Computer Systems Architecture
Assignment 5
Due: Tuesday, May 20

I'm somewhat nervous about assigning this particular problem. It uses Blis. Nonetheless, let's bravely forge ahead.

The assignment is meant to develop your skills as a L1 data cache designer. Be aware, of course, that your assessment of L1 data cache performance is biased by the fact that you are basing it on an evaluation of a single program. (Evan will send out email telling you which benchmark to use for this assignment.)

For this assignment your should work in teams. If possible, I'd like you to pair up with someone you haven't worked with before. But I realize that the possibilities of this are getting slimmer and slimmer.

Part I: Designing the cache

Your cache holds 64KB of data. Using your knowledge of the benefits of cache parameters and your friend Blis, design a configuration (cache size, block size, associativity) that produces the best performance for that amount of data. Don't carry this to extremes, meaning don't simulate every conceivable configuration under the sun; choose the ones you simulate thoughtfully (for example, full associativity might be a nice parameter value to omit). Choose some good alternatives, and experiment with them. Explain in the report why you chose the ones you did.

Discuss how your L1 data cache performs. Does Blis already generate the metrics to enable you to determine how well your cache is functioning? If not, think about what metrics you should add to gauge cache performance and implement them.

Part II: Improving upon the design

Improve on the cache design *without* increasing the total size of the cache or changing the other two parameters by using any of the techniques we have discussed in class. Justify your choice of techniques in light of the performance (or lack of performance) of the machine executing without it (this is an important aspect of this assignment). Discuss the pros and cons of implementing the technique and its effect on all aspects of performance.

When adding a technique, you can either create a new component or add code to the cache component – the latter is likely the path of least resistance.

Report:

Your report should follow the guidelines for the last report, and it can also be short -- only 2-3 pages of text should do it (your figures and tables can be extra). Pay attention to your design hypotheses and data analysis.