Brook for GPUs

Ian Buck, Tim Foley, Daniel Horn, Jeremy Sugerman, Kayvon Fatahalian, Mike Houston, and Pat Hanrahan

Motivation

- General purpose computing
- GPU = general stream coprocessor
- Friendly programming environment
- Programmer need not know GL
- Virtualize or abstract GPU resources
- Cross platform
  - ATI & NVIDIA
  - Linux and Windows
  - DX and OpenGL

In the News

- Open source
- Over 6,300 downloads in 8 months
- 163K page hits
- GPGPU SIGGRAPH Course
- In the news:
  - IEEE Computer
  - http://slashdot.org
  - http://opengl.org
  - http://gpgpu.org

Brook Language

- Streams
  - collection of records requiring similar computation
  - particle positions, voxels, FEM cell, ...
  - Ray r<200>
  - float velocityfield<100,100,100>
  - similar to arrays, but...
  - index operations disallowed: position[i]
  - read/write stream operators
  - streamRead (r, v_ptr);
  - streamWrite (velocityfield, v_ptr);

- Kernels
  - functions applied to streams
  - similar to for_all construct
  - kernel void foo (float a<>, float b<>, float c<>, out float result<>) {
    float array[], iter float n<>, c<100>, b<100>, a<100>
    for (int i=0; i<100; i++)
      c[i] = a[i] + b[i];
  }  

Gather Streams

- Indirect addressing of stream data allowed inside kernels.
- Argument passed with array syntax
  - float array[]

- Maps to depended texture lookup
  - velocityfield<100,100,100>
  - Ray r<200>

Iterator Streams

- Special stream type preinitialized with sequential values (1,2,3,4,...).
- Maps to depended texture lookup
  - _tex_pos, _tex_pos_pos

Running Brook

- BRCC Compiler
  - Based on ctool
  - Leverage vendor shader compilers
    - Microsoft: fx
    - NVIDIA: cgc
  - Converting kernels into shaders
    - stream fetch and store
    - gather operations
    - register mapping
    - stub function

Virtualization

- Stream size and dimensions
  - packing streams into 2D segmented memory space
  - compiler inserts address translation code
    - float matrix<8096,10,30,5>;

Applications

- ray-tracer
- segmentation
- 2D convolutions
- linear algebra
- fft edge detect

GPU Results

- compared against:
  - Intel Math Kernel Library
  - Atlas Math Library
  - cached blocked segmentation
  - FFTW
  - Wald SSE Ray-Triangle code