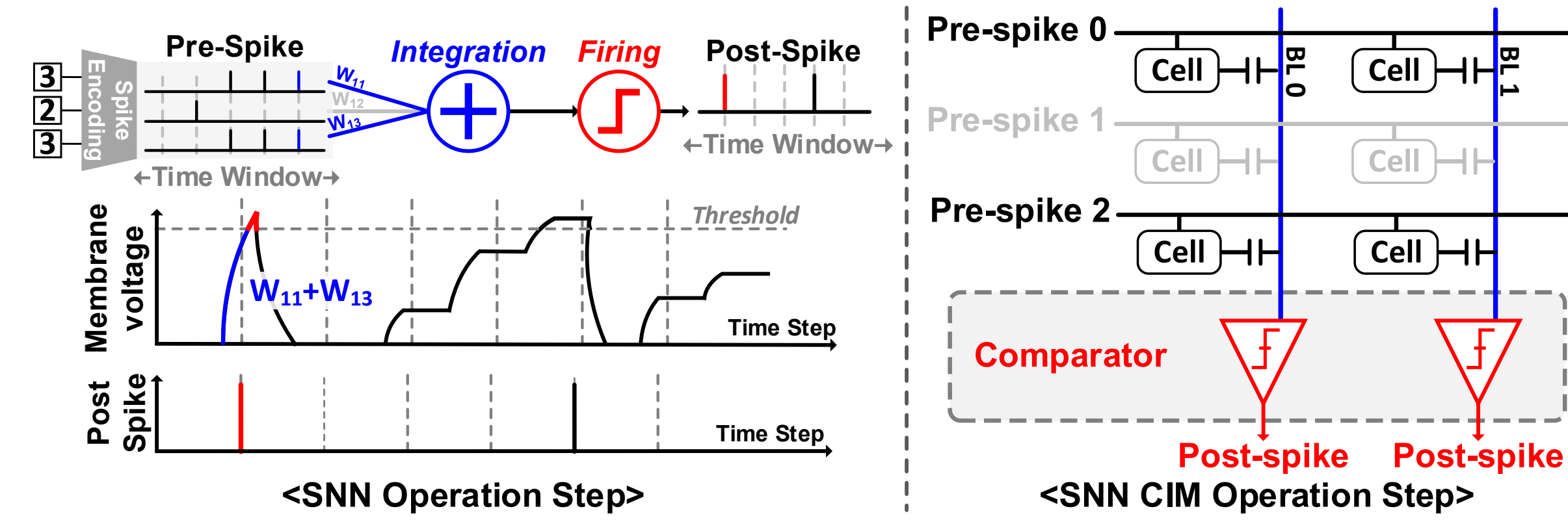
<Problem 1>
High ADC Power (> 50%)
→ Limited Energy Efficiency (TOPS/W)



<Problem 2>
Large ADC Area (> 40%)
→ Limited Area Efficiency (TOPS/mm²)

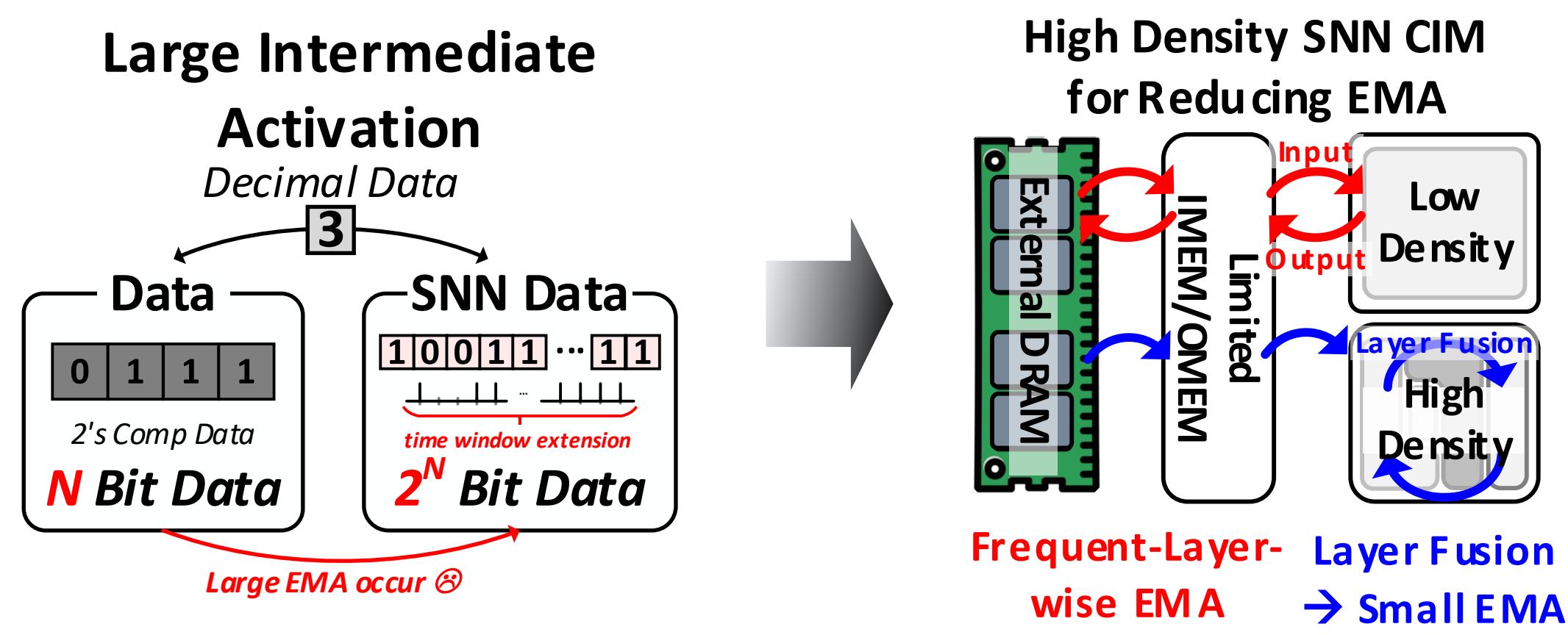


High Macro-Level Efficient SNN CIM



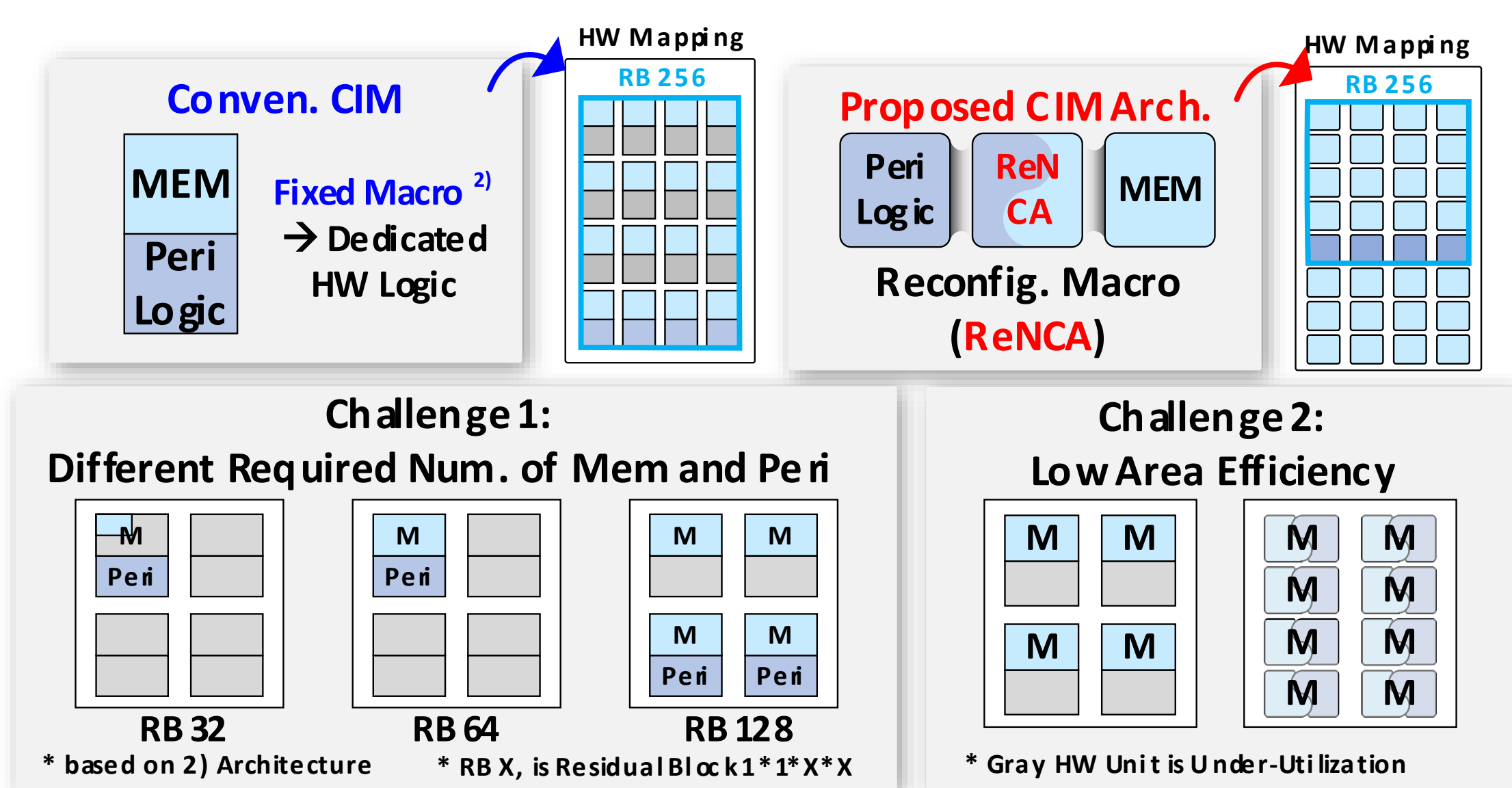
→ ADC is not needed → Power/Area Reduction ☺

Low System-Level Efficient SNN CIM



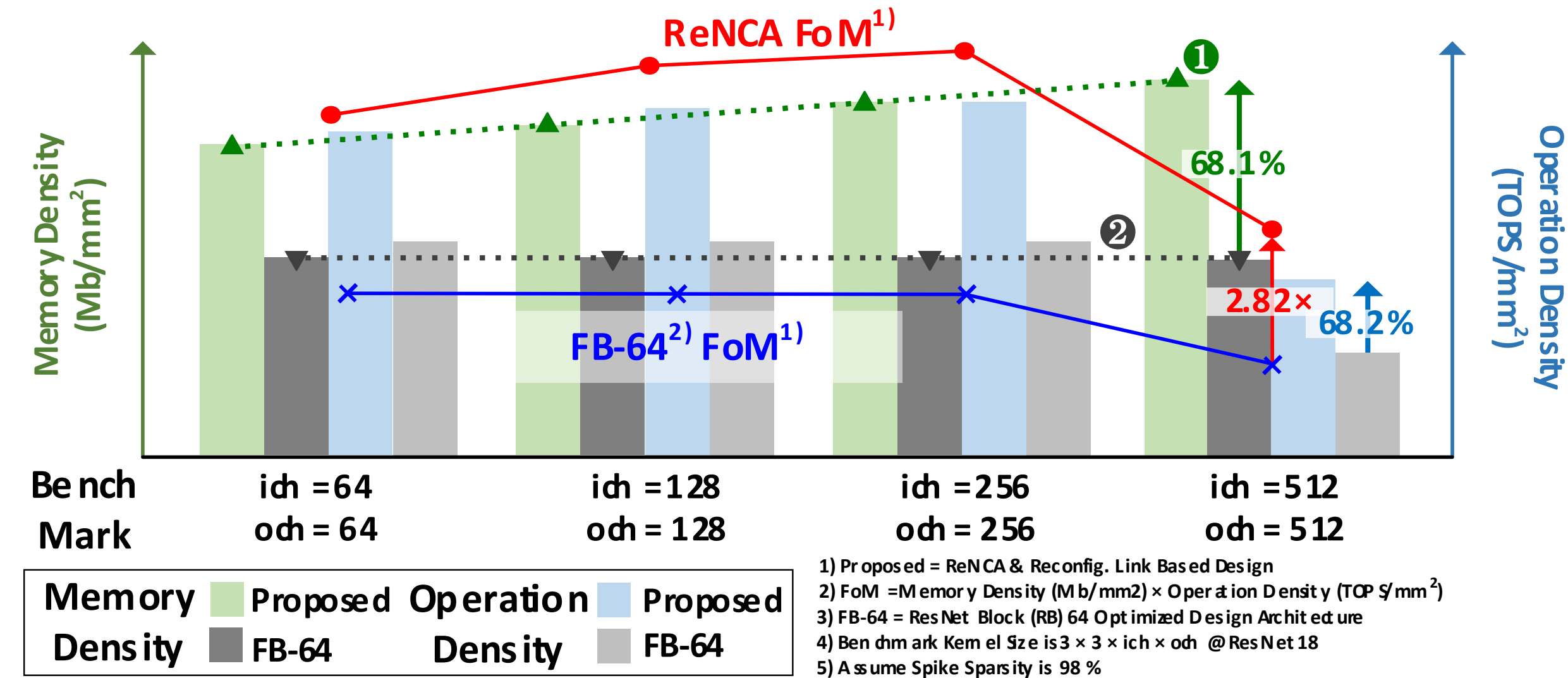
Reconfigurable Membrane Network

Layer-wise Adaptive Accumulation Link



- Previous Fixed Macro Configuration
- HW Under-Utilization according to Various Layer Configuration
- Support Dynamic Accumulation & Data Path

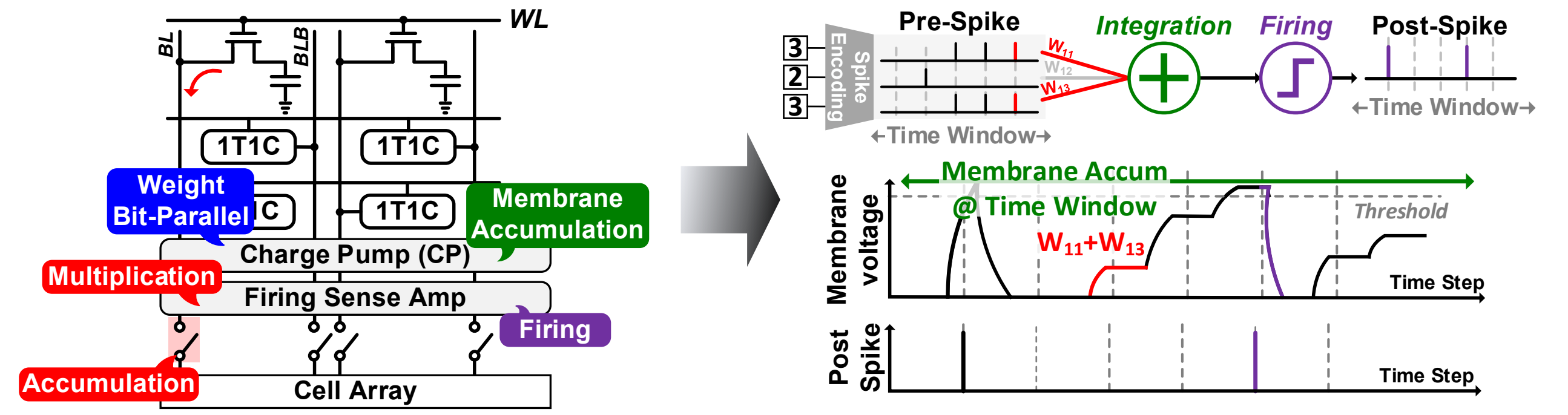
<System Performance Summary>



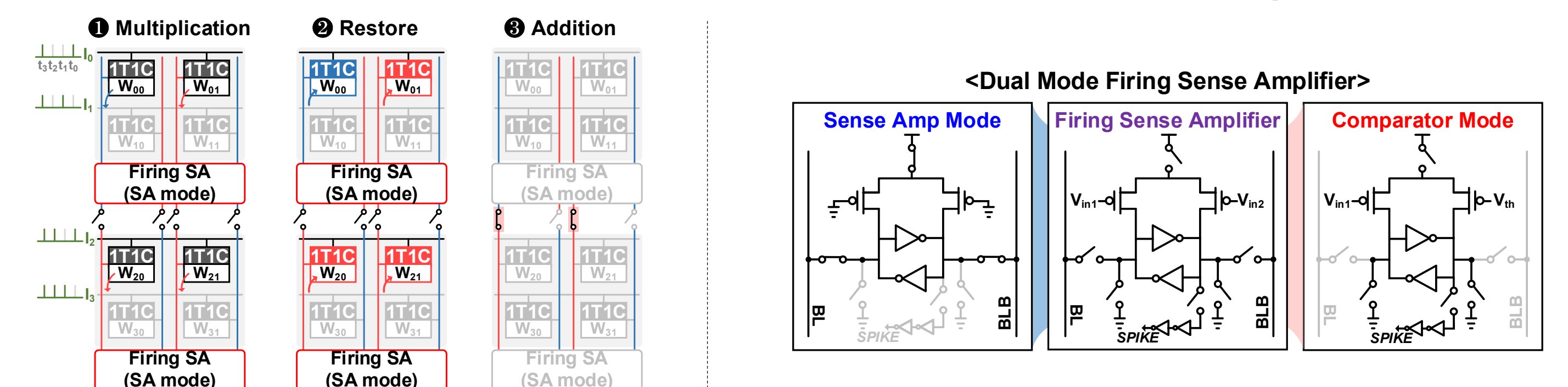
1) Proposed = ReNCA & Reconfig. Link Based Design
2) FoM = Memory Density (Mb/mm²) × Operation Density (TOPS/mm²)
3) FB-64 = ResNet Block (RB) 64 Optimized Design Architecture
4) Benchmark Mem el Size is 3 × 3 × ich × ocd @ ResNet18
5) Assume Spike Sparsity is 98%

Reconfigurable Neuro-Cell Array (ReNCA)

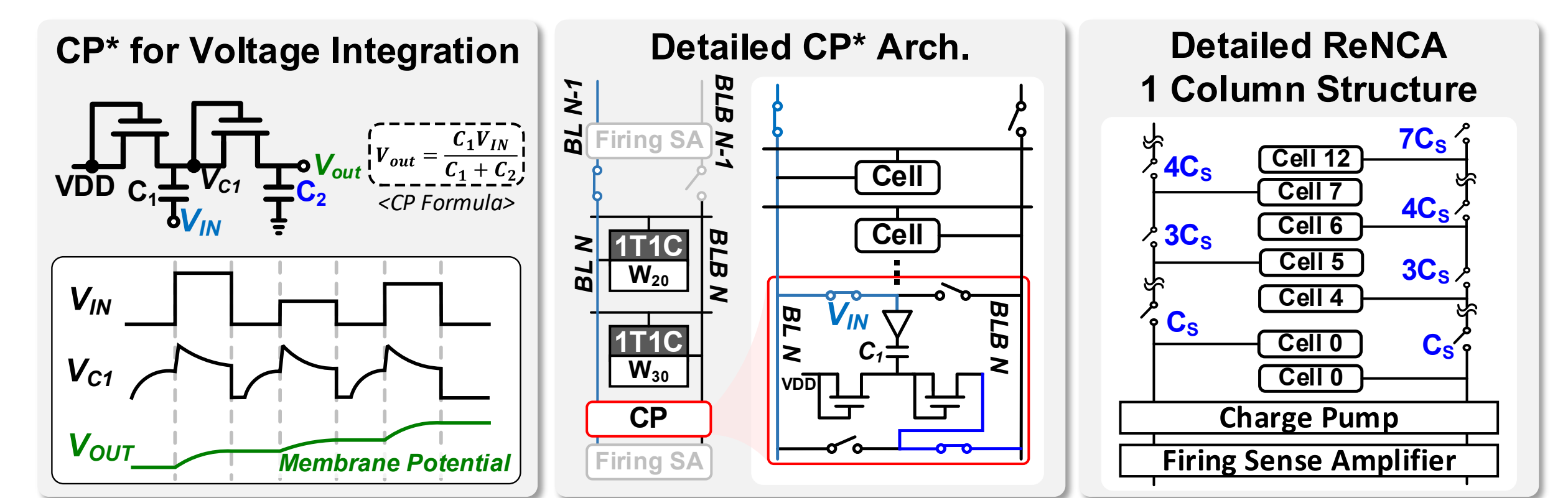
High Density ReNCA Overall Architecture



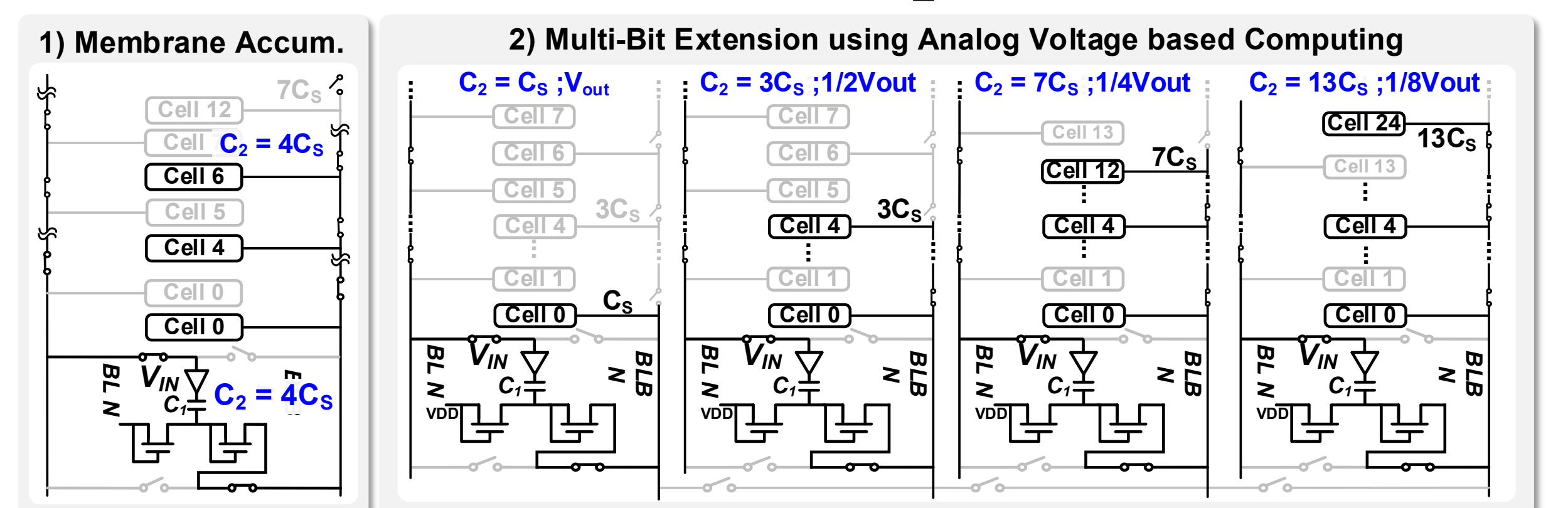
Sense Amplifier based MAC OP. & Firing



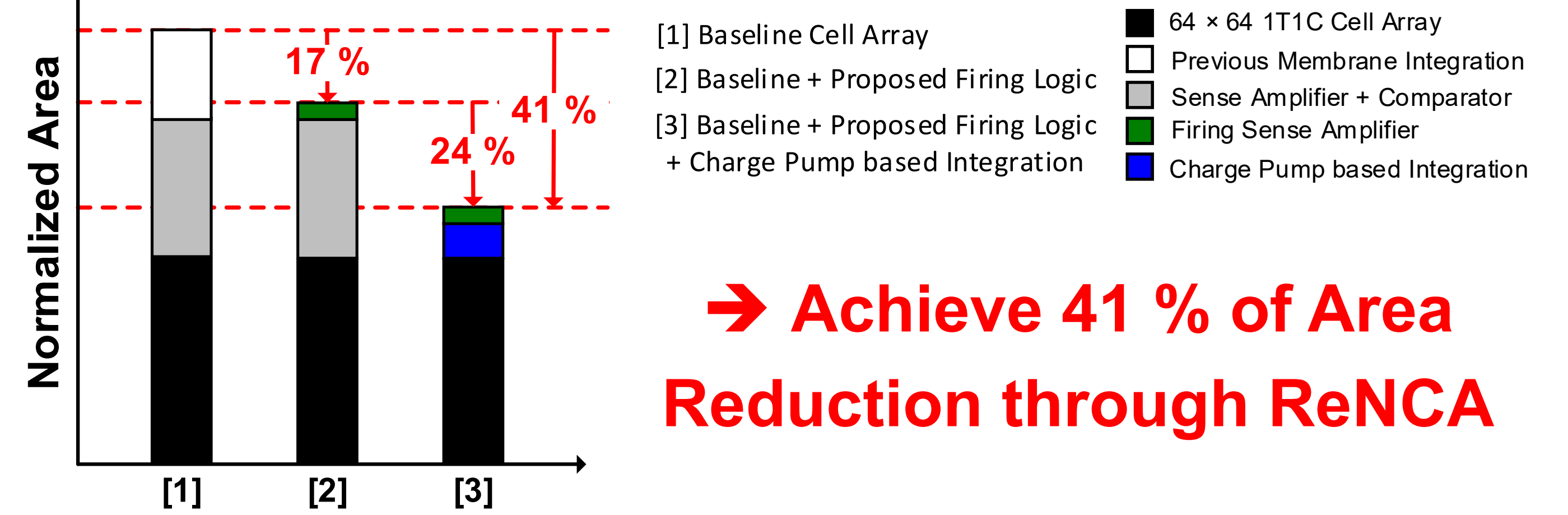
CP* for Membrane Acc. & Multi-Bit Extension



Reusing cell capacitor as CP C₂ using BL switch



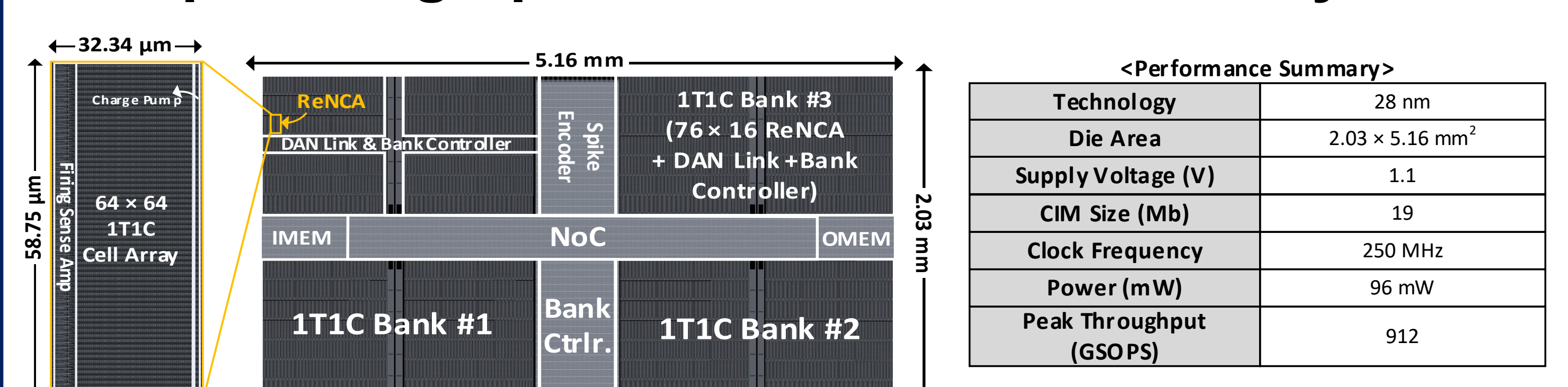
<Area Optimization Summary>



→ Achieve 41% of Area Reduction through ReNCA

Performance & Conclusion

Chip Photograph & Performance Summary



- 1T1C eDRAM CIM Macro for High Density
- Dynamic Adjustable Neuron Link for Reconfigurability
- A 157.15 TOPS/W 1T1C eDRAM-based SNN-CIM Processor for High System-Level Efficiency