CS/EE 147 – GPU Computing and Programming
Introduction

Abbas Mazloumi
amazl001@ucr.edu

Thanks to Daniel Wong, Marcus Chow
Logistics

TA Office hours
https://doodle.com/poll/n7pdhcys6h6gg99e

My Office Hours
Available by appointment

Create ENGR Account
https://intra.engr.ucr.edu/apps/systems/login.php
• Lab Policies
  • No project for this course to reduce workload
  • Labs are designed to be done individually

• Grade Breakdown
  • Labs: 25%
  • Midterm 1&2: 50%
  • Class Participation: 10%
  • Quizzes 15%
  • Extra Credit: 3%
Important Lab Policies

• 3 slip days
• 15% penalty per late day
• All labs are due at the end of the due date (midnight)
• Projects should be uploaded to iLearn and github classroom

(INCLUDE YOUR GITHUB USERNAME IN YOUR ILEARN REPORT)
Last Session

• We use Discord for discussions and group learning (no piazza anymore)
• All discussion and lecture videos are being uploaded on iLearn/YuJa/All Channels/EE_147_001_21S
• The needed instructions and tutorials will be posted on iLearn.
• Other resources
Breakout Groups

• Introduce yourselves
  • Name, Major, Year, anything you want to share
• Find 5 things in common
• Each group turns in an itemized document to answer the following question as their discussion outcome.
Discussion Today

• Why do you think we teach this class at all? The reason why we use GPUs today
• Think about why you wanted to take this class? Deeper than course requirement
GPU Introduction
42 Years of Microprocessor Trend Data

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp
Modern computer architecture is limited by:

- Process Technology / Transistor density
  (End of Moore’s Law)
- Power
  (End of Dennard Scaling)
- Temperature
General-purpose (Easier to program)

Single-core CPU
Multi-core CPU
GPU
FPGA
ASIC

Specialized
Energy-efficient
Examples of Specialization

Bitcoin Mining
Examples of Specialization

Microsoft Catapult
Examples of Specialization

Google TPU - Tensor Processing Unit
Examples of Specialization

**GPUs - Supercomputers**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>System</th>
<th>Cores</th>
<th>Rmax [TFlop/s]</th>
<th>Rpeak [TFlop/s]</th>
<th>Power [kW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DD5/SC/Ng Ridge National Laboratory, United States</td>
<td>Summit - IBM Power System ACX2, IBM POWER9 2.0GHz, NVIDIA V100 Dual-rail Mellanox EDR Infiniband IBM</td>
<td>2,282,564</td>
<td>122,300.9</td>
<td>187,659.3</td>
<td>8,856</td>
</tr>
<tr>
<td>2</td>
<td>National Supercomputing Center in Wuxi, China</td>
<td>Sunway TaihuLight - Sunway NFEP, Sunway SWI6010 2.0GHz 1.45GHz Infiniband Sunway NRCC</td>
<td>10,649,400</td>
<td>93,014.6</td>
<td>125,435.9</td>
<td>15,371</td>
</tr>
<tr>
<td>3</td>
<td>DD5/NCSA/LLNL, United States</td>
<td>Sierra - IBM Power System SE23L, IBM POWER9 2.2GHz, NVIDIA V100 Dual-rail Mellanox EDR Infiniband IBM</td>
<td>1,572,480</td>
<td>71,616.0</td>
<td>119,193.6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>National Supercomputer Center in Guangzhou, China</td>
<td>Tianhe-2A - TH-IVB-EPF Cluster, Intel Xeon E5-2650V2 2.6Ghz, TH Express-3, Marc-H-2000 NUCT</td>
<td>4,981,740</td>
<td>61,444.5</td>
<td>100,678.7</td>
<td>18,482</td>
</tr>
<tr>
<td>5</td>
<td>National Institute of Advanced Industrial Science and Technology (AIST), Japan</td>
<td>All Bridging Cloud Infrastructure (ABCBI) - PRIMERGY CX2360 M1, Xeon Gold 6148 2.4GHz, NVIDIA Tesla V100-SXM3 Infiniband EDR Fugis</td>
<td>391,480</td>
<td>19,880.0</td>
<td>32,578.4</td>
<td>1,649</td>
</tr>
<tr>
<td>6</td>
<td>Swiss National Supercomputing Centre (CSCS), Switzerland</td>
<td>Piz Daint - CRAY XC33, Xeon E5-2698v4 2.3GHz, Aries Interconnect, NVIDIA Tesla P100 Cray Inc.</td>
<td>561,760</td>
<td>19,590.0</td>
<td>25,323.3</td>
<td>2,272</td>
</tr>
<tr>
<td>7</td>
<td>DD5/SC/Ng Ridge National Laboratory, United States</td>
<td>Titan - CRAY XE6, Opteron 6276 14C 2.2GHz, Cray Gemini Interconnect, NVIDIA K20 Cray Inc.</td>
<td>560,480</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,229</td>
</tr>
<tr>
<td>8</td>
<td>DD5/NCSA/LLNL, United States</td>
<td>Sequoia - BlueGene/Q, Power IBC 16C 1.4GHz, Custom IBM</td>
<td>1,972,844</td>
<td>17,172.2</td>
<td>20,132.7</td>
<td>7,894</td>
</tr>
</tbody>
</table>