How to improve data management in the supercomputers of the future

February 1 2017

How to improve data management in the supercomputers of the future. Credit: UC3M

Researchers at Universidad Carlos III de Madrid (UC3M) are establishing new foundations for data management in the supercomputing systems of the future. In recent decades, many scientific discoveries have depended on the analysis of an enormous volume of

data, which is done essentially through computational simulations performed on a large scale in supercomputers. This type of machine is used to study climate models, the development of new materials, research into the origin of the universe, the study of the human genome and new applications in bioengineering.

At present, as an ever-increasing amount of information is collected and stored, scientific [data management](link) confronts a problem: The software that manages the latest generation of supercomputers was not designed for the scalability requirements that are expected in coming years. In fact, in less than a decade, these infrastructures are going to be two orders of magnitude faster than current supercomputers.

"Today, these applications are encountering big problems of performance and scalability due to the exponential increase of data as a result of better instruments, the growing ubiquity of sensors and greater connectivity between devices," explained professor Florin Isaila, from the group ARCOS in the UC3M Department of Computer Science. "These days, a radical redesign of the computational infrastructures and management software is necessary to adapt them to the new model of science, which is based on the massive processing of data."

The objective of the project, "Cross-Layer Abstractions and Run-time for I/O Software Stack of Extreme-scale systems" (CLARISSE), is to increase the performance, scalability, programmability and robustness of the data management of scientific applications to underpin the design of next-generation supercomputers.

Historically, [data management software](link) has been developed in layers with little coordination in the global management of resources. "Nowadays, this lack of coordination is one of the biggest obstacles to increasing the scalability of current systems. With CLARISSE, we research solutions to these problems through the design of new

mechanisms for coordinating the data management of the different layers," said Professor Isaila.

Jesús Carretero, the project's main researcher, UC3M full professor and head of ARCOS, explained, "At present, ARCOS is actively involved in several initiatives around the world to remodel the management software of future supercomputers, including the coordination of the CLARISSE project and the research collaboration network NESUS. The resulting synergies of these efforts are going to contribute substantially to accelerating scientific discoveries in the coming decades."

More information: Further information: arcos.inf.uc3m.es/~florin/clarisse

Provided by Carlos III University of Madrid

Citation: How to improve data management in the supercomputers of the future (2017, February 1) retrieved 25 April 2024 from https://phys.org/news/2017-02-supercomputers-future.html