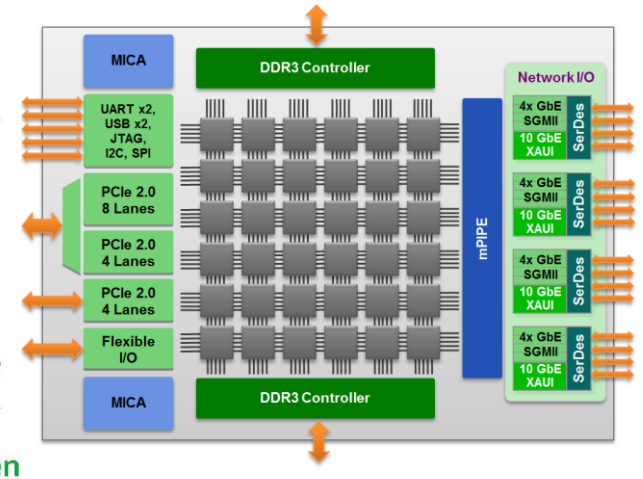
TSHMEM Introduction

HPC acceleration with SHMEM on many-core processors

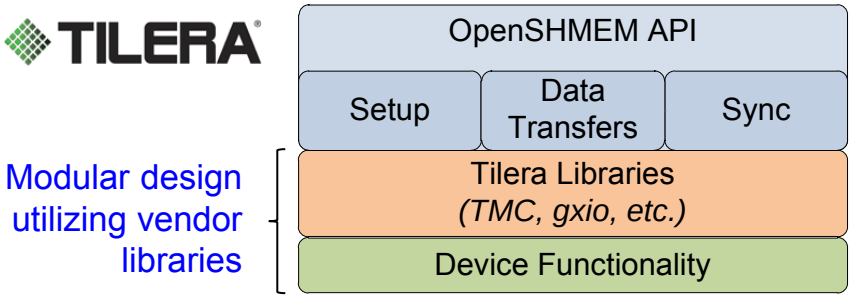


Investigate SHMEM reference design directly over *Tilera TILE-Gx* architecture and libraries

- Stay true to SHMEM principles
  - High performance with low overhead
  - Portability with easy programmability
- Maximize architectural benefits
  - Tile interconnect, mPIPE, MiCA
- Extend design to multi-device systems
  - Evaluate interconnect capabilities
  - Explore optimizations to point-to-point transfers and collectives

Open SHMEM

TILERA



Modular design utilizing vendor libraries

TSHMEM reference design on TILE-Gx36

OpenSHMEM and TSHMEM

Achieved

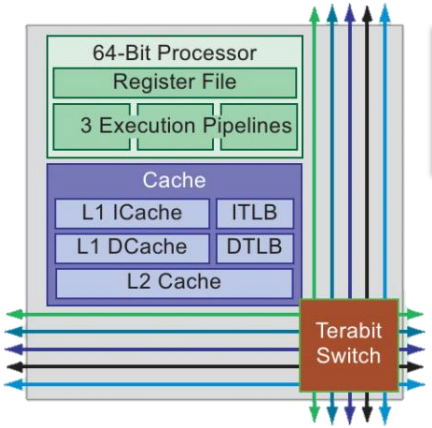
- Dynamic symmetric heap management
- Point-to-point data transfer
- Point-to-Point synchronization
- Barrier synchronization
- Broadcast, Collection, Reduction
- Atomics for dynamic variables
- Extension to multiple many-core devices

Ongoing

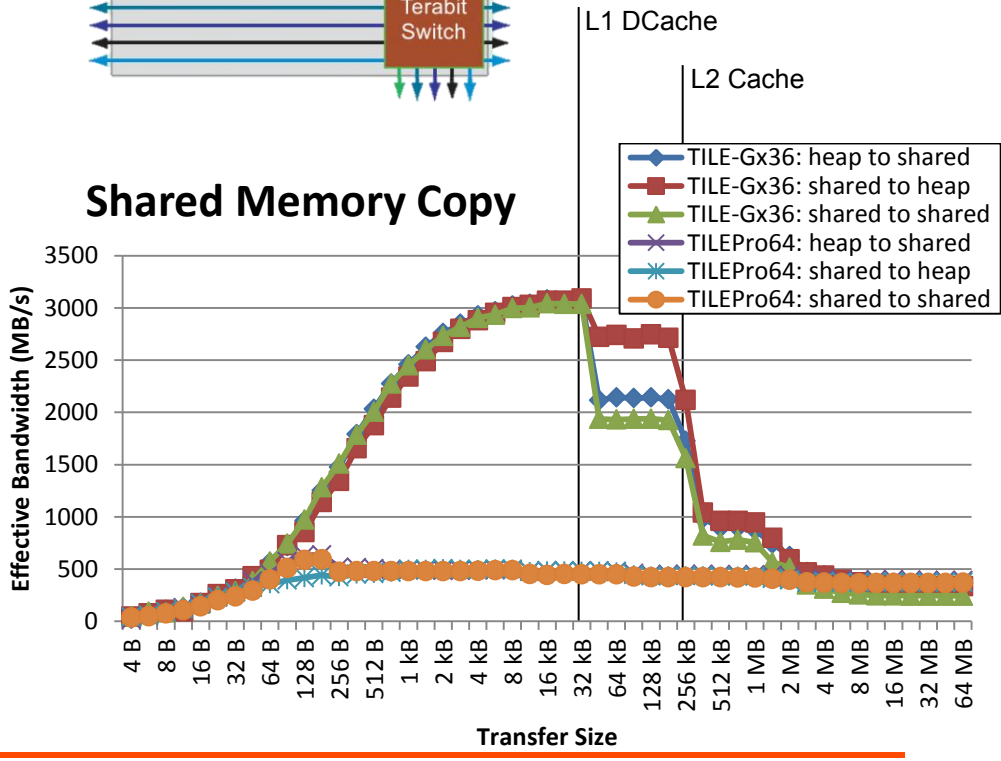
- Optimizations for multiple many-core
- Exploration of new SHMEM extensions

Shared-Memory-Copy Bandwidth

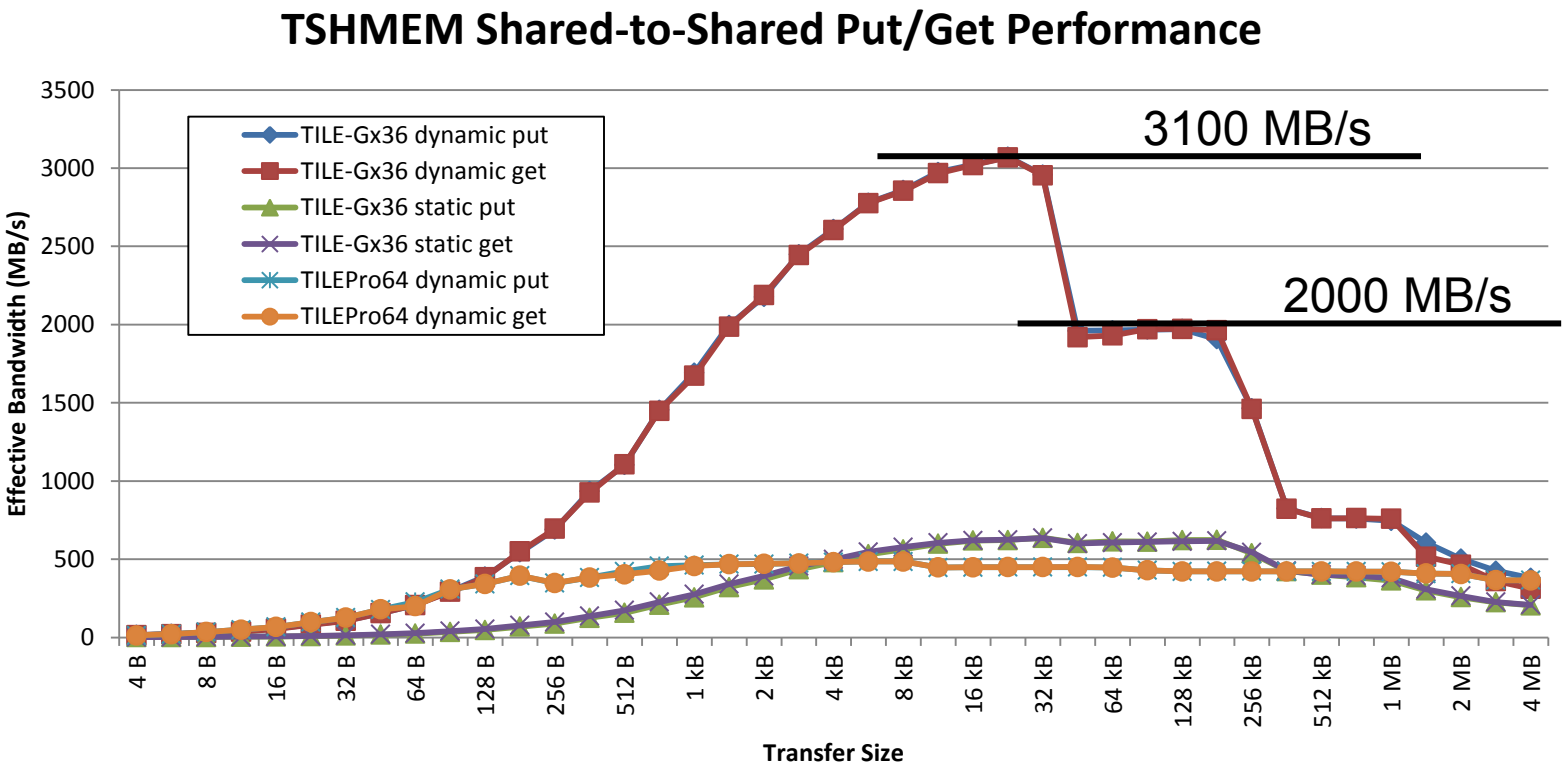
- Bandwidth on iMesh networks to caches and memory controllers
  - Shared memory performance critical for TSHMEM
  - Bandwidth of memory operations influenced by 3 of 5 iMesh networks
    - QDN: memory request network
    - RDN: memory response network
    - SDN: cache sharing network
  - Performance transitions occur at cache-size limits
    - L1 data cache: **3100 MB/s**
    - L2 cache: **2700 MB/s**
    - Tilera DDC L3 cache
    - Memory-to-memory



32k L1i cache
32k L1d cache
256k L2 cache per tile

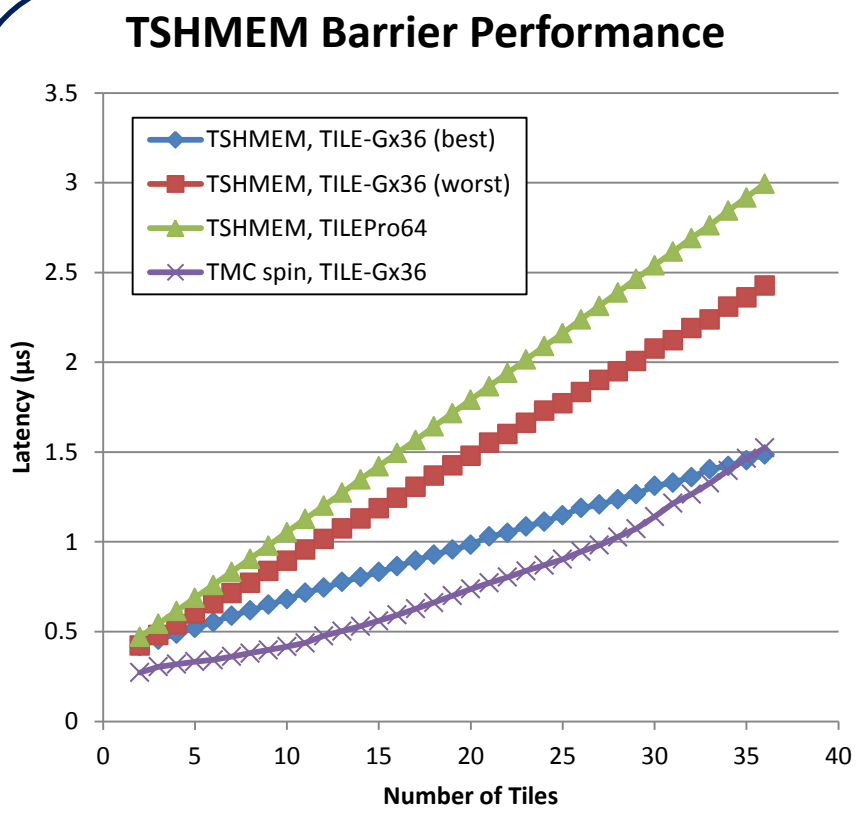


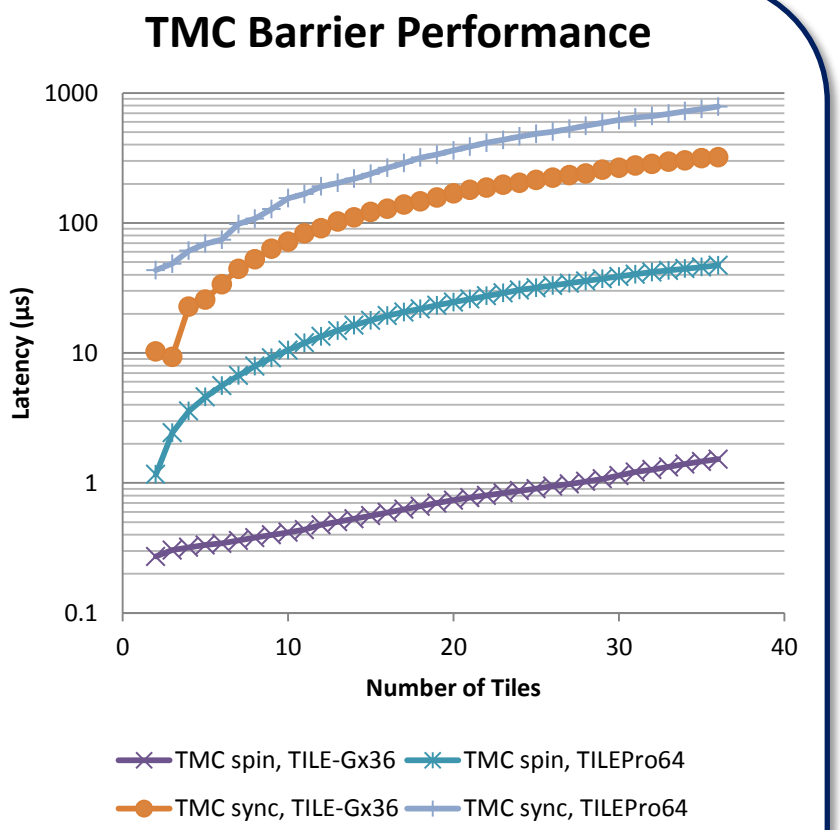
Performance – Put/Get



- TSHMEM dynamic performance ***closely matches shared-to-shared performance*** profile for both TILE-Gx36 and TILE*Pro*64 devices

Performance – Barrier Sync

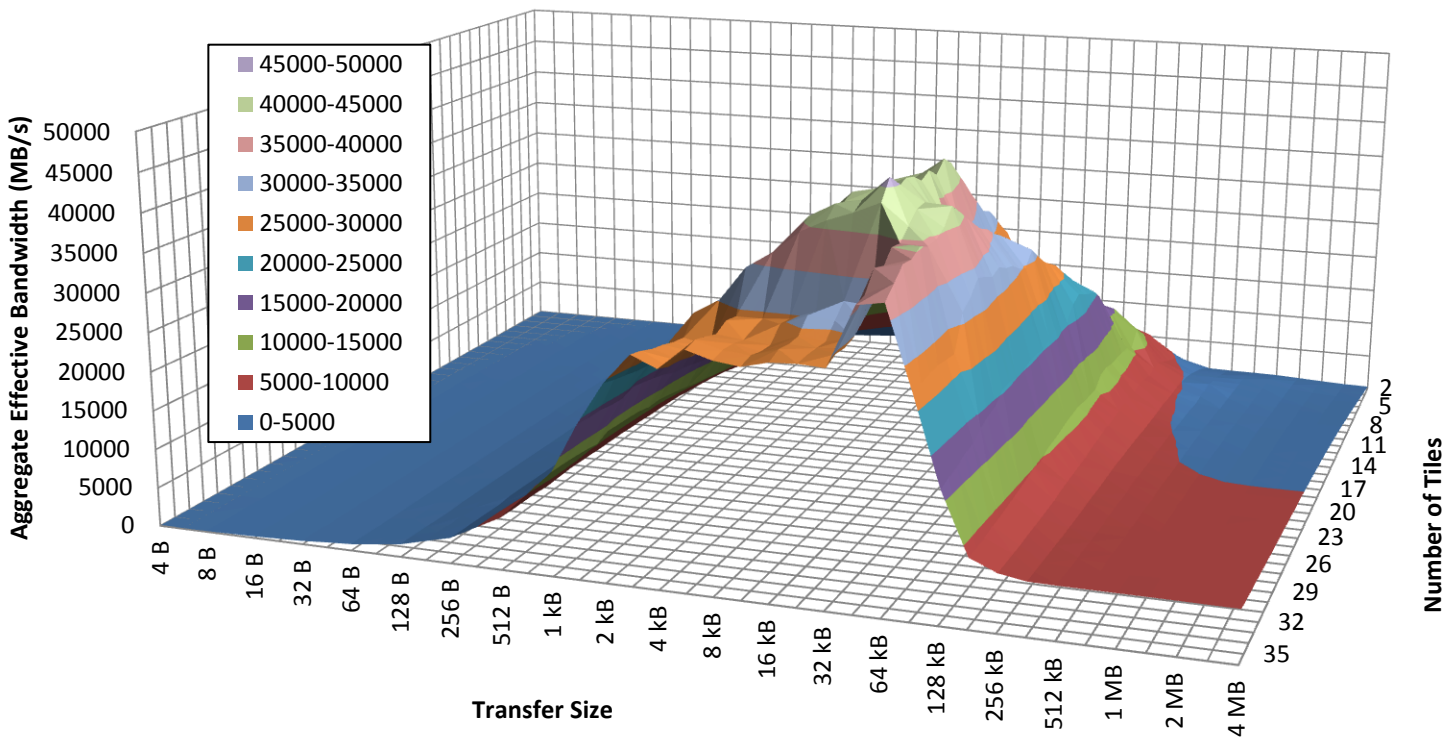




- TSHMEM barriers leverage UDN for ***better scaling*** than most Tilera TMC barriers for TILE-Gx36 and TILE*Pro*64

Pull-based Broadcast

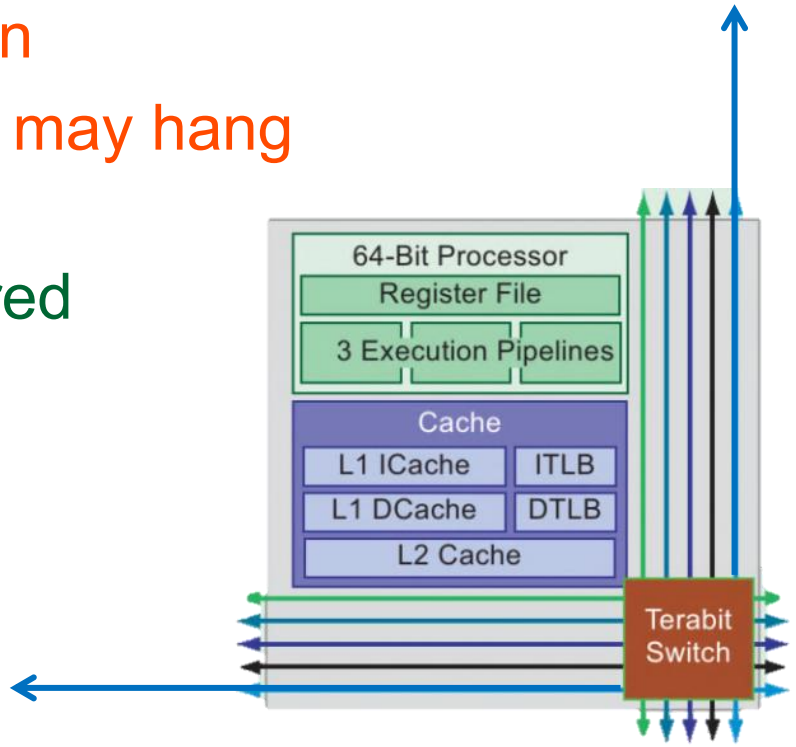
And many more results!



- Single-device broadcast ***up to 45 GB/s aggregate bandwidth*** and 37 GB/s at 36 tiles

OpenSHMEM Extensions

- Tilera's user dynamic network (UDN) needs to be shut down properly
  - TSHMEM uses UDN for barriers and explicit inter-tile communication
  - During termination, processes may hang if UDN is not deactivated
  - *shmem_finalize* support required



Come visit us at SC'12!

- NSF CHREC
  - Booth 2405
- PGAS
  - Booth 2137