Cloud Computing Testbed Chameleon Launches Third Phase with Focus on IoT and Reproducibility

$10 million NSF grant funds next four years of University of Chicago project

Over the past decade, cloud computing transitioned from a technology used by large scientific collaborations to a fundamental Internet resource supporting services across all major sectors of modern economies. That evolution continues today, as researchers use clouds to drive commercial and scientific innovations, including the Internet of Things (Iota), powerful mobile applications, 5G communications, AI, and serverless computing.

Since it launched in 2015, Chameleon has enabled such innovations by providing thousands of computer scientists with basic research resources required to conceptualize, assemble, and experiment with new cloud computing approaches and technologies. Under a new four-year, $10 million grant from the National Science Foundation, the cloud computing testbed led by the University of Chicago will further broaden its scope, adding new features including those for reproducibility, IoT and networking experimentation, and GPUs. This multi-organizational initiative includes the Texas Advanced Computing Center (University of Texas), RENCI (North Carolina), and the International Center for Advanced internet Research (iCAIR), at Northwestern University.

"Chameleon is a scientific instrument for computer science systems research," said, Kate Kealey, Senior Computer Scientist at Argonne National Laboratory and the Consortium for Advanced Science and Engineering (CASE) of the University of Chicago, and principal investigator of the Chameleon project. "Astronomers have telescopes, biologists have microscopes, and computer scientists have Chameleon." Growing interest among the Chameleon community has defined the priorities for the next phase, Keahey said.

In its first five years, Chameleon has attracted a community of more than 4,000 individuals from over 100 institutions, who are addressing more than 500 different research and education projects. Scientists have used the testbed to study power management, operating systems, virtualization, high performance computing, distributed computing, security, machine learning, and many other research areas. Educators have used Chameleon for cloud computing courses, allowing college and high school students to build their own cloud and learn the inner workings of the technology.

The third phase will add expanded tools for reproducible research, allowing easier publication and sharing experimental results so other scientists can replicate and build on it. Because integration with Jupyter notebooks proved popular with the Chameleon community, the team is developing a sharing portal integrated with Zenodo, where experiments can be easily published and discovered. Also, this phase will provide new options for programmable networking including software-defined networking (SDN), including compatibility with networking testbed projects, such as the NSF FABRIC, and new hardware and storage resources at the project's primary two sites, UChicago, and the Texas Advanced Computing Center (TACC), and at a supplemental site at the International Center for Advanced Internet Research, co-located with the StarLight International/National Communications Exchange Facility in Chicago, which allows for large scale testbed federation and collaboration, including for data intensive science at global scale. That site supports specialized research capabilities including the P4 programming language, large flow appliances, 100 Gbps services, high performance Data Transfer Nodes (DTNs), and testbed integration, including with FABRIC.

Chameleon also supports edge computing, fueled by IoT that supplements local computation with cloud resources, which motivated the creation of a new "Bring Your Own Device" processes. These capabilities will support projects such as Sage, a UChicago/Argonne/Northwestern University collaboration studying the deployment of sensor networks in urban and natural environments.

"Some of the most exciting emergent research directions are in developing the cloud to edge continuum," said Pete Beckman, Director of the Sage project and Co-Director of the Northwestern Argonne Institute of Science and Engineering. "Partnering with Chameleon will allow our users to tap into the potential of such a combined platform."

Unlike traditional computer science testbeds, typically configured using local technologies, the CHameleon Infrastructure (CHI) provides capabilities with mainstream open-source OpenStack, a familiar environment supported by a large, active development community.

To expand access to computer science research testbeds, the project will continue development of its "CHI-in-aBox" testbed packaging. "CHI-in-a-Box allows you to not only build a testbed very, very quickly," Keahey said. "But it also contains a packaging of our operations model including testbed monitoring as well as failure detection and mitigation tools. The idea is to lower the cost of this type of testbed so that we can have more of them, and allow more researchers to run more experiments."

"Chameleon is a great example of how shared infrastructure with over 4,000 users can save the academic community time and money while catalyzing new research results," said Deepankar Medhi, Program Director in the Computer & Information Sciences & Engineering Directorate (CISE) at the National Science Foundation. "NSF is pleased to fund Chameleon for four more years in order to extend the platform with new capabilities, thus allowing researchers to conduct new lines of research and students to learn newer technologies."

For additional information, see "Lessons Learned from the Chameleon Testbed" the team's paper from the 2020 USENIX Annual Technical Conference and www.chameleoncloud.org

Contact:

Joe Mambretti,

International Center for Advanced Internet Research, Northwestern University

www.icair.org

j-mambretti@northwestern.edu