Minutes of Honours Progress Meeting - 21 August 2013

1 Attendance

Present:

Benjamin Hugo

Heinrich Strauss

Brandon Talbot

Prof. James Gain

Dr. Patrick Marais

(Prof. James Gain's Office, Building 18, UCT Upper Campus)
from 11h30 to 12h00

2 Agenda

Progress on Zero-Length Encoder (Heinrich)
- James recommends that focus be placed on single-threaded (in terms of accuracy validation). Get that done before attempting others. After that an OpenMP version can be attempted and if time is left a vectorized intrinsics version can be attempted.
- A CUDA version will likely be out of scope, but a decision will be reached at the meeting on the 6th of September.
- Risk mitigation strategy:
  - In the worst case, project can be scaled back at the expense of marks

Progress on Predictive Encoder (Benjamin)
- Good progress for Ben; he has completed his investigation into several linear schemes including: (integer-based) mean, median (linear average time complexity), parallelogram (1D Lorenzo predictor) and a gradient-based approach called Lagrange polynomial extrapolation (which works well for smooth data as he pointed out, but not for the data provided by the SKA). Only the median-based approach achieved a slightly better ratio on some of the data sets, but not enough to warrant it as effective replacement for the current XOR-based difference.
- Work can begin on design chapter and a CUDA implementation. He should be wary of CUDA implementation as a time sink; it usually consumes all available time in the project. It was suggested that Thrust be used to implement parallel prefix sums, but since the data is already on the card (compression will in this case be a post-processing operation) Thrust may not be useful (as Brandon pointed out Thrust will do it's own memory copy from host to device and back).
- James inclined to go with block based - confirmed after the meeting. The consensus was to focus on exploration, but James warned to leave enough time for write-up.

Progress on Arithmetic Encoder (Brandon)
- Brandon going with breakpoint type scheme. With larger files (more unique value); breakpoint over 4-bytes
- Decompression is slower on his home pc, in honours lab pc's are faster. 4.5-5 mb/s on lab pc's
- Compression rate speedup from 3.8 to 4.5 or 5 MiB/s
- Parallel binning and dictionary generation may introduce race conditions.
- Compression much better than Ben's methods: .59 vs. .92 -- .93
- Brandon to determine whether slow decompression is a viable trade off from Jason Manley.
- No GPU for Brandon, so do both approaches (Try both blocking and sequential parallelism, and test the speed difference compared to compression.). (Prof. Gain in favour)

- Brandon to block data stream; need to check whether compression ratio is affected (due to repeated dictionary)

- Check algorithm compared to BZIP2 for the data.
  Which is more easily or effectively parallelizable?

3 Post Meeting Action Items
Action: *Focus on implementation over write-up*
Assigned To: Honours Project Team
Deadline: Until CPU versions are both complete.

4 Decisions Made
Predictive Methods:

- Ben to investigate block-based CUDA approach, given limited portions of the memory and compute capability of a card.
  - Time-permitting this can be expanded to include Thrust library from CUDA.
- Brandon to pursue both block-based parallel approach and aggregated parallel approach
- Heinrich to finish single CPU version first
- Compare algortihms to bzip2 not gzip.

5 Further Meetings
Next meeting to be scheduled for 30 August at 12h30. Further meetings will follow weekly. If possible, Jason Manley should join us at UCT Campus.