

GPU MEMORY HIERARCHY

Local registers

Per threads / Scalar processor

Per blocks / MP

Shared memory

L1 cache









