

Project 2:

Neighbourhood data analysis in a parallel, multi-scale computational model of the kidney:

Background:

This project is part of a larger project where we are using a systems approach to build a model of an entire kidney by modelling individual kidney units (nephrons) and their interactions.

Our current implementation is the size of an entire rat renal system (two kidneys), or 64,000 nephrons. We are looking to find patterns in the output data, so that key events in renal disease can be defined and explored. Disease events might be, for example, a change in function of a few nephrons at the onset of renal disease, or a pattern of progressive malfunction.

The current implementation uses a distributed grid of processors, each simulating multiple nephrons. Because the system is so large, and there are so many time steps, the size of the output data makes it infeasible to store or analyse the data in reasonable time and space. In this project, we exploit the idea of neighbourhood, or nodes that are in close proximity to one another, in parallel computing systems.

Goal:

The goal of this project is to investigate methods to summarize collective function among groups of processing nodes. The aim is to reduce the quantity of output data to a manageable amount by developing a hierarchical neighbourhood structure, such that, not every piece of data needs to be written at every step.

Scope and timelines:

For a three-semester 50-point project, we would expect the timelines to be:

Semester 1:

Literature review of neighborhood analysis in parallel computing systems, with a particular emphasis on the context of the kidney model. Literature review to be submitted.

Semester 2:

Analyze and implement indexing structures suitable for writing out data from the Blue Gene machine. Demonstration of system to be given.

Semester 3:

Implement one or more neighborhood methods for summarizing the kidney data as it is written, and analyse improvements.

Student:

The project would suit a student with a strong computer science background and an interest in parallel computing.