

Thesis Roadmap

Phase 1: Running Local Version (Completion April 30th)

- Change collector class to hold a specific dataset format (April 24th)
- Select/Complete appropriately complex example program (April 30th)

Phase 2: Running Cluster Version (Completion May 10th)

- Modify writing of datasets to utilize HDFS (May 10th)

Phase 3: Robust Cluster Version (Completion June 25th)

- Complete appropriately comprehensive battery of test programs/datasets (June 10th)
 - o Data: Text, Numerical, Graph, combinations
 - o Analysis: Simple, Iterative, Delta-Iterative, combinations
- Expand visualization portion beyond basic functionality (June 25th)

Phase 4: Written Document Completion (Completion July 24th)

Four weeks planned to be devoted to writing, with an additional week after planned completion before the hard deadline. This begins from the submission of my first rough draft, and allows time for two week long reviews of the writing (assuming a busy schedule on the reviewer's end). Most likely, the specifics here will change dramatically based on circumstance and the condition of the first draft.

- Write final rough draft (July 1st)
- Review and receive feedback (July 7th)
- Adjust based on feedback (July 13th)
- Check for additional issues (July 20th)
- Complete last edits (July 24th)

Phase 5: Real-time Version (Completion July 31st)

Assuming that either the previous phases complete ahead of schedule, or that sufficient time is available during the writing/review phase I will direct efforts toward real-time functionality.

- Implement message passing from collectors