

A Neuroscience Gateway for High Performance Computing

*N.T. Carnevale¹, S. Sivagnanam², K.K. Yoshimoto², V. Astakhov³, A.E. Bandrowski³, M.E. Martone³, A. Majumdar²

¹Yale University School of Medicine, New Haven, CT; ²San Diego Supercomputer Ctr., San Diego, CA; ³University of California, San Diego, CA

Introduction

We are developing a Neuroscience Gateway (NSG) that will provide neuroscientists a user-friendly Web-based interface for using essential high-performance computing (HPC) and other cyberinfrastructure (CI) resources. This project is motivated by the emergence of research questions requiring simulations that impose a large computational burden and/or heavy storage requirements. Powerful domain-specific open source tools, such as NEURON, GENESIS, MOOSE, NEST, and PyNN, can be used to construct models suitable for simulation on parallel computers, but only a few investigators have been able to make use of parallel hardware, and even fewer have run simulations on extreme scale HPC machines.

The problem and its solution

The high barrier to entry is the main reason why the broader neuroscience community has not been able to use HPC resources. Common problems include

- the daunting process of requesting time on HPC resources
- complex administrative policies and batch system details
- difficulty installing applications on HPC resources
- difficulty managing workflow that involves multiple remote authentication schemes
- data transfer, storage, and output retrieval issues

These time consuming, confusing details differ significantly from one facility to another. Investigators have to do most of the work themselves, diverting effort from more productive activities.

The NSG will enable computational neuroscientists to access HPC resources, funded by NSF and other support, through a convenient interface to simulation software that is already configured for optimal use. To this end, the we will build an infrastructure layer, called a science gateway [Wilkins-Diehr 2008], that reduces administrative complexities and abstracts away most technical details, so that investigators can easily access the software they need on very large HPC resources. This will be accomplished through a point-and-click browser environment for configuring and running domain-specific programs. All CI/HPC-related complexities will be hidden within the black box of the gateway architecture. This approach has been successful in other research domains, where gateways have catalyzed wider usage of HPC resources. One example is the CIPRES gateway [CIPRES, Miller 2010] (Figure 1).

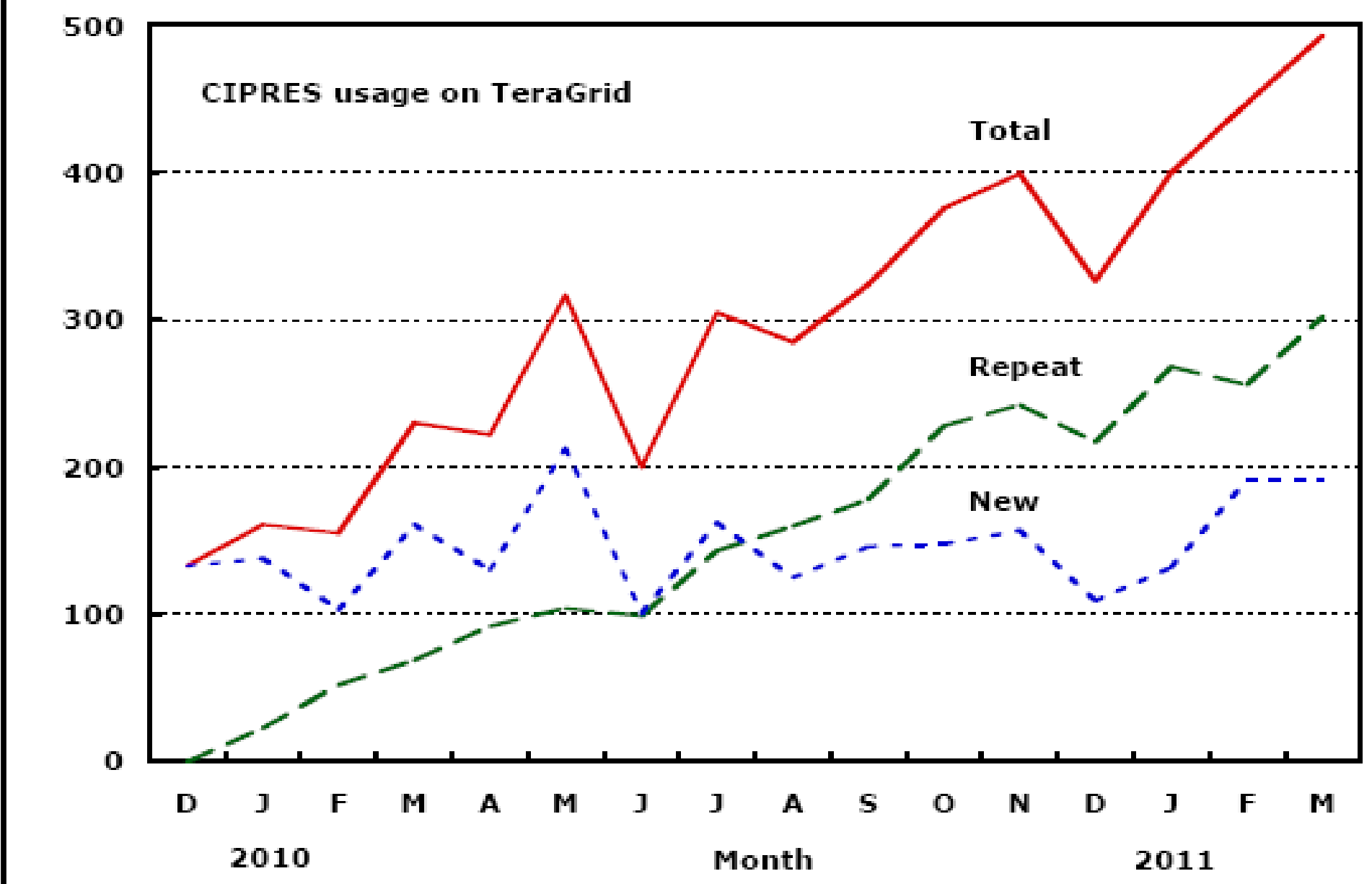


Figure 1. Number of CIPRES Gateway users per month running jobs on the TeraGrid. CIPRES was developed at the San Diego Supercomputer Center (SDSC) for the computational phylogenetics research community. The number of users who submitted one or more jobs per month increased from 132 in 12/2009 to more than 500 by 3/2011. The number of new users/month was >100 from the inception of CIPRES, and the number of repeat users increased steadily.

Design and implementation

The NSG will be designed and implemented to provide a simple web portal based user environment for uploading model source code and running simulations on HPC resources. It will offer a selection of commonly used neural simulators that are optimally installed on various HPC resources. It will be implemented so as to distribute jobs transparently across those resources while minimizing job loss due to system/hardware errors and providing elegant recovery from failures.

The NSG will make it easier for users to

- acquire HPC time under reasonable usage guidelines and policies
- specify parallel simulation parameters, e.g. number of cores, memory per core/node, estimated job run time etc.
- query job status, and request automatic notification of job completion
- securely access and download output results

From the user's point of view, the workflow will be simple (Figure 2).

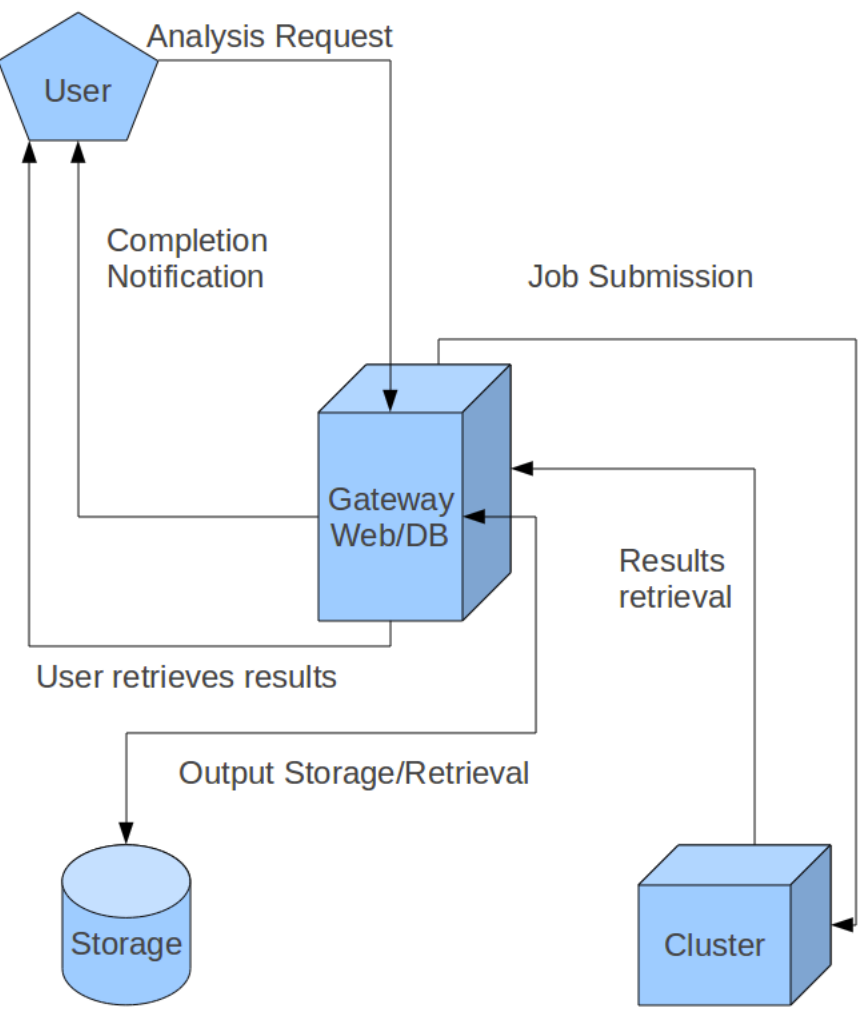


Figure 2. NSG workflow from the user's perspective.
1. User logs into the NSG.
2. User uploads input data.
3. User requests simulation run.
4. NSG front end sends input data and job request to remote cluster.
5. NSG retrieves output data from remote cluster.
6. NSG notifies user of job completion status.
7. NSG provides information about output data location and retrieval.

The NSG will also provide

- documentation and examples of how to parallelize neural models for running on HPC machines
- programmatic interfaces to other neuroscience tools (e.g. ModelDB, WBC, NIF, neuroConstruct)
- interfaces for simulator/tool developers that simplify installation and testing of community codes
- a community environment for collaboration, data sharing, etc.

Benefits

The Neuroscience Gateway will benefit the broader neuroscience research community in several ways.

- Its simple interface will provide streamlined access to HPC resources without forcing investigators to divert effort from their research.
- Users with limited local (university-scale) resources will be able to address pressing research questions that require access to large scale, advanced systems.
- Community code developers (i.e. developers of software such as NEURON, GENESIS, MOOSE, NEST, PyNN) will be able to test/debug, benchmark, and scale codes on large scale resources and, when satisfied, make them available to the broader user community.
- The NSG can also serve as a vehicle for making new community codes available to users.
- It can be used for teaching classes, workshops, and tutorials without having to set up codes on HPC resources or create new accounts for students/participants.
- It will offer a forum for user interactions such as collaboration and sharing of data and curricular materials.

Attic for figures that may yet be used

