TMFab: A Transactional Memory Fabric for Chip Multiprocessors

Sumeet S. Kumar, Rene van Leuken,
Delft University of Technology,
EEMCS, Circuits and Systems Group,
Mekelweg 4,
Delft 2628CD,
The Netherlands
{s.s.kumar, t.g.r.m.vanleuken}@tudelft.nl

Abstract

With the performance of single-core processors approaching its limits, an increased amount of research effort is focused on chip multiprocessors (CMP). However, existing lock-based synchronization methods that are critical to performing parallel computation possess limited scalability and are inherently complex to use while programming. This thesis uses the concept of transactional memory implemented within a synthesizable fabric named TMFab, containing all the requisite hardware components needed to prototype a scalable chip-multiprocessor. Its processor independent nature enables the instantiation and use of any suitable soft-processor core inside the fabric without significant modifications to the fabric hardware. Additionally, the fabric offers scalability on account of its 3D interconnect architecture that supports die-stacking to add additional processor cores to the CMP without increasing its area footprint. The hardware transactional memory system of the fabric reduces performance overheads of transactional operations, allowing transactions to complete execution faster. TMFab is shown to provide speed up as high as 3.44× for a 4 processing element (PE) CMP with correctly partitioned independent transactions and can be used to analyze the points of contention for conflicting transactions. The fabric was synthesized for both Field Programmable Gate Array (FPGA) as well as 90nm semi-custom targets.
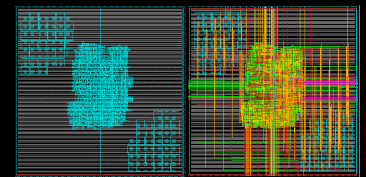
Circuits and Systems
TMFab
A Transactional Memory Fabric for Chip Multiprocessors

Overview

- **Retargetable** infrastructure for fast prototyping of Chip Multiprocessors *(CMP)*

- Speculative lock-free execution using **Hardware Transactional Memory**

- Scalable fabric supporting stacked-die implementations using advanced **Through Silicon Via (TSV)** based interconnect
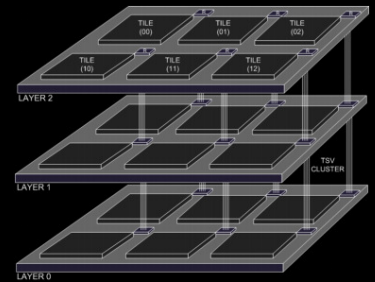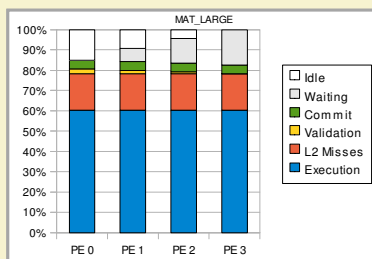

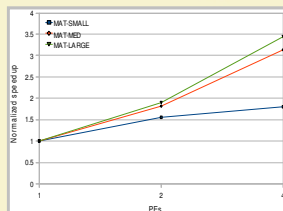Stacked-die implementation using TSV based 3D Network on Chip

Transactional Memory

Stages of transactional execution

- Lazy Version Management
- Pessimistic Conflict Detection
- Aggressive Retry
- Simple programming methodology
- Low performance overhead for Validate-Commit
- *Scalable interconnect* augmented with TM functions
- Processing Element *(PE)* architecture *independent* design

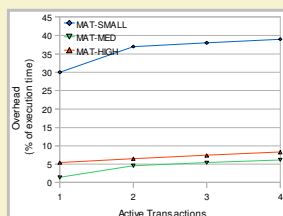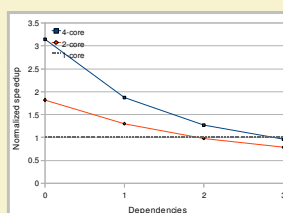Layout of 7-port 3D router with full-custom TSVs

Performance

Normalized speedup

- Performance evaluated with *32-bit pipelined RISC* processor based 4 core CMP. RTL model of complete system simulated at 200MHz

- Best case speedup of *3.44x*

- Observed *reduction in validate-commit time* with progress of execution

- *Cache coherence* maintained by system level transactional memory policy

Reduction in performance overhead with execution
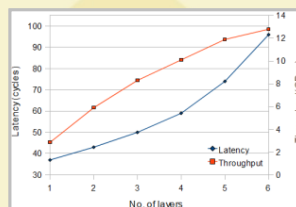
Interconnect performance with stacking

- Stacking observed beneficial upto *five layers* using the *baseline 3D mesh interconnect*

- Interconnect synthesized in *90nm UMC* with *Faraday* standard cell library and *full-custom TSVs*

Effect of inter-transaction dependencies on speedup
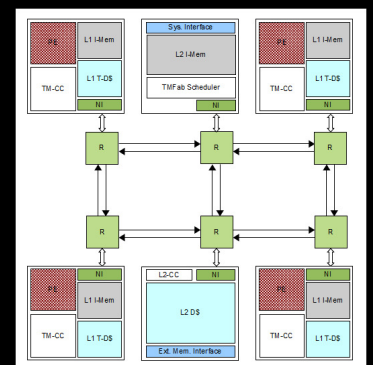
TMFab architecture for 4 PE chip multiprocessor
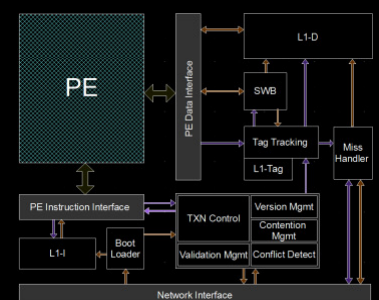
PE architecture independent TMFab Cache Controller (TM-CC)

Authors
Sumeet S. Kumar
Rene van Leuken

Circuits and Systems Group

Faculty of Electrical Engineering, Mathematics, Computer Science

TUDelft