GPGPU, 6th Meeting

Mordechai Butrashvily, CEO

moti@gass-ltd.co.il

GASS Company for Advanced Supercomputing Solutions

Agenda

- 5th meeting
- 6th meeting
- Future meetings
- Activities

5th meeting

- Presenting HMPP by CAPS
- Compiler tool and framework to use with existing C/FORTRAN code and convert to the GPU

6th meeting

- jCUDA – library for accelerating Java applications
- Presenting various library functionalities
- Examples
- Questions

Future meetings

- Building a GPU cluster – a recipe
- Hoopoe – GPU cloud solution and architecture
- OpenCL standard
- More advanced topics
- Looking for ideas ☺

Activities

- Basis for a platform to exchange knowledge, ideas and information
- Cooperation and collaborations between parties in the Israeli industry
- Representing parties against commercial and international companies
- Training, courses and meetings with leading companies

jCUDA
Java library for CUDA

Agenda

- GPU computing with CUDA
- jCUDA – library overview
- Native interface model
- Features and API
- Examples:
  - Vector add
  - Image processing

GPGPU with CUDA

- Using GPU for computations
- CUDA API by NVIDIA
- Works on NVIDIA hardware products
- C language for arithmetic “kernels”
- Works under Windows/Linux/MacOS/Solaris
- Can achieve x10-x1000 performance

jCUDA – library overview

- Motivation – creating a Java interface to the GPU
- Java interface, based on CUDA API and semantics
- Uses native interface (JNI) to communicate with the drivers and hardware
- Works under Linux/Windows, upcoming support for Solaris (?)

Native interface model

- Using JNI
- Split to 2 DLL's / SO's:
  - jcuda – Access to CUDA driver API (CUDA & OpenGL)
  - jcudafft – Access to the CUFFT driver
- Provides a standalone access to CUDA functionality, transparent to the user

Features

- API capabilities:
  - CUDA driver API
  - CUFFT routines (for FFT)
  - OpenGL interoperability
  - Object oriented interface to work with the GPU
- Very light-weight

Features (Cont.)

- Built for performance and real-time applications
- Can be used by server side applications for general computing

CUDA Driver API

- Fully compliant with CUDA 2.1 (not including JIT) and driver interface
- Low level interface to the hardware
- Exposed through *CUDADriver* class
- Should not be used directly by the user
- Allocating memory: *CUDADriver.cuMemAlloc*

CUFFT API

- Provides access to FFT routines implemented on the GPU
- Can be accessed through CUFFTDriver
- Can be used with 1D, 2D and 3D FFT
- Provides access to the driver API
- Low level and should not be accessed directly
- Example for real to complex: *CUFFTDriver.cufftExecR2C*

OpenGL interoperability

- Low level API
- Exposed through OpenGLDriver
- Can be accessed through OpenGLDriver
- Implemented by the CUDA driver
- Allows to integrate CUDA with OpenGL graphics transformations:
  - Replacing fragment shaders
  - Replacing vertex shaders

Object oriented API

- Advanced, simpler access to CUDA API
- Implemented for CUDA, CUFFT and OpenGL driver API
- Can be accessed through classes:
  - *CUDA*
  - *CUFFT*
  - *OpenGL*

OO Example

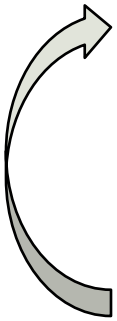

```
CUDA cuda = new CUDA(true);
int devices = cuda.getDeviceCount();
```

Examples

- Vector add
- Image processing:
  - Erode
  - Color convert – RGB → YCbCr

Vector add

- Simple addition between two float vectors
- 256 elements in each
- The result is stored in the first buffer and copied back to host



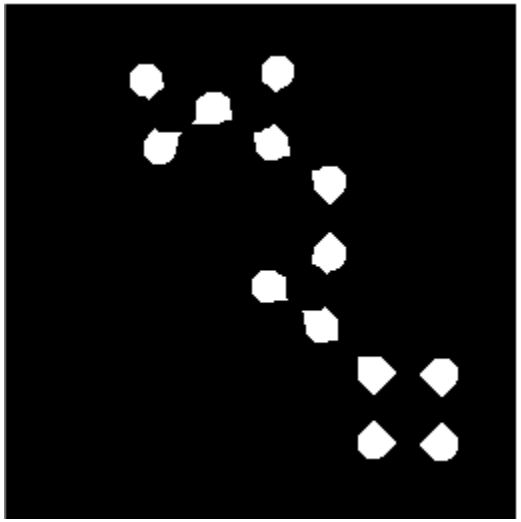
1	0.5	7.8	9.9	10.5	11.2	13.8
+						
2	3	8.9	-5	-0.2	0.5	102.1
=						
3	3.5	16.7	4.9	10.3	11.7	115.9

Erode

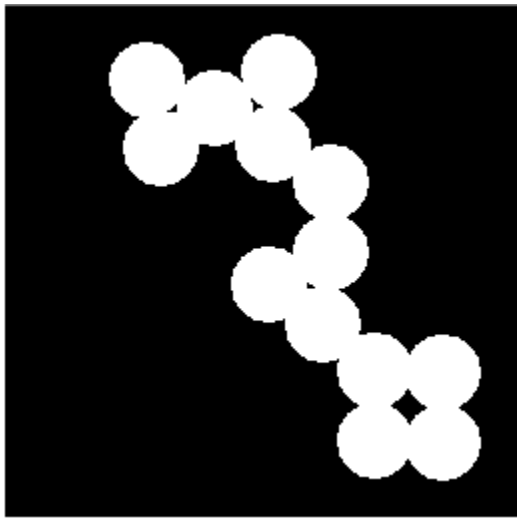
- Image operator to select the minimum value in a region around a pixel
- Used to narrow lines or decrease boldness
- Image size 16x16, grayscale, 32 bit per pixel

Erode example

Before



After



Color conversion

- Converting an RGB image to YCbCr (YUV) color space
- Very useful for video processing or encoding/decoding
- Used by MPEG 4 formats (VC-1 etc.)
- Image features: 16x16, 96 bit (32 bit per channel)

RGB → YUV Example

Summary

- It is possible to accelerate Java applications with jCUDA
- Provided API is simple and efficient
- Real-time performance is here!

Questions