Trusted CI Success Story
Pegasus

Trusted CI engagement with Pegasus focuses on credentials

Whether it's physics, machine learning, astronomy, genomics, weather patterns, or seismology, these research disciplines, and many more, require National Science Foundation (NSF) researchers to run complex workflows through a variety of computing environments.

What if a scientist wants to run her computations through a university data center on Monday but wishes to run them through a cloud storage service on Tuesday?

Pegasus can help with that. A workflow management system funded by NSF, Pegasus helps scientists manage their data from end to end through a variety of distributed-computing environments.

A key Pegasus feature is data management, says Karan Vahi, Pegasus lead architect and senior computer scientist at the University of Southern California Information Sciences Institute. "We use a variety of transfer services that exist on the user's infrastructure to ship data. We manage workflow so users can securely copy data," he said.

In 2012, Pegasus architects realized they needed better management of user credentials in order to safeguard secure file copy, so they contacted Trusted CI, the NSF Cybersecurity Center of Excellence.

Pegasus' workflow management system helps scientists run data through a variety of computing environments.

In collaboration with Trusted CI, Pegasus did a deep dive into its architecture over a period of six months. "We looked at how to avoid storage of secure shell (SSH) credentials to the local file system and what could be implemented to support a more robust and uniform credential process," said Vahi.

Trusted CI provided a variety of recommendations for managing SSH credentials. "We looked at all the options, but we couldn't do all of them. If a secure solution becomes too complex, then the end goal is not achieved in terms of user satisfaction. Trusted CI helped us weigh the pros and cons of options and helped us prioritize the solutions that were practical and not too costly or complex for the users," remarked Vahi.

Before the engagement, according to Vahi, Pegasus' credential management was ad hoc specific to a particular client. "Trusted CI helped us identify and formalize solutions that served as a blueprint for uniformly handling credentials for data management and job submission. The direct impact—we adopted first-class credential handling," Vahi emphasized.

"Above all," he said, "Trusted CI helped us look at the larger picture and come up with a uniform solution. After we did that work, it was much easier for users to input their credentials and much easier for us to support and manage credentials."

"We follow commonsense cybersecurity principals to this day. The Trusted CI team helped us think about uniformity, having a common interface to manage credentials, and making it easier for our users. Our Trusted CI solution is flexible as technology changes. We can adjust. Lessons learned over the years are still relevant," added Vahi.

"It was easy to work with Trusted CI, and that had a great impact on Pegasus because we managed to introduce integrity protection," added Vahi. "In fact, we later worked with members of the Trusted CI team to spin out a new project called Scientific Workflow Integrative with Pegasus and won a Phil Andrews Most Transformative Contribution Award at PEARC19".