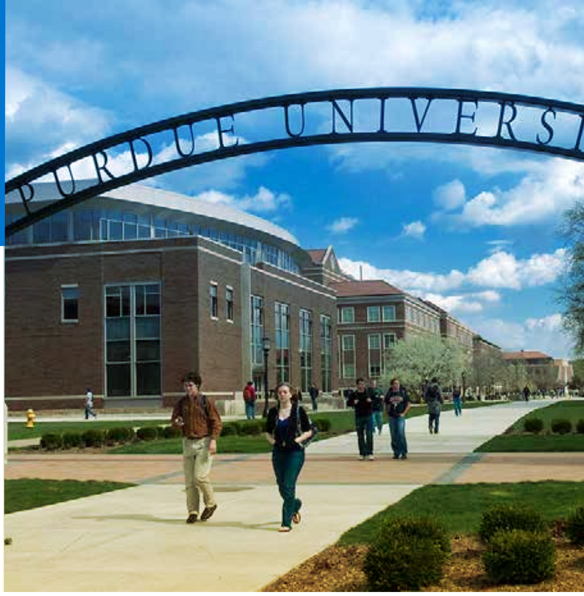
DELL EMC intel



Community clusters

University arms its researchers with leading-edge HPC systems to drive discovery and innovation.

PURDUE UNIVERSITY

High-Performance Computing | United States

Business needs

Researchers at Purdue University need access to high-performance computing clusters for compute- and data-intensive investigations.

Solutions at a glance

- Dell EMC PowerEdge servers
- Intel® Xeon® Scalable Processors
- Lustre storage
- InfiniBand and Ethernet networking

Business results

- Making leading-edge HPC resources available to diverse communities
- Accelerating throughput for compute- and data-intensive applications

The university's Brown cluster has

550
cores

The Brown cluster debuted at No. 302 on the

TOP500
list

A service-driven mission

As a university founded in 1869 under the Morrill Land Grant Act, Purdue University has had a public-service mission from its earliest days. The Act, signed by President Abraham Lincoln in 1862, turned public lands over to states that wanted to launch colleges dedicated to teaching agriculture and the "mechanic arts."

In the century and half that followed, Purdue grew into a top-tier research institution known as a national and global leader in discovery and innovation. Today, thanks to the breadth and depth of its research program, Purdue continues to drive advances in fields ranging from engineering and artificial intelligence to sustainability and health.

To enable these world-class research programs, Purdue and its faculty researchers invest in leading-edge high-performance computing systems that drive compute- and data-intensive scientific investigations. And this is where Information Technology at Purdue (ITaP) Research Computing organization enters the picture.

ITaP Research Computing provides advanced computational resources and services to support Purdue faculty, staff and student researchers. As part of this mission, the organization operates a suite of top-ranked cooperative supercomputing clusters. It also provides the Purdue research community with expert assistance and an array of tools and services for intensive computation, modeling, simulation, data management and analysis, data visualization, virtual collaboration and more.

Community clusters

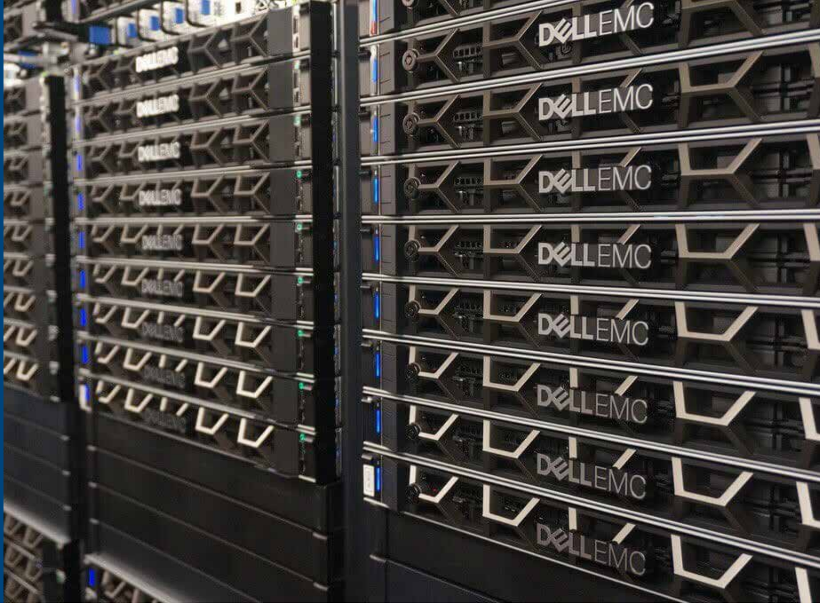
ITaP Research Computing delivers high performance computing resources via the university's Community Cluster Program. Through this program, ITaP funds and operates HPC clusters in partnership with faculty researchers. Faculty members contribute funding to provide the nodes and central IT contributes storage, network, professional administration and support staff to make these resources accessible across the campus.

Community Cluster systems are built with versatile Dell EMC PowerEdge servers with Intel Xeon Scalable processors, optimized for HPC and AI workloads, which allow users to manage large, complex datasets, gain insights more quickly, accelerate product innovation and drive scientific explorations that weren't previously possible. Lustre storage, InfiniBand and Ethernet networking were also included in the solution. The clusters are integrated into one large supercomputer with the right technologies to solve some of the biggest problems in science, engineering, physical sciences and life sciences, according to Preston Smith, director of research computing services and support at Purdue.

> *"Our Community Cluster model builds an HPC system every year, so our faculty members always know that they are going to have access to a cutting-edge supercomputer, either just deployed or about to be deployed, to accomplish their research."*
>
> Preston Smith
> Director of Research Computing Services
> and Support, Purdue University

Dell EMC systems at Purdue

In any given year, Purdue operates up to five flagship Community Cluster systems. Currently, two of these systems are based on Dell EMC solutions — clusters known as Brown and Gilbreth.



Brown cluster

The Brown cluster, built through a partnership with Dell EMC and Intel, is optimized for researchers running traditional, tightly coupled science and engineering applications. When it went into service in the fall of 2017, Brown ranked at No. 302 on the TOP500 list of the world's largest supercomputers.

The core components of Brown include:

- 550 Dell EMC PowerEdge compute nodes with two Intel Xeon Scalable Processors, 6146, 12-core, 4.2 GHz (24 cores per node) and 96 GB of memory
- 100 Gbps EDR InfiniBand interconnect on all nodes

Gilbreth cluster

The Gilbreth cluster, in service in the first half of 2019, is optimized for machine learning, artificial intelligence and other GPU-based applications.

The core components of Gilbreth include:

- 50 Dell EMC PowerEdge compute nodes featuring the Intel® Xeon® Scalable Processors, 6142, 16-core, with 192 GB of RAM and 100 Gbps EDR InfiniBand interconnects
- Three model-training nodes and 8 TB of local flash storage
- Nodes from two existing GPU clusters operated by ITaP Research Computing
- A multi-petabyte Lustre parallel file system and a shared flash-based storage system for processing large datasets

Realizing the benefits

Researchers across the Purdue community put the Community Clusters to work to drive discovery in virtually all academic disciplines.

"Purdue University has faculty from every college who need high performance computing to advance their research programs," Smith says. "As our organization builds systems for artificial intelligence and machine learning workloads, we are seeing researchers coming from domains that we have not previously served. For example, people solving problems in humanities with text analysis, or people doing image processing — they need larger systems with more storage and larger amounts of data to get results faster."

All told, Purdue has about 1,300 researchers using the HPC systems, including graduate students, undergraduate students and faculty members.

"With the power of Community Clusters, researchers at Purdue are taking giant leaps in discovery in engineering, artificial intelligence, sustainability, and health and longevity," Smith says. "We have researchers advancing new fuels, discovering new ways to sustain and feed the world, and finding cures to disease."

In one such research breakthrough, announced in 2018, Purdue researchers used cryo-electron microscopy (cryo-EM) and HPC systems to create the most accurate picture of the deadly Zika virus to date.

This high-resolution picture identified probable drug-binding pockets on the surface of the virus, potentially paving the way for the design of a vaccine.

This discovery built on the findings of earlier scientific investigations conducted at Purdue, in which researchers were the first to discover the structure of the Zika virus. In this earlier work, researchers used cryo-EM to create an image of the virus at a resolution of 3.8 Ångstrom, a unit used to express the size of atoms and molecules. And now they've scaled that number down to 3.1, bringing them even closer to the atomic details of the virus.

In another initiative, using the Gilbreth cluster, agriculture researchers plan to fly drones over Indiana fields to gather images of crops. They will then offload their data onto an edge computer out on the farm and use a high-speed network to transfer the data to a central location, then load the data into Gilbreth's parallel file system and let the cluster process the images and run analytics.

Ultimately, this digitally-driven agricultural research will yield valuable insight into the health of crops and their potential yields. These insights can help producers cut costs and boost yields — two all-important goals in a time when a growing population and a changing climate are posing severe challenges to the world's food supply.

Working with Dell EMC

Purdue's ITaP Research Computing organization has a long-running relationship with Dell, and now Dell EMC. Since 2004, six of the organization's Community Cluster systems have been built by Dell or Dell EMC. In building these systems, ITaP Research Computing worked closely with Dell EMC engineers to optimize designs to yield more performance per dollar.

"One of the most important things that we have to consider when we build a supercomputer for our campus is to do something in a cost-effective manner," Smith says. "Our model is built around selling access to the node hardware to the individual faculty, so we have to optimize our designs to be cost-effective so the researchers can afford the computing resources that they need."

In addition, Dell EMC has always been there with the service and support that an organization needs fo have a successful experience with technology, Smith notes.

"We have been very pleased with the service we have gotten from Dell EMC," he says. "Delivery is always quick. And working with the support organization on the hardware has been very satisfactory."