Thesis Topic:

Hardware and Software Support for Hybrid PIM-CPU Architectures

---

Background:

The integration of Processing-In-Memory (PIM) with traditional CPU architectures is emerging as a suitable solution to meet the increasing demands of data-intensive applications. This hybrid approach leverages the strengths of both data-centric and CPU-centric processing, enabling efficient handling of large datasets and complex computations while reducing memory-processor data transfers. However, one of the significant challenges in these hybrid systems is ensuring data reliability, especially in the presence of silent data corruptions (SDCs) caused by bit errors.

While the addition of on-die Error Correction Code (ECC) in memory systems provides the illusion of errorless memory, capable of correcting and detecting single and double bit errors, respectively, it cannot fully eliminate the risk of SDCs. In reality, varying degrees of bit error rates persist, which, if undetected, can lead to system crashes, not to mention that ECC is not fully avaliable in some occasions. These bit flips are often localized, affecting neighboring bit positions within a word. To mitigate this risk, transitioning from a single continuous data representation (word view, e.g., 32-bit word) to a distributed representation (radix view, e.g., 8-bit radix), coupled with ECC protection at the radix level, can significantly reduce the likelihood of multiple bit flips within a single word, thereby enhancing system reliability.

Project Description:

This thesis will focus on developing hardware mechanisms and software strategies that optimize data representation to improve the reliability and efficiency of hybrid PIM-CPU architectures. The research will cover several key areas:

1. Data Representation Optimization: Investigate and implement a dual representation of data in both word view and radix view.

   ```
                              The aim is to reduce the impact of bit errors by distributing data more granularly and applying ECC at the radix level,
                              thus minimizing the risk of multiple bit flips within a single word.
   ```

2. Memory Management Techniques: Create memory management strategies that support dynamic transitions between word and radix views.

   ```
                 These techniques will be designed to maximize the efficiency of data-centric operations in PIM-CPU systems while safeguarding against SDCs.
   ```

3. Compiler Support: Implement compiler optimizations that facilitate the transparent allocation and management of memory across different data representations,

   ```
                 ensuring that applications can leverage these improvements without manual adjustments.
   ```

Key References:

1. High-Density Image Storage Using Approximate Memory Cells
2. Zero-Space Cost Fault Tolerance for Transformer-based Language Models on ReRAM
3. HEART: Hybrid Memory and Energy-Aware Real-Time Scheduling for Multi-Processor Systems

---

Student Profile:

Skills Required:

- Proficiency in C/C++ programming

Desired Knowledge:

- File systems
- Computer Architecture
- System Programming

Supervisor/Contact Information:

- **Name:** Dr. Ir. Kuan-Hsun Chen
- **Email:** k.h.chen@utwente.nl
- **Name:** Dr. Yun-Chih Chen (Remote in Germany)
- **Email:** yunchih.chen@tu-dortmund.de

How to Apply:

Send an e-mail to k.h.chen@utwente.nl, write a bit about your motivation to work on the subject, your background, and at what stage of the master program you are. We will then set a meeting.
