

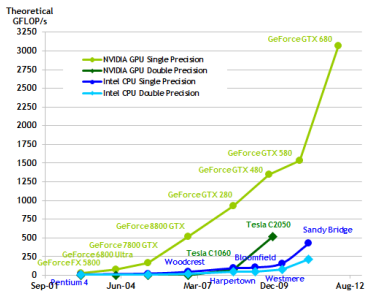
Chapter 4



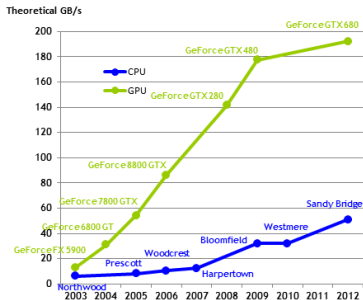
GPU Computing and Accelerators: Part I



Why use accelerators?



(a) Floating point operations



(b) Memory bandwidth

Figure: Throughput comparison of Multicore CPUs and CUDA enabled GPUs (taken from CUDA C Programming Guide)



Why use accelerators?

Architecture	GFLOPS	GFLOPS/Watt	Utilization
Core i7-960	96	1.14	95%
NVIDIA [®] GTX280	410	2.6	66%
Cell	200	5.0	88%
NVIDIA [®] GTX480	940	5.4	70%
TI C66x DSP	74	7.4	57%

Table: Power efficiency comparison of Multicore CPUs and accelerator chips (taken from Conference Poster by F. Igual and M. Ali)



Memory Hierarchy with Accelerators

Common Features

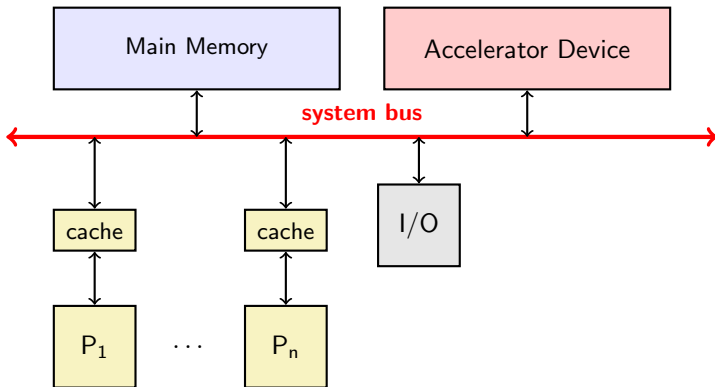


Figure: Schematic of a general parallel system



Memory Hierarchy with Accelerators

Graphics Processing Units (GPUs)

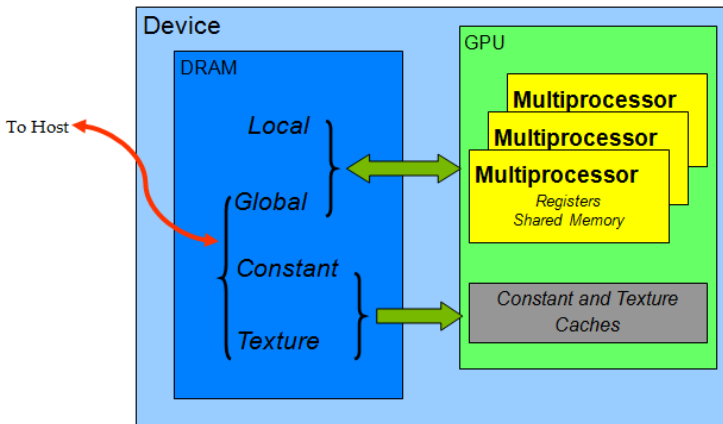


Figure: Memory configuration of a CUDA Device (taken from CUDA C Programming Guide)



Memory Hierarchy with Accelerators

Field Programmable Gate Arrays (FPGAs)

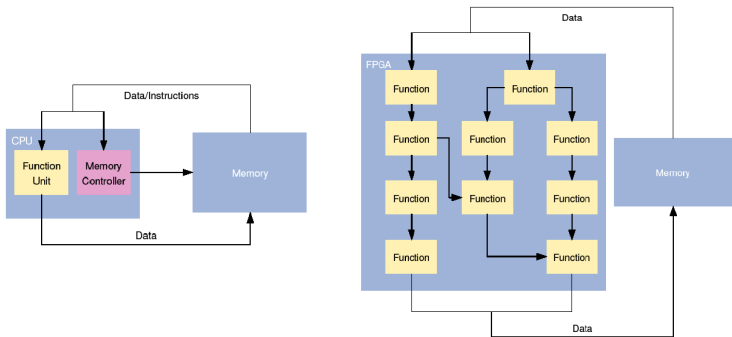


Figure: Comparison of CPUs and FPGA execution models.