

Ottimizzazione della gerarchia della memoria di un processore ad elevate prestazioni

Il progetto di ricerca si concentrerà sull'ottimizzazione della gerarchia di memoria in un processore di classe applicativa, intervenendo direttamente sul design RTL e sulla sua integrazione architetturale. Nella fase iniziale, verrà definita una microarchitettura di riferimento con obiettivi di throughput e latenza, identificando colli di bottiglia legati alla gerarchia di memoria. Si esploreranno soluzioni di caching multi-livello, buffer locali, e interconnessioni ottimizzate, con particolare attenzione alla compatibilità con pipeline profonde e parallelismo dei dati. Le modifiche saranno implementate in RTL e validate tramite simulazioni funzionali e timing-accurate. Successivamente, il progetto prevede la sintesi fisica e la place & route per valutazioni su area, potenza e frequenza operativa. La soluzione verrà infine mappata su FPGA per la validazione con workload realistici, incluse applicazioni embedded e AI. I risultati guideranno l'identificazione di configurazioni architetturali ottimali per futuri tape-out.

Optimization of the memory hierarchy of an application class processor

The research will focus on optimizing the memory hierarchy of an application-class processor, acting directly on the RTL design and its architectural integration. The initial phase will define a reference microarchitecture with performance goals in terms of throughput and latency, identifying bottlenecks related to the memory subsystem. Solutions such as multi-level caches, local buffers, and optimized interconnects will be explored, with a focus on compatibility with deep pipelines and data parallelism. These enhancements will be implemented at RTL level and verified through functional and timing-accurate simulations. The project will then proceed to physical synthesis and place & route to assess area, power, and operating frequency. The final architecture will be mapped on FPGA platforms for validation with realistic workloads, including embedded and AI applications. Results will support the identification of optimized memory configurations suitable for future tape-outs.