

HPCC Mid-Morning Break

GPGPU on HPCC

Dirk Colbry, Ph.D.
Research Specialist
Institute for Cyber-Enabled Research
colbrydi@msu.edu

MICHIGAN STATE
UNIVERSITY



Agenda

- What is GPU programming?
- What do we have at HPCC?
- How do you gain access to the GPU cores?

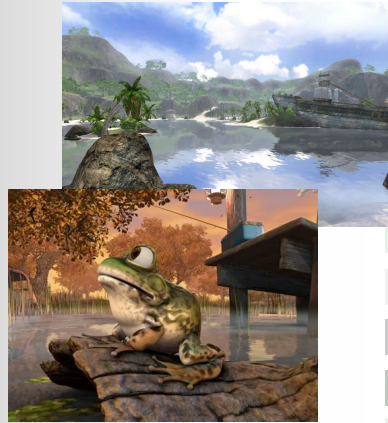


MICHIGAN STATE
UNIVERSITY

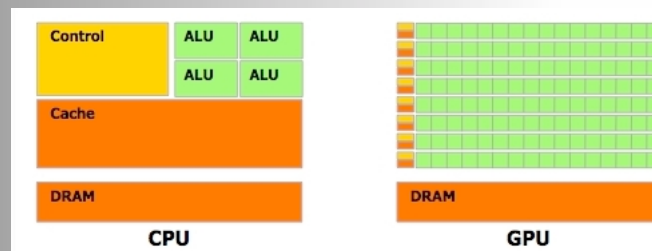


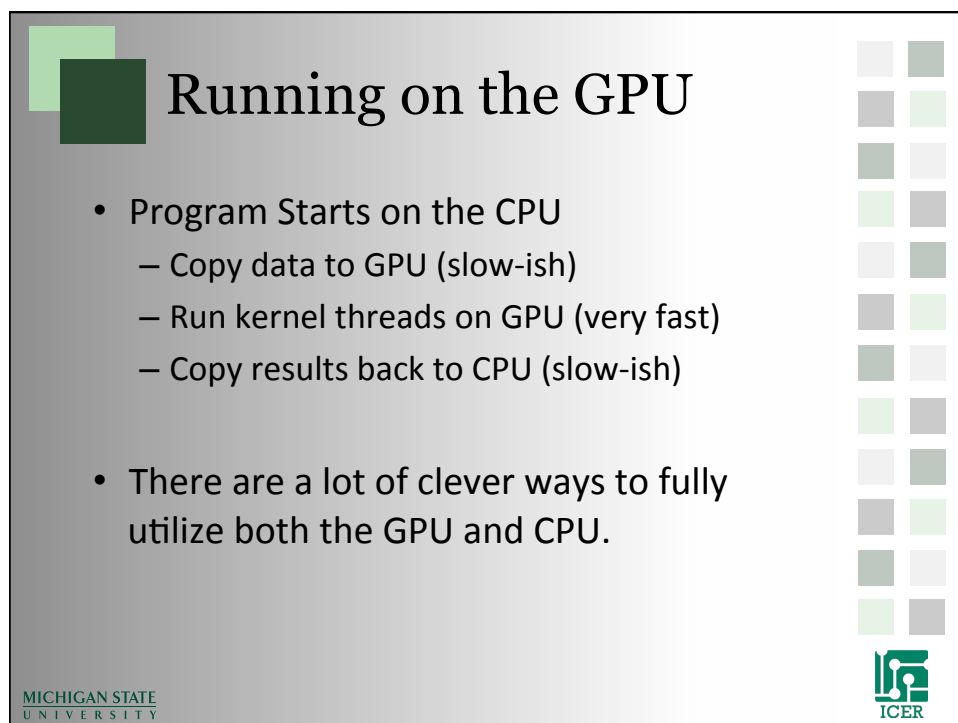
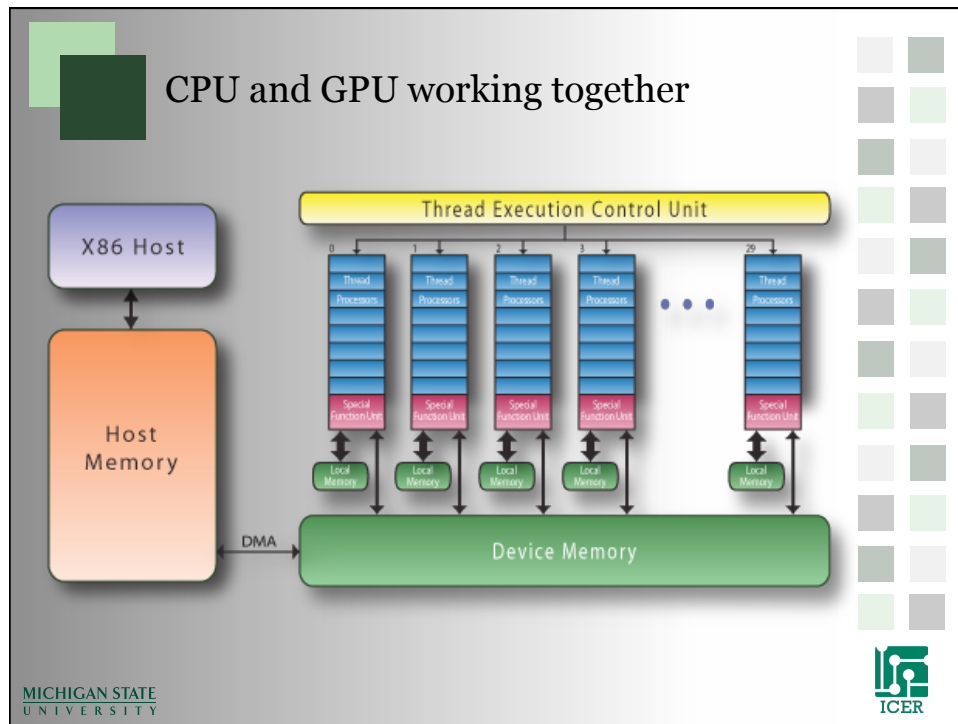
What is a GPU?

- Graphics Processing Unit
- Originally designed to make Video Games
- Uses many processing cores to parallelize the math required for real time game play.
- Early researchers made general programs that looked like graphics so they could run in the GPU.
- In 2006 nVidia released the CUDA programming interface to allow users to easily make scalable general purpose programs that run on the GPU (GPGPU).



GPU vs CPU





Pros and Cons

- Benefits
 - Lots of processing cores.
 - Works with the CPU as a co-processor
 - Very fast local memory bandwidth
 - Large online community of developers
- Drawbacks
 - Can be difficult to program.
 - Memory Transfers between GPU and CPU are costly (time).
 - Cores typically run the same code.
 - Errors are not detected (on older cards)
 - Double precision calculations are slow (On older cards)

HPCC Systems



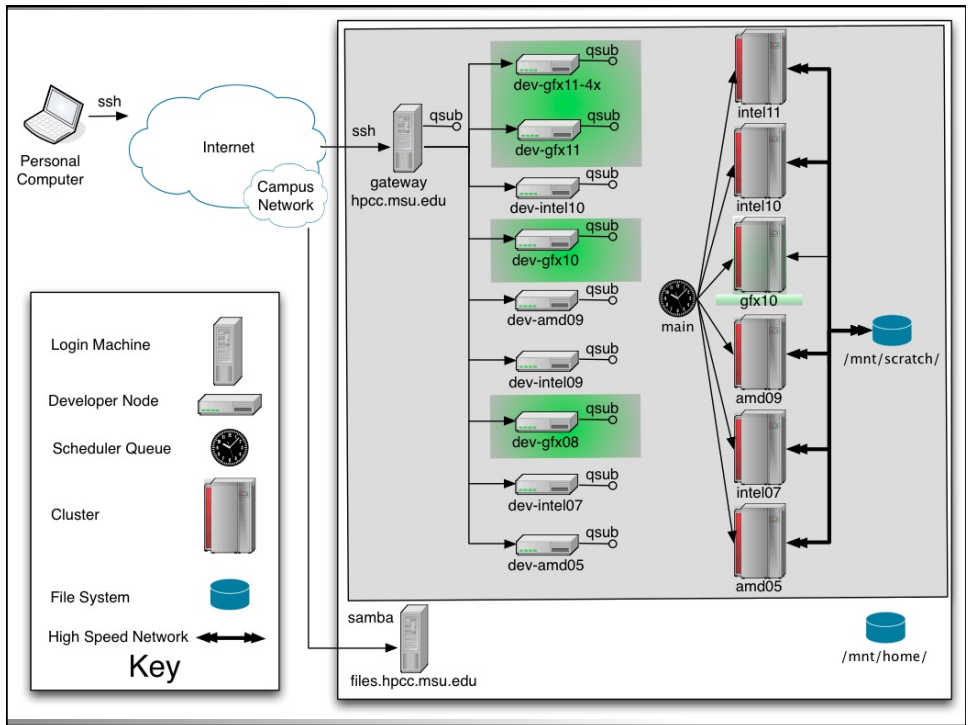
'GFX10', 15,616-core GPGPU cluster



The 'GFX10' cluster consists of 32 compute nodes with two 240-core nVidia Tesla M1060 GPGPU accelerators, eight Intel Xeon cores and 18 GB of RAM per node. The peak aggregate performance of the entire cluster is 59 trillion single-precision floating-point operations per second.

Typical Nvidia Tesla M1060

- Number of Streaming Processor Cores **240**
- Frequency of processor cores **1.3 GHz**
- Single Precision peak floating point performance **933 gigaflops**
- Double Precision peak floating point performance **78 gigaflops**
- Dedicated Memory **4 GB GDDR3**
- Memory Speed **800 MHz**
- Memory Interface **512-bit**
- Memory Bandwidth **102 GB/sec**
- System Interface **PCIe**






GPGPU Developer Nodes

Name	Cores	Memory	GPUs	Notes
dev-gfx08	4	8GB	3	3 Nvidia 480 GTX Cards
dev-gfx10	8	18GB	2	2 Nvidia M1060 Tesla Cards
dev-gfx11	4	8GB	2	1 Nvidia 480 GTX Card
dev-gfx11-4x	8	18GB	4	2 Nvidia C2075 Tesla Cards, 1 Nvidia C2070 Tesla Card and 1 Nvidia 580 GTX card

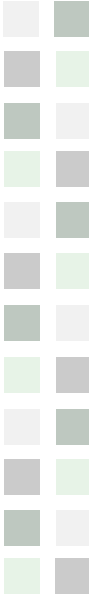
Installed Software on HPCC

- Cuda toolkit
 - For programming in c/c++ and fortran
- PGI compilers with graphics accelerator support.
- cublas – Cuda version of blas libraries
- cufft – Cuda version of fft libraries
- pycuda – Python Cuda Interface
- Openmm/gromacs – Molecular Dynamics Program optimized for GPUs





Other GPGPU Software

- MATLAB
- Mathematica




MICHIGAN STATE UNIVERSITY





Other Available Software

- OpenCL
 - c/c++ interface
- Jacket
 - Matlab GPU wrapper
- Lattice Boltzmann
 - pde solver
- OpenVIDIA
 - Machine Vision
- Many Many others
- Cuda Zone
 - ~90 thousand cuda developers.
 - Lots of software examples
 - Developer Forms
 - Tutorials
- http://www.nvidia.com/object/cuda_home.html
- www.gpucomputing.net



MICHIGAN STATE UNIVERSITY



Running on the GPU cluster

```
#!/bin/bash -login
#PBS -l nodes=1:ppn=1:gfx10,gres=gpu:1,feature=gpgpu
#PBS -l walltime=01:00:00

module load CUDA

mycudaprogram myarguments
```

Getting Examples

```
module load powertools
getexample
getexample cuda
getexample cuda_hybrid
cd cuda
./README
cd ../cuda_hybrid
./README
```




GPGPU Study Group

- Weekly Meetings
 - Wednesdays 4-5pm (1442 BPS)
- Email listserve
 - <http://xsrv.icer.msu.edu/mailman/listinfo/gpgpu-users>
- On-line collaboration
 - <http://wiki.hpcc.msu.edu/display/gpgpu/>
- Voice in iCER GPGPU policy decisions

- Please email me if you are interested in participating and/or willing to help organize the group.