

# GPU chip

## Multiprocessor

CUDA cores

Warp  
scheduler

Warp  
scheduler

Load-store unit

L1 \$ / shared memory

## Multiprocessor

CUDA cores

Warp  
scheduler

Warp  
scheduler

Load-store unit

L1 \$ / shared memory

## Multiprocessor

CUDA cores

Warp  
scheduler

Warp  
scheduler

Load-store unit

L1 \$ / shared memory

## Interconnection network

L2 cache

Memory controller

L2 cache

Memory controller

L2 cache

Memory controller

L2 cache

Memory controller

DRAM  
chip

DRAM  
chip

DRAM  
chip

DRAM  
chip

DRAM  
chip

DRAM  
chip

DRAM  
chip

DRAM  
chip