

# Review: “Processing in Memory: The Terasys Massively Parallel PIM Array”

## Summary

The paper addresses the memory bottleneck. And proposes an in-memory-processing solution called Terasys. The paper provides a physical implementation as well as a C extension called dbC as an API with the accelerator. It also explores applications and gives a back of the envelope cost estimation. - The paper proposes Terasys an in-memory-processor that consists of a large number of processing elements (256x256) which are located in memory and perform single-instruction-multiple-data, abbreviated as SIMD, operations. All the processing elements are controlled by a global control unit. The system is scalable as the computational power grows in tandem with the memory. - The paper provides dbC which is an extension to C that allows interfacing with the Terasys accelerator. It introduces SIMD structures into C to allow tailored code to optimize the accelerator usage. - The paper outlines that Terasys has superior performance to conventional systems in image processing and matrix operations. As Terasys is located in memory the system achieves higher performance and lower energy usage than multiprocessor systems. - The paper provides a proof of concept real-world implementation to back up the claims in the paper. This backs up their claims about performance and power usage.

## Strengths

The paper is an early instance of in-memory-processing. It provides a fast and working solution to real world problems such as the MRI imaging problem outlined in the paper. The paper lies important ground work for later work in the area of in-memory-processing. Further it also provides a cost analysis giving transparency about use cases of the Terasys.

## Weaknesses

The paper has fairly few concrete numbers and results. This is an issue as it complicates the assessment of the system’s actual performance and viability in diverse scenarios. The proposed extension to C does not follow a general standard which can be more confusing than helpful in certain scenarios and which also greatly reduces the interoperability of the system. The paper also claims the system to be scalable and easy to integrate into greater systems. While their anecdotal evidence of integrating Terasys into an existing com-

pute system and integrating it into a supercomputer seems to back this up, there is a lack of evidence to substantiate such a claim in its general form.

## My POV

The paper was an early adopter of in-memory-processing and it is easy to propose improvements from a modern viewpoint. But this viewpoint wouldn’t exist if not for papers such as this that lay ground work for all the following research. While dbC did not catch on, the concept of in-memory-processing very much did. I would prefer it if the paper would have provided more actual data so that I could further assess the actual performance of the system. Nonetheless it is clear that Terasys’s concepts were a success as we find similar architecture in modern state-of-the-art GPUs and compute capable DRAM.

## Takeaways

- Terasys was an important early proof of concept for in-memory-processing that fueled research to this day and provided the foundation for many modern architectures.
- The paper highlights the importance of backing concepts with physical implementations to provide palpable results on which future research can build and to which it can compare itself.