IU partnership results in faster Trinity RNA sequencing software

July 18 2012

Key software used to study gene expression now runs four times faster, thanks to performance improvements put in place by a team from the Indiana University Pervasive Technology Institute (PTI), the Broad Institute of MIT and Harvard and Technische Universität Dresden.

The timesaving breakthroughs will allow bioinformaticians and biologists who study RNA sequences to analyze more data in a shorter amount of time. This will speed the understanding of biological processes in fields as diverse as ecology, evolution, biofuels and medicine.

Robert Henschel and Richard D. LeDuc, of PTI and IU's National Center for Genome Analysis Support (NCGAS), announced the findings today at the XSEDE12 conference in Chicago. Henschel and LeDuc, along with partners from the Broad Institute and the Center for Information Services and High Performance Computing (ZIH) at Technische Universität Dresden, teamed up to announce this advance in a fast-growing area of computational biology.

The software, known as Trinity, was developed by researchers at the Broad Institute and Hebrew University. It produces high-quality RNA sequence assemblies used by scientists studying gene expression. These RNA sequence assemblies allow scientists to know which genes are active within a living creature. Trinity is especially useful for studying organisms without a complete genome sequence, such as agricultural pests, ecological indicator species and human parasites.

The software has long been considered a leader in the field, but it needed some finetuning.

"IU research technologists strive to deliver tools and services that accelerate discoveries for scientists all over the world. By collaborating with our counterparts at Broad and ZIH, we were able to do just that with Trinity. This is just one example of how the various centers affiliated with PTI—such as NCGAS—improve the capabilities of scientists at home and abroad," said Craig Stewart, executive director of IU's Pervasive Technology Institute and principal investigator of the National Science Foundation grant that funds NCGAS.

"In the past, Trinity was a high quality tool but the run time was too long," said Henschel. "Now with our performance improvements, it runs as fast as the competition—if not faster—and still produces superior quality sequence assemblies."

The partners first used standard high performance computing techniques to improve the software's speed. Specifically, this involved building Trinity with an optimizing compiler for the Intel® Xeon® architecture and using optimizing compiler flags. In addition, the team properly configured the application to take full advantage of multicore, multisocket compute nodes in today's clusters.

Next, the team finetuned each part of the Trinity package to improve the overall scalability of the application. They used Vampir performance analysis tools, developed at ZIH, to gain insights into the software's performance. The optimizations included improving and parallelizing input/output, simplifying data structures for better performance and optimizing parallel regions in the application.

Henschel is hopeful that IU's work with Trinity will continue. "We are working on establishing a continued collaboration between IU, Broad

and ZIH to further optimize Trinity," said Henschel. "We hope these performance improvements are just the beginning of a longer term relationship that will continue to benefit biological research."

More information: www.xsede.org/

Provided by Indiana University

Citation: IU partnership results in faster Trinity RNA sequencing software (2012, July 18) retrieved 23 June 2024 from https://phys.org/news/2012-07-iu-partnership-results-faster-trinity.html