

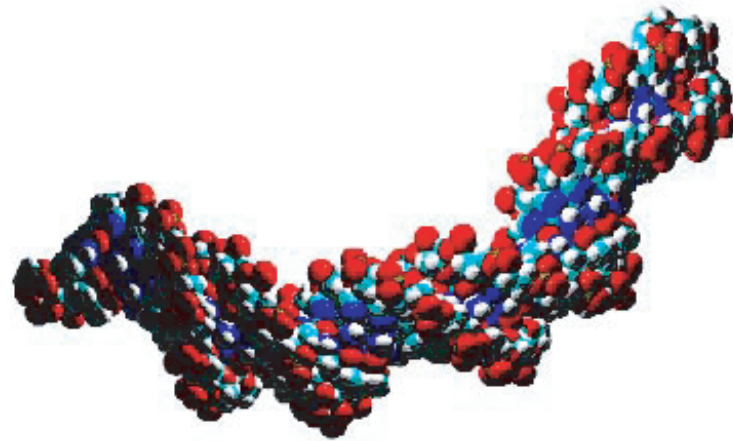
Caltech
Department of
Computer
Science

CS 101.3
Hacking the GPU
Class Lecture

Monday,
November 25th,
12:30 - 2pm
Lauritsen 123

Pat Hanrahan

Department of
Computer Sciences
Stanford University



Streaming Scientific Computing on GPUs

Graphics processors are very fast and powerful. A modern GPU has 8 parallel programmable pipelines running 4-way vector floating point instructions and achieves over 50 GFLOPs. Graphics processors use VLSI resources efficiently because they use data parallelism and carefully orchestrate communication between processors. For these reasons, over the last two decades, graphics processor performance has been increasing significantly faster than microprocessors, with an annual increase in performance rate of over 2.4.

A natural question to ask is whether other important algorithms can take advantage of the techniques used to achieve high performance in graphics processors. The Stanford Streaming Supercomputer project, co-lead with Bill Dally, is investigating this approach. In this talk I will describe our approach to the problem. First, we abstract the processor as a stream processor. Second, we have developed a high-level programming environment for stream processing called Brook. Brook will run a variety of architectures, including next generation GPUs. Finally, we have evaluated three applications: molecular dynamics, Galerkin finite element codes on triangular meshes, and a multigrid-based fluid flow solver. These algorithms all seem to map well to stream processors. Our conclusion is that it is possible to build much more cost-effective supercomputers for scientific calculations.

