Upgrades to Tufts High Performance Computing Cluster Benefit Researchers

UIT recently completed a research cluster upgrade project, that will greatly benefit departments across the University working with large data sets, model building, and simulations.

High Performance Computing (HPC) connects a set of computers together via a high-speed network to provide a single, shared resource. Over the past three years, UIT has observed an increasing demand for the HPC research cluster and the new High-Performance Computing (HPC) research environment brings more than 1000-cores of computing power to the Tufts community.

HPC in Action

Several departments across the University are engaged in projects using HPC. These include the simulation of string networks for the investigation of string theory, analysis of data from the Large Hadron Collider in Physics and Astronomy, and simulation of blood flow in the kidneys in Chemical and Biological Engineering. Health Sciences and Mathematics use HPC to perform next-generation sequencing analysis and the calculation of partial differentiation equations in fluid dynamics, respectively.

"TO ACHIEVE THE PHYSICS GOALS OF THE TUFTS GROUP, SEVERAL HUNDRED TERABYTES OF DATA WILL NEED TO BE PROCESSED SEQUENTIALLY. GOOD NETWORK PERFORMANCE IS CRUCIAL FOR THE ANALYSIS TO PROCEED IN A TIMELY FASHION AND THE NEW TUFTS CLUSTER PROVIDES 10 GB INTERCONNECTIVITY."

— Professor Austin Napier, Physics and Astronomy

Research Associate Professor Ken Olum, from the Tufts Institute of Cosmology, makes an observation about the Cluster shared by several researchers across the University: “It makes much more sense to have this facility available as a university-wide resource than to build such a cluster in our own group or department. As a university-wide resource, it can be shared by many different groups so that a lot of resources are available when you need them, but when you don’t, the resources are used by others.”

Professor Olum uses the Cluster to do large-scale simulations of cosmic strings for the Tufts Institute of Cosmology: “We simulate a string network containing up to 10 billion string segments. Obviously this takes a lot of computational power. To get the data presented in our last paper, we used about 27 years of CPU time. To do that in a reasonable period of time, you need a large cluster of processors. We typically use several hundred processors for a few days at a time.”

Lakshmanan (Lax) Iyer and Chao Qiang Lai, from Tufts Center for Neuroscience Research, use HPC to perform next-generation sequencing analysis. "In collaboration with Dr. Ron Lechlan the Chief of Endocrinology at the Tufts Medical Center, we are applying cutting edge technology to determine the gene expression profile of Tanyeya, a special cell of glial origin in the brain. While much is known about the anatomy of these cells, their physiologic functions remain speculative and enigmatic. The results of these studies would provide clues towards their function. This analysis requires considerable amount of disk storage and CPU time. Without the Tufts high performance cluster and the associated storage it would be impossible to make sense of this data."

The High Energy Physics group at Tufts University is using the new HPC cluster to analyze particle interactions from the ATLAS experiment at the CERN Large Hadron Collider in Geneva, Switzerland. The LHC is engaged in the search for the lost missing particle in the Standard Model of particle physics (the Higgs particle). The Tufts group, which consists of professors Krysztof Sliwa, Austin Napier, and Pierre-Hugues (Hugo) Beauchemin, Dr. Sharka Todorova (residing at CERN), graduate students Samuel Hamilton and Jeffrey Wetter, and undergraduate George Thompson, is studying the production and properties of top quarks, searching for new particles containing bottom quarks, and conducting model-independent searches for "new physics," such as supersymmetry (SUSY).

Says Professor Napier, "To achieve the physics goals of the Tufts group, several hundred Terabytes of data will need to be processed sequentially. Good network performance is crucial for the analysis to proceed in a timely fashion and the new Tufts cluster provides 10 GB interconnectivity. Tufts is a Tier 3 site and we are able to access large datasets stored at BU and process the data on the HPC cluster at Tufts." About 2 Petabytes of data are already available for analysis. Besides a copy at CERN, additional copies of the data are stored worldwide at Tier 1 and Tier 2 Computing Sites. HPC clusters at Boston University and Harvard provide a local Tier 2 site in the Boston area.

Dr. Michael Levin describes how the Tufts Center for Regenerative and Developmental Biology uses HPC: “While we use the cluster for some statistical analysis, we use it mostly to run models in developmental, regenerative, and synthetic biology. We use Matlab and other tools to run agent-based models and genetic algorithms to simulate cell-to-cell communication using bioelectrical events (endogenous voltage gradients).”

Dr. Daniel Lobo, a Postdoctoral Associate in the Biology Department and Tufts Center for Regenerative and Developmental Biology, expanded on the work they do using HPC to run computational models: "We are planning to use the cluster in a project about regenerative biology; more precisely, about planarian (flatworm) regeneration. The aim of our project is to discover a comprehensive model of planarian regeneration, which would pave the way for advances in the biomedical areas of cancer, birth defects, and regenerative medicine but also in the areas of cybernetics, computer science, robotics, and many areas of engineering."

The research cluster user website is available at http://research.uit.tufts.edu

Self-provisioning: http://research.uit.tufts.edu

The self-service research storage webform is available at http://researchstorage.uit.tufts.edu