# Review: "Ambit: In-Memory Accelerator for Bulk Bitwise Operations Using Commodity DRAM Technology"

## Summary

The paper outlines the restrictions imposed by the memory bandwidth bottleneck and proposes a solution through bitwise memory operations on entire DRAM lines. It does this in the following steps:

- The paper explains that bulk bitwise memory operations — AND, OR, NOT, etc. — performed on large stretches of contiguous data are very inefficiently handled in conventional systems, as these are simple instructions that must be repeated over large amounts of data and thus become bottlenecked by the memory bandwidth.

- The paper introduces the concept of in-memory bulk bitwise operations. It explains a system using standard DRAM design to utilize simultaneous loading/enabling of memory lines to perform AND and OR operations on entire memory lines. In short: this is possible because the bitlines span orthogonally to the word lines through the DRAM and share a single sense amplifier for all entries. By activating three lines at the same time, the voltage over the three lines is averaged. This gives us a control line **C** and two input lines **A** and **B**, on which we can perform the following operations: `C(A+B) + C'(AB)`. These correspond to a logic **AND** for `C = 0` and a logic **OR** for `C = 1`.

- The paper also extends the DRAM structure to allow for logic **NOT** operations by connecting the inverted signal of the bitline to a dual-contact DRAM cell.

- The paper names the accelerator **Ambit** and shows that it can be easily integrated into a system, as its interface is identical to that of common DRAM, allowing for seamless integration.

- The paper shows that Ambit has both low hardware and control overhead compared to conventional DRAM.

- The paper evaluates the performance of Ambit by performing circuit-level SPICE simulations as well as analyzing real-world example problems simulated on the system. The SPICE simulation shows that the error rate is low for reasonable process variation and further elaborates that faulty modules could still be used as conventional DRAM, as the Ambit accelerator is an extension of conventional DRAM. The runtime performance analysis shows that for database access with bitmap indices and BitWeaving, Ambit provides significant speedup. Furthermore, a bitvector implementation with Ambit is compared to an RB-Tree implementation, and it is shown that for large sets Ambit yields a significant speedup.

## Strengths

Ambit uses existing technology by being an extension to conventional DRAM. It is also a very smart utilization of existing processes, resulting in a new way to compute within memory, allowing it to alleviate the memory bandwidth bottleneck for specific types of computations. The paper also proposes that faulty Ambit accelerators could still be used as conventional DRAM, thereby reducing yield costs for the manufacturer.

## Weaknesses

The paper does not provide an actual manufactured Ambit accelerator, and therefore any of the results presented are based solely on simulations. It is fair to assume that these simulations are highly accurate, but simulated data can never fully replace real-world measurements; therefore, the claims must be taken with a grain of salt.

Ambit only implements a very limited instruction set with AND, OR, and NOT operations. While other binary operations can be constructed from these, it is also outlined that the successive loading of memory into the relevant DRAM lines is time-intensive. The paper itself does not propose any such implementation, and therefore I have to assume that the use case of the accelerator is limited to situations where these specific operations are the primary bottleneck.

## My POV

I would see great value in testing a physical implementation of the Ambit accelerator to obtain real-world data to back up the simulations. In general, the paper alludes to steps that manufacturers still need to take to actually produce a working system, which somewhat undermines the simulated results.

Further exploration of chaining the AND, OR, and NOT operations would also have great value, as it would demonstrate to what degree more complex logic could be performed directly on the accelerator.

## Takeaways

- Ambit introduces a very powerful solution to parts of the memory bandwidth bottleneck. The concept of computing not just *in memory* but *with memory* seems very powerful and worthy of further exploration.

- The results of the paper are based on simulation rather than physical integration, meaning that all results must be taken with a grain of salt.