Fire and Ice: How Temperature Affects GPU Performance

Danny Price
Harvard-Smithsonian Center for Astrophysics
dprice@cfa.harvard.edu

CfA: Danny Price, Ben Barsdell, Lincoln Greenhill
NVIDIA: Mike Clark, Ron Babich
Introduction

• Watts, leakage & performance
• GPUs in radio astronomy
• Experimental setup
• Results & conclusions
The Green 500

Listed below are the November 2013 The Green500's energy-efficient supercomputers ranked from 1 to 10.

<table>
<thead>
<tr>
<th>Green500 Rank</th>
<th>MFLOPS/W</th>
<th>Site*</th>
<th>Computer*</th>
<th>Total Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4,503.17</td>
<td>GSIC Center, Tokyo Institute of Technology</td>
<td>TSUBAME-KFC - LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x</td>
<td>27.78</td>
</tr>
<tr>
<td>2</td>
<td>3,631.86</td>
<td>Cambridge University</td>
<td>Wilkes - Dell T620 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20</td>
<td>52.62</td>
</tr>
<tr>
<td>3</td>
<td>3,517.84</td>
<td>Center for Computational Sciences, University of Tsukuba</td>
<td>HA-PACS TCA - Cray 3623G4-SM Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x</td>
<td>78.77</td>
</tr>
<tr>
<td>4</td>
<td>3,185.91</td>
<td>Swiss National Supercomputing Centre (CSCS)</td>
<td>Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x</td>
<td>1,753.66</td>
</tr>
<tr>
<td>5</td>
<td>3,130.95</td>
<td>ROMEO HPC Center - Champagne-Ardenne</td>
<td>romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x</td>
<td>81.41</td>
</tr>
<tr>
<td>6</td>
<td>3,068.71</td>
<td>GSIC Center, Tokyo Institute of Technology</td>
<td>TSUBAME 2.5 - Cluster Platform SL390s G7, Xeon X5670 6C 2.930GHz, Infiniband QDR, NVIDIA K20x</td>
<td>922.54</td>
</tr>
<tr>
<td>7</td>
<td>2,702.16</td>
<td>University of Arizona</td>
<td>iDataPlex DX360M4, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR14, NVIDIA K20x</td>
<td>53.62</td>
</tr>
<tr>
<td>8</td>
<td>2,629.10</td>
<td>Max-Planck-Gesellschaft MPI/IPP</td>
<td>iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x</td>
<td>269.94</td>
</tr>
<tr>
<td>9</td>
<td>2,629.10</td>
<td>Financial Institution</td>
<td>iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x</td>
<td>55.62</td>
</tr>
<tr>
<td>10</td>
<td>2,358.69</td>
<td>CSIRO</td>
<td>CSIRO GPU Cluster - Nitro G16 3GPU, Xeon E5-2650 8C 2.000GHz, Infiniband FDR, Nvidia K20m</td>
<td>71.01</td>
</tr>
</tbody>
</table>

Source: [http://www.green500.org/lists/green201311](http://www.green500.org/lists/green201311)
The Green 500

#1

**TSUBAME-KFC**
Tokyo Institute of Technology
Oil-cooled

#2

**Wilkes Cluster**
University of Cambridge
Air-cooled

Source: [http://www.green500.org/lists/green201311](http://www.green500.org/lists/green201311)
The Square Kilometer Array

Image: www.skatelescope.org
Radio Interferometry - lots of computations
Radio Interferometry - lots of computations

- More antennas = lots more baselines

\[ N_B = \frac{N_A(N_A - 1)}{2} \]

For example, \( N_A = 512 \Rightarrow N_B = 130816 \)

- Computations done on a frequency channel basis

- Trivially parallelizable: over frequency and baseline
xGPU - a GPU correlator for radio astronomy

xGPU is a GPU-based cross correlator for radio astronomy.

It powers LEDA (the world’s largest Nant radio telescope), PAPER, and the MWA (among others).


Authors:
Michael Clark (NVIDIA)
Paul La Plante (Loyola University Maryland)
Lincoln Greenhill (Harvard-Smithsonian Center for Astrophysics)
David MacMahon (University of California, Berkeley)
Ben Barsdell (Harvard-Smithsonian Center for Astrophysics)
Watts, leakage & performance

\[ P = P_{\text{dynamic}} + P_{\text{static}} \]

Leakage power (bad)
Watts, leakage & performance

\[ P_{dynamic} = C f V^2 \]

- Switching frequency
- Transistor capacitance
- Voltage
Leakage mechanisms:
- Reverse-biases junction leakage
- Gate-induced drain leakage
- Gate direct-tunneling leakage
- Subthreshold (weak inversion) leakage
Leakage mechanisms:
- Reverse-biases junction leakage
- Gate-induced drain leakage
- Gate direct-tunneling leakage
- Subthreshold (weak inversion) leakage
How temperature affects performance

\[ I_{\text{subthreshold}} = A_s \frac{W}{L} \left( \frac{kT}{q} \right)^2 e^{\frac{q(V_S - V_{th})}{nkT}} \]
How temperature affects performance

\[ I_{\text{subthreshold}} = A_s \frac{W}{L} \left( \frac{kT}{q} \right)^2 e^{\frac{q(V_S - V_{th})}{nKT}} \]

Thermal voltage
8.62 × 10^{-5} eV/K
26 mV at room temperature
How temperature affects performance

Device specifics:
- A: Technology specific constant
- L, W: Device channel length & width

\[ I_{\text{subthreshold}} = A \frac{W}{L} \left( \frac{kT}{q} \right)^2 e^{\frac{q(V_S - V_{th})}{nkT}} \]

Thermal voltage
- 8.62 \times 10^{-5} \text{ eV/K}
- 26 mV at room temperature
How temperature affects performance

Device specifics:
- \( A \): Technology specific constant
- \( L, W \): device channel length & width

Voltage terms:
- \( V_s \): Gate to source voltage
- \( V_{th} \): Switching threshold voltage
- \( n \): transistor subthreshold swing coeff

\[
I_{\text{subthreshold}} = A_s \frac{W}{L} \left( \frac{kT}{q} \right)^2 e^{\frac{q(V_s - V_{th})}{nkT}}
\]

Thermal voltage
- \(8.62 \times 10^{-5} \text{ eV/K}\)
- \(26 \text{ mV at room temperature}\)
Results Part 0:

Green GPU: A water-cooled test rig
Green GPU: A water-cooled test rig

- GreenGPU specifications:
  - Gigabyte GA-Z68MX motherboard
  - Intel Core i7-2600
  - 16GB RAM
  - NVIDIA Tesla K20m
  - EK-FCTK20 water block
Warranty void: K20m from a HP server
Warranty void: K20m from a HP server

GK110 die image: news.softpedia.com
Green GPU
Green GPU
Green GPU

Tuna Can
The experiment

• Run xGPU in benchmark mode
• Run SMI tool in loop
• Monitor GPU die temp and power
NVIDIA System Management Interface

```
<table>
<thead>
<tr>
<th>Fan</th>
<th>Temp</th>
<th>Perf</th>
<th>Pwr:Usage/Cap</th>
<th>Memory-Usage</th>
<th>GPU-Util</th>
<th>Compute M.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Tesla K20m</td>
<td>Off</td>
<td>0000:01:00.0 Off</td>
<td>340MB / 479MB</td>
<td>99%</td>
</tr>
<tr>
<td>N/A</td>
<td>31C</td>
<td>P0</td>
<td>151W / 225W</td>
<td>0000:01:00.0 Off</td>
<td>340MB / 479MB</td>
<td>99%</td>
</tr>
</tbody>
</table>

Compute processes:
- GPU
- PID
- Process name
  - 0 9159 ./cuda_correlator

Thu Feb 13 15:23:05 2014

NVIDIA-SMI 5.319.37 Driver Version: 319.37
GPU Name Persistence-M Bus-ID Disp.A Volatile Uncorr. ECC
- 0 Tesla K20m Off 0000:01:00.0 Off 0 9159 ./cuda_correlator

Thu Feb 13 00:16 2014

NVIDIA-SMI 5.319.37 Driver Version: 319.37
GPU Name Persistence-M Bus-ID Disp.A Volatile Uncorr. ECC
- 0 Tesla K20m Off 0000:01:00.0 Off 0 9159 ./cuda_correlator

Thu Feb 13 15:23:05 2014

NVIDIA-SMI 5.319.37 Driver Version: 319.37
GPU Name Persistence-M Bus-ID Disp.A Volatile Uncorr. ECC
- 0 Tesla K20m Off 0000:01:00.0 Off 0 9159 ./cuda_correlator
```

GPU Memory Usage
- 0 9159 ./cuda_correlator 325MB

GPU Util Compute M.
- 99% | Default

```
NVIDIA System Management Interface
NVIDIA SMI — GPU Boost on K20m

• Can control “boost clock” through nvidia-smi:
  
  nvidia-smi –q –d SUPPORTED_CLOCKS
  nvidia-smi –ac <MEM clock, Graphics clock>

• Changing boost clock changes voltage levels (more on this later)

  • **Undervolting** uses less power, but becomes unstable at higher temperatures

  • **Overvolting** uses more power, but is more stable at higher temperatures. There is a big “overclocker” community that exploits this.

• **V4** is the K20m default. Note max power (TDP) is 225W.
How to *really* void your warranty

Don’t try this at home

GPU-Z
(www.techpowerup.com/gpuz/)

Kepler Bios Tweaker v1.27
(overclock.net forums)
Overclocking and xGPU performance

![Graph showing the relationship between clock frequency and xGPU GFLOPS/s for CUDA 5.5 and CUDA 6.0.]
Drifting temperature, drifting power

- Turned off water pump
- Turned on water pump
Results: clock & voltage control
Measured power vs temperature

$V = V_4$
Measured power vs temperature

Power (W)

Temperature (C)

V = V4
Measured power vs temperature

Power (W) vs Temperature (°C) for different frequencies:
- 705MHz
- 805MHz
- 905MHz

Arrows indicating a change of 14W.
Measured power vs temperature

\[ P = P_{\text{dynamic}} + P_{\text{static}} \]

-14W
+14W

V = V4
How temperature affects performance

\[ V = V_4 \]
How voltage affects performance

![Graph showing the relationship between temperature and performance for different voltage levels. The x-axis represents temperature in Celsius (C), and the y-axis represents GFLOPS per Watt (GFLOPs/W). The graph includes data for v1, v2, v3, v4, and v5, with v1 having the highest performance at 800 MHz.]
Undervolted, different clock frequencies

$V = V_1$

![Graph showing the relationship between temperature (C) and GFLOPs/W for different clock frequencies. The graph includes lines for 614 MHz, 714 MHz, 814 MHz, and 914 MHz.]
Heavily undervolted

- Data!
- Invalid

Graph showing performance in GFLOPs/W vs. Temperature (C) for two configurations:
- Red dots: v=875mV, f=705MHz
- Blue dots: v=875mV, f=920MHz

The shaded region indicates invalid data.
Green GPU and the Test Equity 1000 Chamber of Doom

Temperature range:
-100F to 350F
(-73C to 175C)
Green GPU and the Test Equity 1000 Chamber of Doom
or
Thermal shock: how to really really void your warranty

Temperature range:
-100F to 350F
(-73C to 175C)
Green GPU and the Test Equity 1000 Chamber of Doom

or

Thermal shock: how to really really void your warranty

Temperature range: -100F to 350F (-73C to 175C)
Overclocked & Undervolted

\[ P = 0.18T + 157.95 \]

\[ P = 24.0 \left( \frac{k}{\theta} \right) T^2 + 160.50 \]

926 MHz clock
Voltage v1
Overclocked & Undervolted

926 MHz clock
voltage v1

[Graph showing the relationship between temperature and GFLOPs/W, with a trend line indicating a decrease in GFLOPs/W as temperature increases.]
Part III: Comparisons & Conclusions
Kepler vs. Fermi

Tesla K20m
GK110, 28nm

Geforce GTX 580
GF110, 40nm

$P = 243 + 0.148T + 0.00397T^2$

$P = 235 + 0.548T$

926 MHz clock
t voltage v1
Maxwell results (preliminary)

GTX 750 Ti
- 640 CUDA Cores
- 1020 MHz base Clock (MHz)
- 1085 MHz boost Clock (MHz)
- 60W TDP

xGPU
- 1062 GFLOP/s performance
- 76% efficiency of code
- expect >80% efficiency with code optimization
- already >16 GFLOPS/W
- can’t control clock / voltage (yet…)

GTC 2014 | MARCH 24-27 | SAN JOSE, CA
Conclusions

• 12.96 —> 18.34 GFLOPs/W, i.e. **41% increase** by overclock + undervolt on K20m for xGPU.

• Maxwell is >16 GFLOPs/W for xGPU out of the box!
Conclusions

- xGPU: a useful CUDA benchmark and “burn-in” tool
- Lower temperature = lower power consumption
- Overclocking w/o overvolting = good
- Overclocking + unovervolting = better
- Tesla K20m is pretty hard to brick
Acknowledgements

• Robert Kimberk (CfA lab tech)
• Overclocking community
• NVIDIA

CfA: Danny Price, Ben Barsdell, Lincoln Greenhill
NVIDIA: Mike Clark, Ron Babich
dprice@cfa.harvard.edu