

## SDSC's CIPRES Gateway provides window to the past -- and fast

March 28 2011

---

A novel supercomputing resource created by researchers at the San Diego Supercomputer Center at the University of California, San Diego, is allowing scientists to study evolutionary relationships among large populations of living things in significantly shorter times – and without having to understand how to operate large, complex computer systems.

The new resource, called the CIPRES (Cyberinfrastructure for Phylogenetic RESearch) Gateway, is an Internet portal that allows scientists anywhere in the world to upload their data via a standard Web browser and perform phylogenetic analyses. The most time-consuming analyses use supercomputers, such as SDSC's new Trestles system, that are part of the National Science Foundation's TeraGrid, the world's most powerful collection of high-performance computing resources dedicated to academic research.

"In addition to answering the age-old questions of how all living things are related to each other, understanding evolutionary relationships has some very important practical benefits," said Mark Miller, principal investigator in SDSC's Research, Education and Development group, and leader of the CIPRES Gateway project. "For example, knowing the evolutionary relationships among a group of viruses or bacteria can help doctors understand where an infection came from, effectively treat patients who are infected, and work to contain the spread of disease during an outbreak."

Moreover, understanding how individual species adapt for survival in a

specific geographic location can help scientists manage a species for long-term survival in that location, or engineer crops for higher productivity in a particular location.

Evolutionary relationships are uncovered by comparing DNA sequences from individuals under study. Just as a single DNA sequence can be used to identify a criminal with a very high degree of accuracy, a group of DNA sequences can be used to determine just how closely related any group of living things are with great precision.

"DNA sequences from individuals can be prepared so quickly and cheaply now, we can understand [evolutionary relationships](#) more accurately than ever before," according to Miller. "The problem is, the number of computations required grows quickly as the amount of data grows. There are only three possible relationships between any four individuals, but there are more than two million different relationships between 10 individuals. A computer that could analyze a million trees per second would require about 20 billion years to test all the possible relationships for just 22 individuals!"

Solving this problem is where the CIPRES Gateway and TeraGrid supercomputers come in. The power of supercomputers comes from parallel computing, in which large analyses are broken into smaller pieces that are run simultaneously on many processor cores. Under the TeraGrid's Advanced User Support program, Wayne Pfeiffer, a distinguished scientist at SDSC, helped improve the parallel performance of RAxML and MrBayes, two widely used phylogenetics codes.

"Most RAxML analyses submitted to the CIPRES Gateway now run on 60 cores of Trestles," said Pfeiffer. "With a typical speedup over a single core of about 30, this means that analyses that would require a month on a laptop can be completed in a day via the gateway."

"This is an excellent example of how science is being transformed through new ways of leveraging the capabilities of today's supercomputers," said Richard Moore, SDSC's deputy director.

"Significantly reducing the time it takes researchers to run such complex analyses, while freeing them from having to fully understand all the intricacies of today's supercomputers, means greater scientific productivity. This is what makes the CIPRES Gateway such a valuable phylogenetic resource."

Although SDSC's CIPRES Gateway has been in operation for a little more than a year, it has already provided immediate benefits to the scientific community, both in time savings and in new discoveries. To date, more than 2,000 scientists have run more than 35,000 analyses for approximately 100 completed studies. These studies span a broad spectrum of biological and medical research.

One study, recently published in the journal *Parasitology*, focused on research that showed humans are much more likely to infect apes with malaria, than the reverse. Details of that study can be found [here](#).

"Without the CIPRES Gateway, this work, and the other projects I am working on, would not go as quickly or as smoothly," said James B. Munro, a researcher from the University of Maryland School of Medicine, who was part of the research team that reported the new insights into the complex relationship between the malarial parasites and their mammalian hosts.

Provided by University of California - San Diego

Citation: SDSC's CIPRES Gateway provides window to the past -- and fast (2011, March 28) retrieved 12 May 2024 from <https://phys.org/news/2011-03-sdsc-cipres-gateway-window-.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.