Scientific computing in the cloud gets down to Earth

April 8, 2019

LOS ALAMOS, N.M., April 8, 2019—In a groundbreaking effort, seismology researchers have conducted a continent-scale survey for seismic signatures of industrial activity in the Amazon Web Services commercial cloud (AWS), then rapidly downloaded the results without storing raw data or needing a local supercomputer.

"Using a traditional workflow, to download-store-calculate on a desktop, this work would've taken more than 40 days to do. Using the cloud service, it took just under 7 hours," said Jonathan MacCarthy of the Earth and Environmental Sciences division at Los Alamos National Laboratory. "To our knowledge, this is the first application of streaming cloud-based research in seismology."

While Los Alamos is home to its own gaggle of computers of every size, this was an effort to think outside the substantial local capabilities. The team wanted to explore streaming the research workflow in the Amazon commercial cloud, MacCarthy explained, where a cluster of up to 200 on-demand computers were coordinated to request about 2 terabytes of data directly from the data center. The raw data were never stored, and only the results of the specific calculation were sent back to the researcher's computer. "It's like instructing a swarm of temporary seismologists to do many small parts of a much larger research problem, then disappear," MacCarthy said.

Increasingly, cloud computing services have been used to carry out varied types of scientific research, MacCarthy noted, putting the power of computational clusters into the hands of any researcher with a credit card, not just those who have access to traditional high-performance-computing (HPC) resources. And while Los Alamos is home to multiple supercomputers and even a quantum computer, exploring new options to accomplish the science is always part of a national laboratory's capabilities.

In this case, the specific application has produced the first large-scale map of industrial noise in the US. The project is aimed at understanding seismic harmonic tonal noise from industrial activity, developed using a detection algorithm originally developed by Marcillo and Carmichael. The work was done in coordination with the Incorporated Research Institutes for Seismology (IRIS) Data Management Center (DMC), a community data repository for seismology.

Link to EOS story.