Minutes of Honours Progress Meeting - 20130917

1 Attendance

Present:

Benjamin Hugo
Heinrich Strauss
Brandon Talbot
Prof. James Gain

Unavailable:

Dr. Patrick Marais

(Prof. James Gain's Office, Building 18, UCT Upper Campus)
from 12h00 to 12h30

2 Agenda

Progress on Zero-Length Encoder (Heinrich)

- SSE/AVX in- progress
- GPU versions not started.

(concern raised by Prof. Gain)

- Nose to the grindstone!

Progress on Predictive Encoder (Benjamin)

- Refactoring documents after draft
- Sort implementation-wise
- CUDA progress slow
  - Calidated and debugged
- Own kernels
  - Prefix sums are as previously discussed with James
- Needs more detail in background
- Small intro - not 10 pages
  - ~3 for Ben
  - ~2 paragraphs on SKA

Progress on Huffman Encoder (Brandon)

- Brandon working on version using thrust::sort
  - sort method requires 2x memory
  - not in-place
  - used STL sort instead
  - 10-11x slower

- Speed given considers cost of transfer
  - only do binning on GPU

- SSE no basic methods on vectorised data
- Binning size-disparity in on input/output between CPU and GPU versions
- No hashmaps in SSE
    - Consider Plain-Old Data-structure?
- Will benchmark against own kernel to determine transfer speed
- Time kernel and subtract data transfer time.
- Check occupancy
    - consider normal CUDA kernel
    - sorting and binning based on STL implementation
    - because algorithm is largely cpu-bound, brandon must consider transfer costs
- Consider disk impact
    - having a disk implementation will simplify comparisons against bzip2 and gzip
    - bzip2 is the definite thing to beat (speed and compression ratio)
- Brandon claims that the files have only 9000 unique values, regardless of file size
- Kernel Invocations
    - occupancy 3
- Prefix sum - usually implemented using multiple kernel calls
    - best sort for GPU is parallel radix sort
- Key thing is to use NVIDIA Visual Profiler

3 Further Meetings

Next meeting scheduled for 24th September 2013 at 12h00 (meetings to follow weekly).

Notes

- Thesis must be in single-column format - see tree-sketching from 2011 or 2012 (Mark Danoher)
- Reasonable abstract length.
- Use of ``we'' is signal of novelty in a paper
- Never use ``I''
- use the ``elegant variation''