

# Parallel Programming and Computing: Group Project

Christopher D. Carothers  
Department of Computer Science  
Rensselaer Polytechnic Institute  
110 8th Street  
Troy, New York U.S.A. 12180-3590

March 28, 2019

**HARD DEADLINE: 11:59 p.m. Friday, April 26th, 2019**

## 1 Description

The central key requirement of your group project is that it **must involve** the design, development in software and experimentation of a massively parallel system. **A key factor is that your parallel system must involved a high degree of interaction and not be considered “embarrassingly parallel”**. This means that projects involving “wide-area”, “campus-area” or other asynchronous, master-slave approaches are not a viable project for this course.

*The end goal is that you are making progress towards a result that is of high enough quality to be published.* Beyond that, you are free to explore any particular “non-embarrassingly” parallel application that is relevant to your research. More specifically, use the format of papers we have read as part of this course as a template.

Examples of a project might be:

- Re-implement a current serial/research code using MPI to execute effectively on the Blue Gene/Q system.
- Parallelization of components in an existing HPC code using threads with MPI.
- Parallelize an existing sequential data algorithm using threads, or MPI (CUDA could be allow under special instances - talk w/ Prof. Carothers) parallel programming frameworks or a hybrid combination of 2 or 3 of them.

- Design a performance benchmark that tests how MPI + threads with MPI parallel I/O perform together on the Blue Gene/Q system. What are the overheads as you change the mix of MPI ranks and threads and how are reading and writing data rates to the parallel I/O system effected ? Here, you'll want to consider a number of point-to-point communications and collective calls.
- Construct a new PARALLEL discrete-event model using a parallel simulator (e.g., ROSS). For example, create a human “agent” model of a large-scale city that is being evacuated due to a weather disaster.
- Final note: if your project does not involve MPI, threads or CUDA at some level you are probably on the wrong track for this project and should speak with the Prof. Carothers.

The key components of your project submission paper:

1. First and foremost, your project **MUST** be about High-Performance Computing and performance. So, this excludes any sorts of Cloud Computing projects or embarrassingly parallel “over the web” projects. While those are valid types of computation they are not what this course is about. We are concerned about high-performance computing on supercomputing class systems.
2. **It is critical you pick something you and your team can complete.**
3. List your team members. Teams can have up to 3 members. For each team member describe their contribution to the overall project. That is, what did you do besides attend meetings?.
4. Describe your parallel implementation both code and algorithms used.
5. We have read a number of papers this semester. So, provide a review of related published articles. Do a **Google Scholar** search. In terms of related work, pick key older papers provided they have a very high citation count otherwise pick the newest results possible from key journals and conferences - e.g., look for IEEE, ACM and SIAM conference and journal publications.
6. **Performance results:** This will include your sequential and parallel results and indicate your overall speedup. You should perform a “strong scaling” study (i.e., problem size remains the same and processor count increases).

7. **Analysis of performance results:** Provide additional information on why your performance turned out the way it did. In particular, you should understand how much communication overhead your program incurred versus doing real computational work. This should be measured and quantified. Use the cycle counter `GetTimeBase()` to measure these aspects.
8. Summary and future work. Provide a summary of what you did and directions of where you think this project could go in the future. That is what problems did you not have time to solve?
9. **You can expect your paper/project write-up to be a similar length regular conference paper with the following provisos: For undergraduate teams your report should be at least 6 pages double column, single spacing, 10 point font with performance graphs and references. For graduate teams, your report should be at least 9 pages double column, single spacing, 10 point font with more performance graphs and more references. As paper examples, use the papers we discussed in class.**
10. **Submit a final copy of your report in PDF format on [Submitty.cs.rpi.edu](http://Submitty.cs.rpi.edu).**
11. **Place your code on [kratos.cs.rpi.edu](http://kratos.cs.rpi.edu) and note in the write-up which team member's account has the code. Also, provide some sort of README on how to run the code.**
12. EMAIL Prof. Carothers early if have questions or want any sort of guidance about your project. Waiting until a few days before the deadline to ask how or what you should is not acceptable.
13. The deadline (11:59 p.m. April 26th, 2019) is hard because it is the last day of class and class assignments are not suppose to extend beyond that point in the semester. So, please plan your use of the supercomputer accordingly and expect the time for 128 node jobs to take many hours to get through the system.