Using the Roofline Model and Intel Advisor

Date: August 16, 2017
Presented by: Sam Williams and Tuomas Koskela

The presentation material along with the video will be made available in two places:
https://exascaleproject.org/event/using-the-roofline-model/
https://ideas-productivity.org/events/

You may use this Google Doc to ask questions as we go along and one of our presenters will either answer here or ask the presenter at certain intervals during the presentation.

A: 0.166 = 2 flops/24 bytes

But the Flop per Byte doesn't change does it? Which number changes?

Q: Slide 11, are the two the same kernels? How do you get different AI for same kernels for different memories?

A:

Q: Must the roofline model be run utilizing all cores on a node or can it be used on one core? Because the bandwidth and peak FLOP/s differs on a serial program over a parallel one

Q: On slide 12, what is DLP?

A: “Data Level Parallelism” and TLP is “Thread Level Parallelism”

Q: Why does the green line (mcdram) have a y-offset for AI=0? Slide #16

A: When we do a real example, you’ll see that it’s usually presented a log-log scale. Slide 16 for example.

Q: On slide 19, what cache model(s) were used for the byte denominator for AI?

Q: KNL has no FLOP hardware counter, so how did you attain flop counts on that? Related to that, how were the flops measured on the Haswell?

A: We use the Intel sde tool, that is a software counter that counts the FLOPS from the instructions issued by the program

Q: Can you reiterate how the 0,44 AI was obtained?

Q: Slide 20: Did you conclude that all of the apps are memory bound?

Q: How does spatial locality effect AI?

Q: Is there a tool to calculate DRAM AI or L1/L2 AI for a complex application code (not small kernels)?

Q: Is the roofline plot you showed for the whole application or the loop you have highlighted? Related to this question - can we annotate a specific region of code to be profiled instead of a whole application. Similar to vtune macros?

Q: Can these tools be used for python/cython-cased code?

Q: What are the differences between the stencil loops?

Q: My code is written in Fortran + C++. Can I use Advisor with it?

Q: In Roofline Advisor, how were the flops measured on the Haswell?

A: SDE was used to calculate flops on Haswell I think

Q: Can we get the codes for stencil, ver0-4?