

Diagnosing supercomputer problems

November 13 2017



Sandia National Laboratories computer scientist Vitus Leung and a team of computer scientists and engineers from Sandia and Boston University won the Gauss Award at the International Supercomputing conference for their paper about using machine learning to automatically diagnose problems in supercomputers. Credit: Randy Montoya

A team of computer scientists and engineers from Sandia National

Laboratories and Boston University recently received a prestigious award at the International Supercomputing conference for their [paper](#) on automatically diagnosing problems in supercomputers.

The research, which is in the early stages, could lead to real-time diagnoses that would inform supercomputer operators of any problems and could even autonomously fix the issues, said Jim Brandt, a Sandia computer scientist and author on the paper.

Supercomputers are used for everything from forecasting the weather and cancer research to ensuring U.S. nuclear weapons are safe and reliable without underground testing. As supercomputers get more complex, more interconnected parts and processes can go wrong, said Brandt.

Physical parts can break, previous programs could leave "zombie processes" running that gum up the works, network traffic can cause a bottleneck or a computer code revision could cause issues. These kinds of problems can lead to programs not running to completion and ultimately wasted supercomputer time, Brandt added.

Selecting artificial anomalies and monitoring metrics

Brandt and Vitus Leung, another Sandia computer scientist and paper author, came up with a suite of issues they have encountered in their years of supercomputing experience. Together with researchers from Boston University, they wrote code to re-create the problems or anomalies. Then they ran a variety of programs with and without the anomaly codes on two supercomputers—one at Sandia and a public cloud system that Boston University helps operate.

While the programs were running, the researchers collected lots of data on the process. They monitored how much energy, processor power and

memory was being used by each node. Monitoring more than 700 criteria each second with Sandia's high-performance monitoring system uses less than 0.005 percent of the [processing power](#) of Sandia's supercomputer. The cloud system monitored fewer criteria less frequently but still generated lots of data.

With the vast amounts of monitoring data that can be collected from current supercomputers, it's hard for a person to look at it and pinpoint the warning signs of a particular issue. However, this is exactly where [machine learning](#) excels, said Leung.

Training a supercomputer to diagnose itself

Machine learning is a broad collection of computer algorithms that can find patterns without being explicitly programmed on the important features. The team trained several machine learning algorithms to detect anomalies by comparing data from normal program runs and those with anomalies.

Then they tested the trained algorithms to determine which technique was best at diagnosing the anomalies. One technique, called Random Forest, was particularly adept at analyzing vast quantities of monitoring data, deciding which metrics were important, then determining if the [supercomputer](#) was being affected by an [anomaly](#).

To speed up the analysis process, the team calculated various statistics for each metric. Statistical values, such as the average, fifth percentile and 95th percentile, as well as more complex measures of noisiness, trends over time and symmetry, help suggest abnormal behavior and thus potential [warning signs](#). Calculating these values doesn't take much computer power and they helped streamline the rest of the analysis.

Once the machine learning algorithm is trained, it uses less than 1

percent of the system's processing power to analyze the data and detect issues.

"I am not an expert in machine learning, I'm just using it as a tool. I'm more interested in figuring out how to take monitoring data to detect problems with the machine. I hope to collaborate with some machine learning experts here at Sandia as we continue to work on this problem," said Leung.

Leung said the team is continuing this work with more artificial anomalies and more useful programs. Other future work includes validating the diagnostic techniques on real anomalies discovered during normal runs, said Brandt.

Due to the low computational cost of running the machine learning algorithm these diagnostics could be used in real time, which also will need to be tested. Brandt hopes that someday these diagnostics could inform users and system operation staff of anomalies as they occur or even autonomously take action to fix or work around the issue.

This work was funded by National Nuclear Security Administration's Advanced Simulation and Computing and Department of Energy's Scientific Discovery through Advanced Computing programs.

More information: Ozan Tuncer et al, Diagnosing Performance Variations in HPC Applications Using Machine Learning, *High Performance Computing* (2017). [DOI: 10.1007/978-3-319-58667-0_19](https://doi.org/10.1007/978-3-319-58667-0_19)

Provided by Sandia National Laboratories

Citation: Diagnosing supercomputer problems (2017, November 13) retrieved 13 May 2024

from <https://phys.org/news/2017-11-supercomputer-problems.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.