Parallel Computing with GPUs

Parallel Patterns
Part 1 – Parallel Patterns Overview

The University Of Sheffield.

Dr Paul Richmond

http://paulrichmond.shef.ac.uk/teaching/COM4521/

This Lecture (learning objectives)

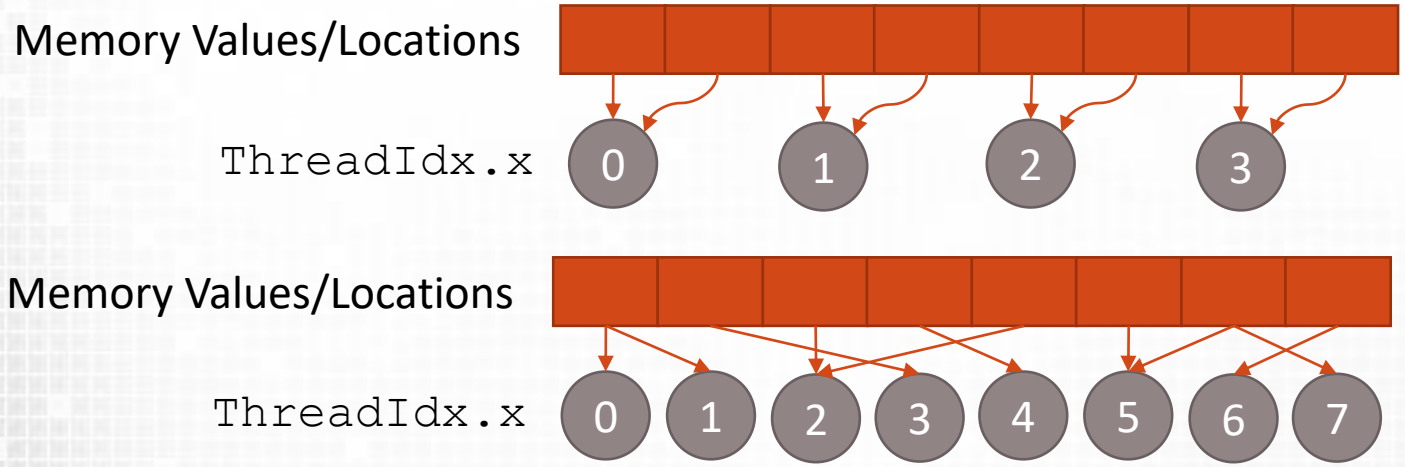
- Parallel Patterns
  - Define a parallel pattern as building blocks for parallel applications
  - Give examples of common patterns

What are parallel Patterns

- Parallel patterns are high level building blocks that can be used to create algorithms
- Implementation is abstracted to give a higher level view
- Patterns describe techniques suited to parallelism
  - Allows algorithms to be built with parallelism from ground up
  - Top down approach might not parallelise very easily…
- Consider a the simplest parallel pattern: *Map*
  - Takes the input list *i*
  - Applies a function *f*
  - Writes the result list *o* by applying *f* to all members of *i*
  - Equivalent to a CUDA kernel where *i* and *o* are memory locations determined by `threadIdx` etc.

Gather

- Multiple inputs and single coalesced output
- Might have sequential loading or random access
  - Affect memory performance
- Differs to map due to multiple inputs

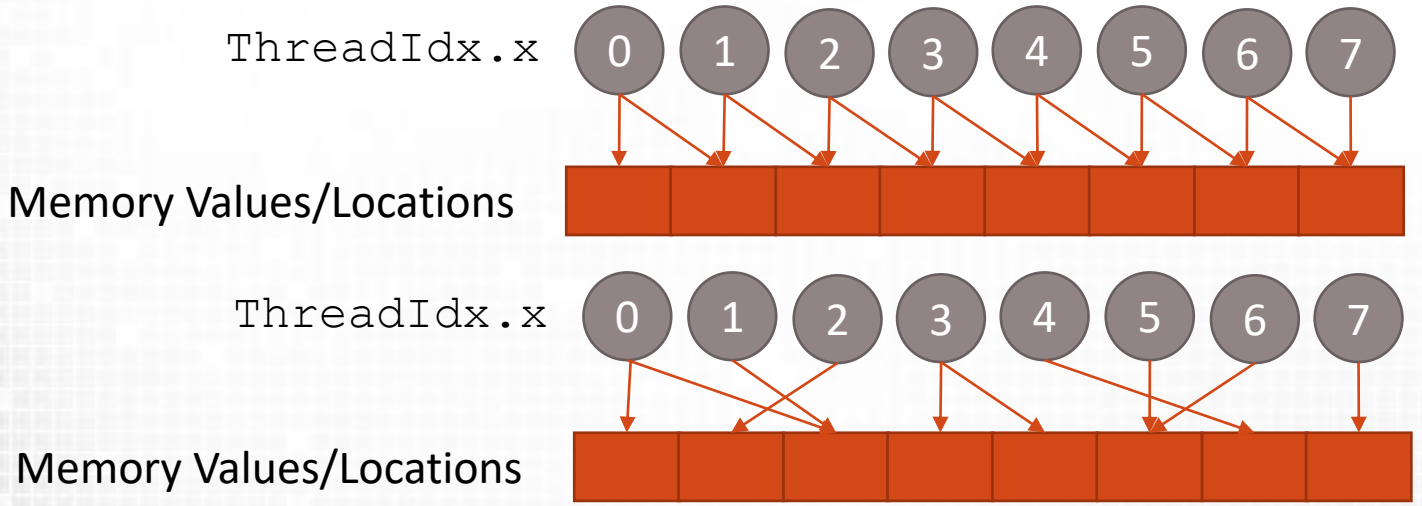


Gather operation
- Read from a number of locations

Gather operation
- Read from a number of locations
- Random access load

Scatter

- Reads from a single input and writes to one or many
- Can be implemented in CUDA using atomics or block/warp co-operation
- Write pattern will determine performance



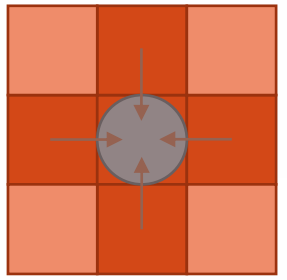
Scatter operation
- Write to a number of locations
- Collision on write

Scatter operation
- Write to a number of locations
- Random access write?

The University Of Sheffield.

Other Parallel Patterns

- Stencil
  - Gather a fixed pattern, usually based on locality
  - See 2D shared memory examples



Stencil Gather

- **Reduce**
  - Reduce value to a single value or set of key value pairs
  - Combined with Map to form Map Reduce (often with intermediate shuffle or sort)
- **Scan**
  - Compute the sum of previous value in a set
- Sort (*later*)
  - Sort values or <value, key> pairs

Summary

- Parallel Patterns
  - Define a parallel pattern as building blocks for parallel applications
  - Give examples of common patterns

- Next Lecture: Reduction