Minutes of Honours Progress Meeting - 24 September 2013

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 13h00

2 Agenda

Progress on Zero-Length Encoder (Heinrich)

- think fine-grained, not multithreaded
- level of proactivity must be high

Progress on Predictive Encoder (Benjamin)

- Using GF670 in the Lab to get speeds
- Last memory transfer kills throughput - same goes for Brandon's implementation
- Is GPU -> CPU slower than CPU -> disk? Prof. Gain thinks that's highly unlikely
- Transfer cost is an issue again.
- Consider slowing compression to inflating throughput (better goals)
- elaborate on multiple running kernels
- Two kernels should use seperate memory to prevent locking
- Prefix Sums compacted at bit level; residuals (variable length) compacted at byte level.
- Reduction phase - using CPU instruction for loading into counters

Progress on Arithmetic Encoder (Brandon)

- Started Redocumentation
- All timings on thrust (done); found a flag for thrust with openmp
- 2GB data set on CPU runs in 27s; Data copy to GPU copy is 17.6 s
- Mapping back w/ hashmap is 0.25 s - Can't pre-process this like Ben can
- If task is CPU bound, can asynchronously copy data to GPU
- Speed of GPU 2GB is similar to CPU implementation (because of memCpy)
- Consider moving generat tree to GPU; (tree will be smaller; might cancel hashing of codes)
- Implementation (and data) might not all fit on 1 SM.
- Might be possible to bring down to ~2GB in 9 seconds without the copy
- *Brandon's note -> Generate tree on CPU*
- use GPU for float-> code boolean value issue [try array instead of hashmap]
- Output is ~40% of float size

3 Post Meeting Action Items

Action: *GPU Direct - possibility for consideration*
Assigned To: Project Team
Deadline: 01 October 2013

Action: *Focus on design doc for CUDA over implementation*
Assigned To: Heinrich
Deadline: 01 October 2013

Action: *Focus on CUDA implementation over write-up*
Assigned To: Benjamin, Brandon
Deadline: 01 October 2013

Action: *Find disk access speed (SKA disk array);is it comparable to PCIe bus speed?*
Assigned To: Project Team
Deadline: 01 October 2013

Action: *Ensure that implementations write out to disk (for comparison to gzip, bzip2, etc)*
Assigned To: Project Team
Deadline: 01 October 2013

4 Decisions Made

TOPIC 1:

Decision 1
Decision 2

5 Further Meetings

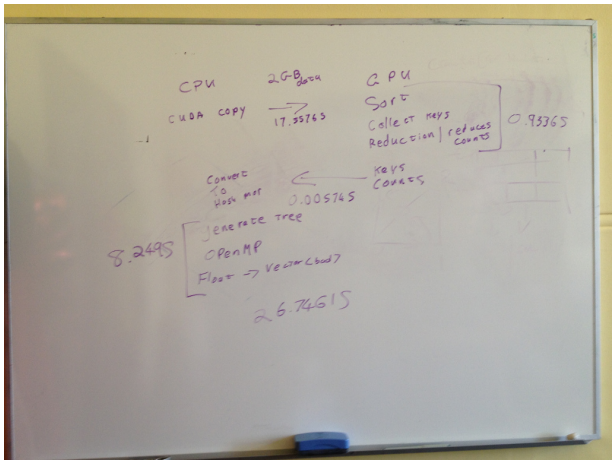
Next meeting scheduled for 01 October 2013 at 12h00 in Prof. Gain's office..

Notes

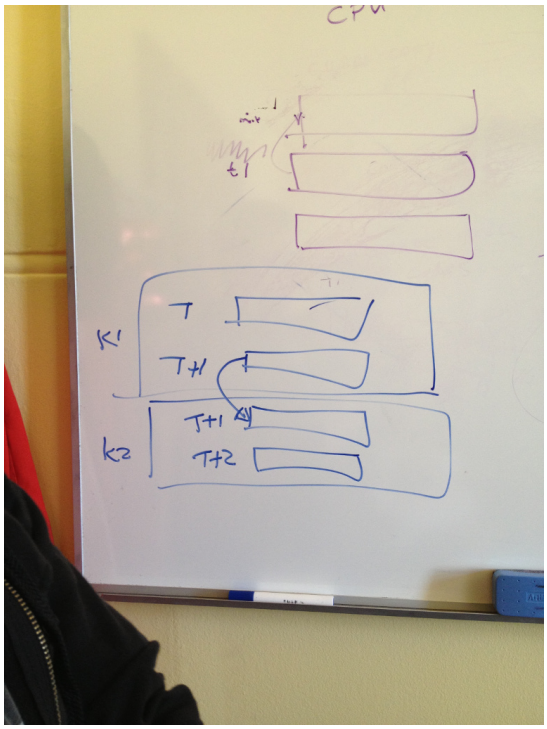
About Writing:

Pseudo code and structure diagrams are good
implementation capped at ~40% of document

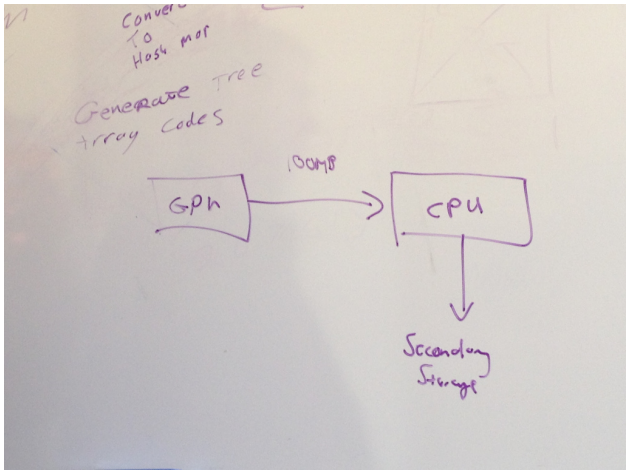
Consider the possibility of using texture memory for some operations



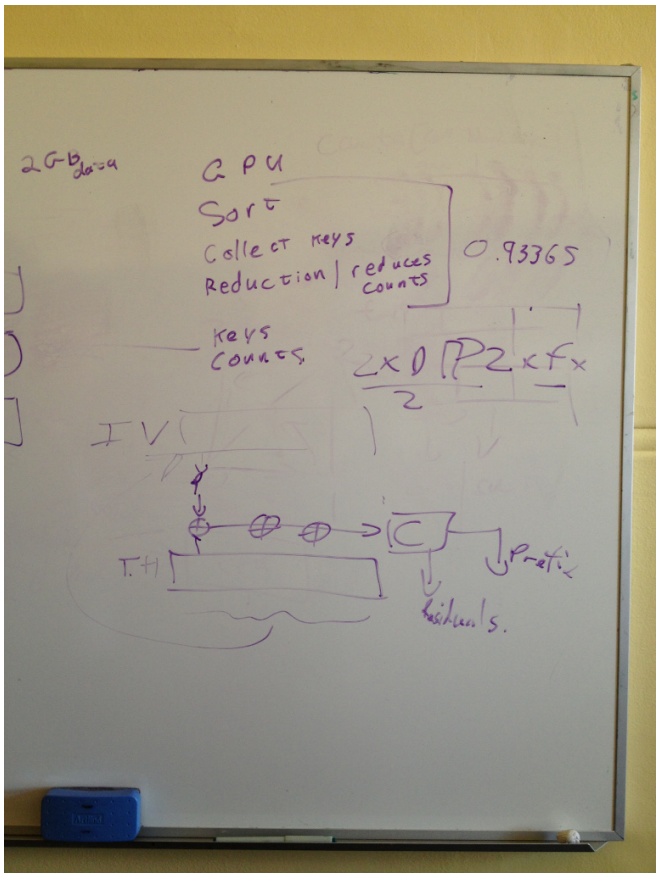
Brandon: discussion of algorithm
float -> vector<bool> (in image)



Brandon: reducing number of copies / hiding the transfer cost



Brandon: double-take on algorithm



Ben: algorithm layout