

For the following assignment read the description of the GeForce 3 in Erik Lindholm, Mark J. Kilgard, Henry Moreton, "A User-Programmable Vertex Engine," *SIGGRAPH 2001*, p.149-158 and also read the description of the GeForce 6800 in John Montrym and Henry Moreton, "The GeForce 6800," *IEEE Micro Magazine*, vol. 25, no. 2, March 2005, pp. 41-51. Both papers are linked to the course references page, <http://www.ece.lsu.edu/gp/ref> and require a password if accessed from outside `lsu.edu`.

**Problem 1:** According Montrym 2005 the GeForce 6800 was designed specifically so that a particular resource would likely be the bottleneck. (They don't use the word bottleneck.)

Ideally there would be no bottlenecks, but there is no way a particular GPU design could have no bottlenecks for all of the different code that might run on it. This resource we are talking about is an expensive one so the designers don't want it idle unless there is nothing to do.

(a) What is the resource?

(b) Suppose for some candidate design and GPU code that resource is not the bottleneck. How might the design make the resource the bottleneck by *adding* to other parts of the design. (Be reasonably specific.)

**Problem 2:** Compare the vertex processor design in the GeForce 3 and the GeForce 6800.

(a) Describe two interesting similarities and two interesting differences between them.

(b) For each difference explain why the design changed. Try to be reasonably specific.

**Problem 3:** The 6800 has six vertex processors, each operating independently, as an MIMD (multiple instruction, multiple data) group. An alternative would be to operate the six vertex processors as a SIMD (single instruction, multiple data) group, that is, a single PC would be used for all six VPs. (This has nothing to do with multithreading).

There are certain additions to the instruction set of the 6800 that would not be nearly as useful if the six VPs operated as a SIMD group.

(a) What are those extensions?

(b) Why would SIMD make it difficult?