SUCCESS STORY | CLEMSON UNIVERSITY

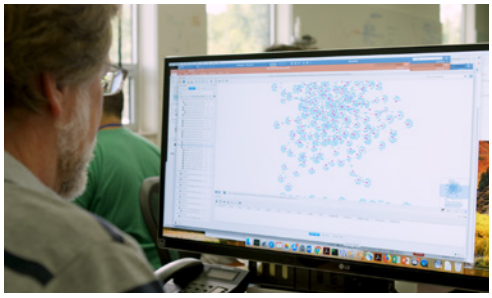
SIMPLIFYING HPC APPLICATION DEPLOYMENTS WITH NVIDIA GPU CLOUD CONTAINERS

Clemson University

NVIDIA

Clemson University HPC administrators support GPU-optimized containers to help scientists accelerate research

FINDING AN ANSWER FOR HPC RESOURCE ACCESSIBILITY



Researcher running simulations to predict a population's propensity to develop diseases

When faculty and graduate students from over 50 university departments are hounding you for high-performance computing (HPC) resources and applications to support their research, things can get complicated in a hurry.

At Clemson University, IT administrators manage access to the school's Palmetto HPC cluster of more than 2,000 NVIDIA GPU-accelerated nodes including the latest Tesla V100s. Scientists working on a variety of research projects frequently request time on the cluster to conduct their most demanding simulations across molecular dynamics, quantum chemistry, computational fluid dynamics, and deep learning.

"They want to solve their problems, and they're too big to solve on laptops or workstations," said Ashwin Srinath, a research facilitator in Clemson's IT department. "Most of our researchers don't really want to get very much into how HPC works. They just want to run their program on a really big machine."

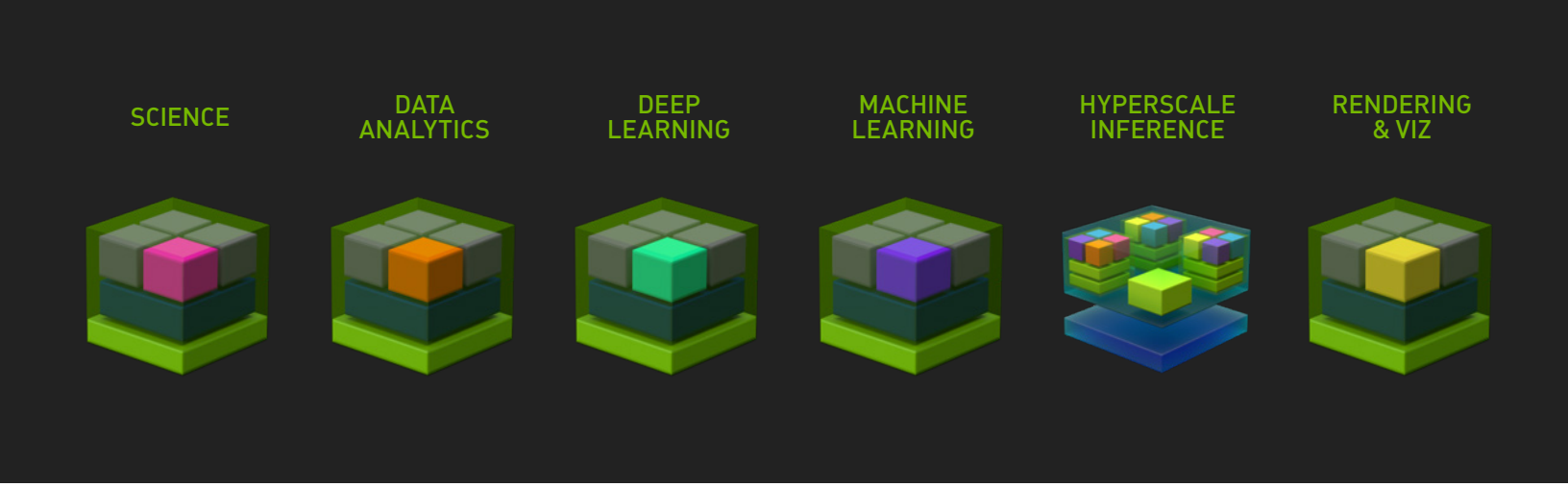
But striving to provide that kind of service to HPC resources is challenging. Installing or updating applications on bare metal in an already complex network of environment modules bogs down IT and delays application access, in turn, slowing down research. More often than not, updating the dependencies inadvertently breaks the dependency networks, interrupting other applications. And with so many different types of users and projects to support, HPC administrators spend inordinate amounts of time fulfilling new install and app upgrade requests for a huge variety of applications.

Things are complicated further by the fact that some users may want to run simulations on older versions of applications while other users want access to the latest features. Having to install, update and maintain multiple versions of the application is a secondary priority that can drain researcher productivity.

CUSTOMER PROFILE

CLEMSON COMPUTING AND INFORMATION TECHNOLOGY

Organization	HPC System	Industry	Location	Organization Size	Website
Clemson University – HPC Center IT Department	Palmetto cluster powered by NVIDIA® Tesla® V100 GPUs	Education	Clemson, South Carolina	19,402 undergraduate students, 4,985 graduate students, 1,431 academic staff	www.clemson.edu



Software containers on NGC

PRODUCTS

NVIDIA GPU Cloud:
- Deep Learning
  TensorFlow, PyTorch, and more
- HPC
  NAMD, GROMACS, and more
- HPC Visualization
  ParaView, VMD, and more
- Machine Learning
  RAPIDS, MATLAB

NVIDIA Tesla V100

NVIDIA Tesla P100

Perhaps even more importantly, this added IT workload impedes application users' productivity by forcing them to wait longer to run simulations. They also sometimes find themselves having no choice but to run simulations on older, feature-limited versions of applications, resulting in limited insights and sub-optimal results.

NGC CONTAINERS BOOST HPC CLUSTER PROVISIONING

Srinath and the Clemson HPC admin team needed a solution that delivered a seamless HPC experience, and maximized the efficiency of IT and researchers alike. It should also simplify their work and enable users to conduct research more quickly. They found one: software containers in the NGC container registry.

GPU-optimized application containers remove much of the grunt work for HPC administrators. Applications and all of their relevant dependencies are packaged in self-contained environments that are agnostic to any underlying hardware or other installed software. As such, containers eliminate the installation process and application deployments can be completed without impacting other applications on the cluster.

For instance, a container would enable an HPC administrator to upgrade an application such as NAMD without having it interfere with another, say, VMD. And when one user wants to use the latest version of NAMD while another user wants to continue working with the older version, both versions can run in parallel in their own containers, giving researchers flexibility in how they approach their simulations.

The deep learning containers on NGC enable users to run their favorite deep learning frameworks on HPC clusters without relying on the system admins. This self-service aspect further simplifies IT tasks and increases scientists' productivity.

"Containers from NVIDIA GPU Cloud have automated application deployment, making users self-sufficient while allowing us to focus on other critical priorities. With the adoption of NGC, we're able to help researchers with real problems."

Ashwin Srinath,
Research Facilitator,
Clemson University IT Department

Another important benefit containers bring to HPC is reproducibility, which is a key consideration for researchers but is very hard to achieve in bare-metal settings. Because applications running in a container always use the same environment, their performance becomes reproducible and portable, which enables users to share their research and build upon each other's work.

DEPLOYING NGC CONTAINERS ON PALMETTO IS EFFORTLESS

Clemson has seen all of these benefits, but perhaps the most important has been the ease with which Srinath and the HPC admin team are able to deploy containers on their Palmetto cluster.

By running NGC containers on Singularity, which is designed for the HPC environment, Srinath simply points users toward step-by-step instructions on downloading containers from NVIDIA GPU Cloud and running the container on Singularity.

NGC containers are also enabling Srinath and his team to respond more quickly when users clamor for updates to frequent releases of software like TensorFlow. "One of the favorite things about NGC containers besides the fact that NGC containers work in Singularity, is the availability of the latest version of the application," said Srinath.

It's true, NVIDIA works directly with the developers to host the latest versions of their applications. On the deep learning side, NVIDIA releases a new version of the popular frameworks every month so that user can easily get access to the latest features.

CONTAINERS BENEFIT ADMINS AND USERS ALIKE

It doesn't take long for organizations to experience the benefits of NGC containers, and Clemson immediately realized numerous payoffs, starting with enabling IT to greatly reduce housekeeping chores.

"By spending less time on software installs, we're definitely able to take care of things that have more priority," said Srinath. "We're able to help researchers with real problems—for instance, improving workflows or parallelizing code."

Instead of spending a half-day or more installing applications such as GROMACS, which simulates proteins, lipids and nucleic acids for molecular dynamics research, the users are now able to deploy the application in minutes, or the time it takes to download the image and run it on the cluster. Srinath said that ratio — reducing half-day tasks to minutes — is representative of how long it takes to install most apps on bare metal versus with NGC containers.

That said, some applications provide even more dramatic examples.

"I remember spending at least two weeks getting Caffe installed the first time I tried to install it on the cluster," said Srinath. And with an NGC container? "A few minutes," he said.

On the user side, researchers are enjoying faster access to the programs they need to run their experiments, improved reusability of the resulting work, and unprecedented ability to share their project environments with collaborators.

Clemson's experience underscores the diverse benefits of NGC containers. They're optimized to support the most recent versions of applications and to ensure maximum performance. They make it easy to put the latest features and capabilities in the hands of researchers who want them, but still support those who want to continue using older versions of apps. And they enable IT to deliver all of this relatively effortlessly.

In short, NGC containers bring new levels of agility, responsiveness, and usability to the provisioning of HPC resources. In doing so, they enable HPC administrators and the researchers they support to work more efficiently, effectively, and ultimately, deliver greater impact.

NGC containers support HPC applications, HPC visualization, deep learning and machine learning and they are available free of charge. For more information on how NGC containers can accelerate your HPC research activities, visit **www.nvidia.com/en-us/gpu-cloud/.**

www.nvidia.com

NVIDIA