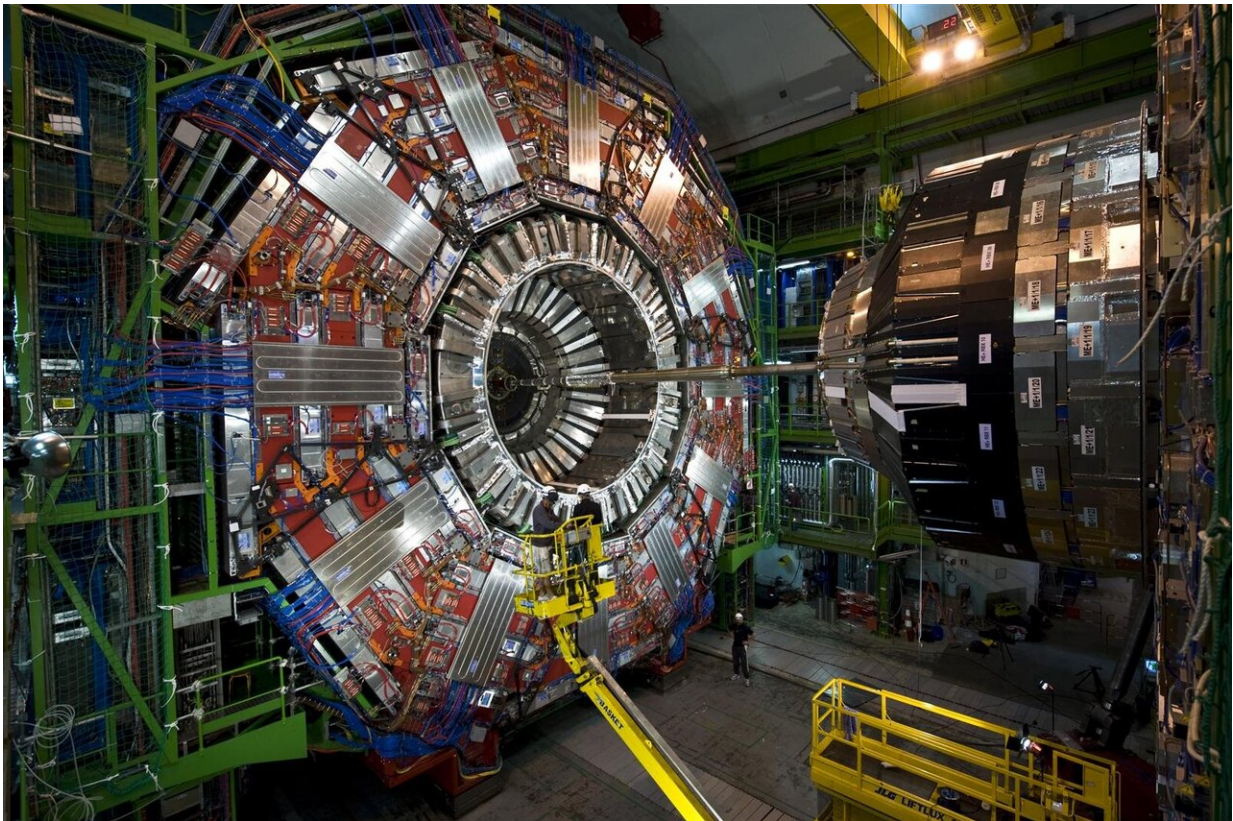


# A glimpse into the future: Accelerated computing for accelerated particles

August 16 2019, by Leah Hesla

---



Particles emerging from proton collisions at CERN's Large Hadron Collider travel through through this stories-high, many-layered instrument, the CMS detector. In 2026, the LHC will produce 20 times the data it does currently, and CMS is currently undergoing upgrades to read and process the data deluge. Credit: Maximilien Brice, CERN

Every proton collision at the Large Hadron Collider is different, but only a few are special. The special collisions generate particles in unusual patterns—possible manifestations of new, rule-breaking physics—or help fill in our incomplete picture of the universe.

Finding these collisions is harder than the proverbial search for the needle in the haystack. But game-changing help is on the way. Fermilab scientists and other collaborators successfully tested a prototype [machine-learning](#) technology that speeds up processing by 30 to 175 times compared to traditional methods.

Confronting 40 million collisions every second, scientists at the LHC use powerful, nimble computers to pluck the gems—whether it's a Higgs particle or hints of dark matter—from the vast static of ordinary collisions.

Rifling through simulated LHC collision data, the machine learning technology successfully learned to identify a particular postcollision pattern—a particular spray of particles flying through a detector—as it flipped through an astonishing 600 images per second. Traditional methods process less than one image per second.

The technology could even be offered as a service on external computers. Using this offloading model would allow researchers to analyze more data more quickly and leave more LHC computing space available to do other work.

It is a promising glimpse into how machine learning services are supporting a field in which already enormous amounts of data are only going to get bigger.

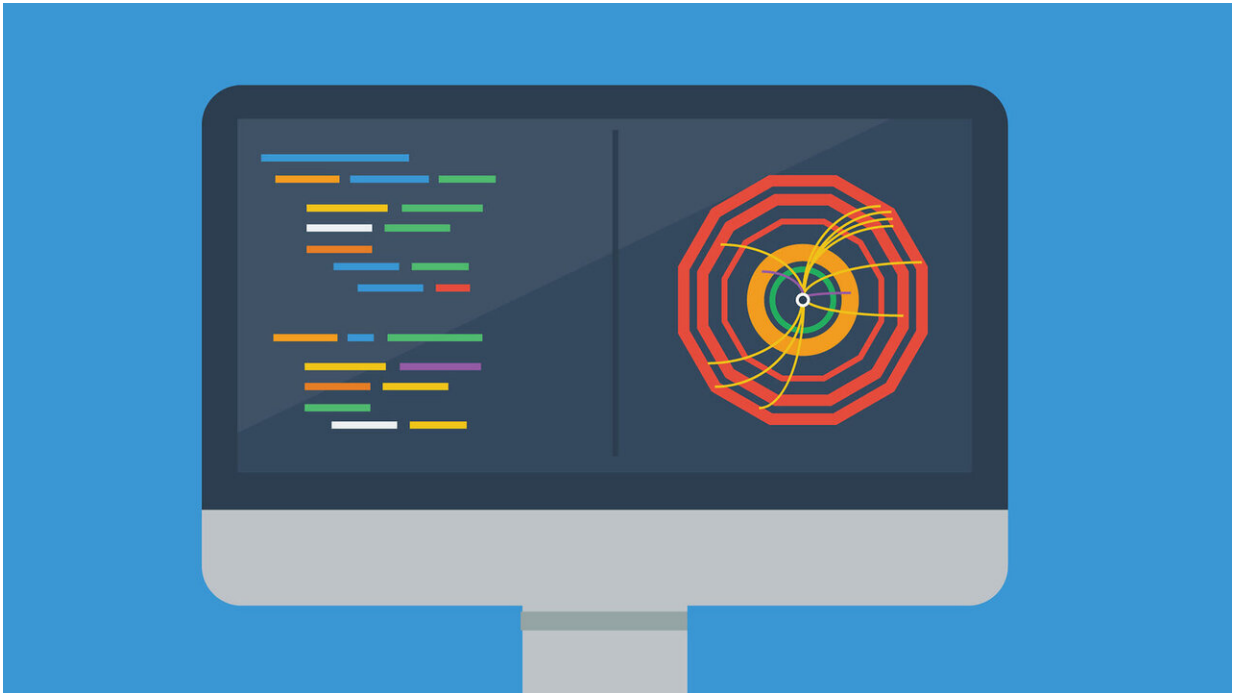
## **The challenge: more data, more computing power**

Researchers are currently upgrading the LHC to smash protons at five times its current rate. By 2026, the 17-mile circular underground machine at the European laboratory CERN will produce 20 times more data than it does now.

CMS is one of the particle detectors at the Large Hadron Collider, and CMS collaborators are in the midst of some upgrades of their own, enabling the intricate, stories-high instrument to take more sophisticated pictures of the LHC's particle collisions. Fermilab is the lead U.S. laboratory for the CMS experiment.

If LHC scientists wanted to save all the raw collision data they'd collect in a year from the High-Luminosity LHC, they'd have to find a way to store about 1 exabyte (about 1 trillion personal external hard drives), of which only a sliver may unveil new phenomena. LHC computers are programmed to select this tiny fraction, making split-second decisions about which data is valuable enough to be sent downstream for further study.

Currently, the LHC's computing system keeps roughly one in every 100,000 particle events. But current storage protocols won't be able to keep up with the future data flood, which will accumulate over decades of data taking. And the higher-resolution pictures captured by the upgraded CMS detector won't make the job any easier. It all translates into a need for more than 10 times the computing resources than the LHC has now.



Particle physicists are exploring the use of computers with machine learning capabilities for processing images of particle collisions at CMS, teaching them to rapidly identify various collision patterns. Credit: Eamonn Maguire/Antarctic Design

The recent prototype test shows that, with advances in machine learning and computing hardware, researchers expect to be able to winnow the data emerging from the upcoming High-Luminosity LHC when it comes online.

"The hope here is that you can do very sophisticated things with machine learning and also do them faster," said Nhan Tran, a Fermilab scientist on the CMS experiment and one of the leads on the recent test. "This is important, since our data will get more and more complex with upgraded detectors and busier collision environments."

## Machine learning to the rescue: the inference difference

Machine learning in particle physics isn't new. Physicists use machine learning for every stage of data processing in a collider experiment.

But with [machine learning technology](#) that can chew through LHC data up to 175 times faster than traditional methods, particle physicists are ascending a game-changing step on the collision-computation course.

The rapid rates are thanks to cleverly engineered hardware in the platform, Microsoft's Azure ML, which speeds up a process called inference.

To understand inference, consider an algorithm that's been trained to recognize the image of a motorcycle: The object has two wheels and two handles that are attached to a larger metal body. The algorithm is smart enough to know that a wheelbarrow, which has similar attributes, is not a motorcycle. As the system scans new images of other two-wheeled, two-handled objects, it predicts—or infers—which are motorcycles. And as the algorithm's prediction errors are corrected, it becomes pretty deft at identifying them. A billion scans later, it's on its inference game.

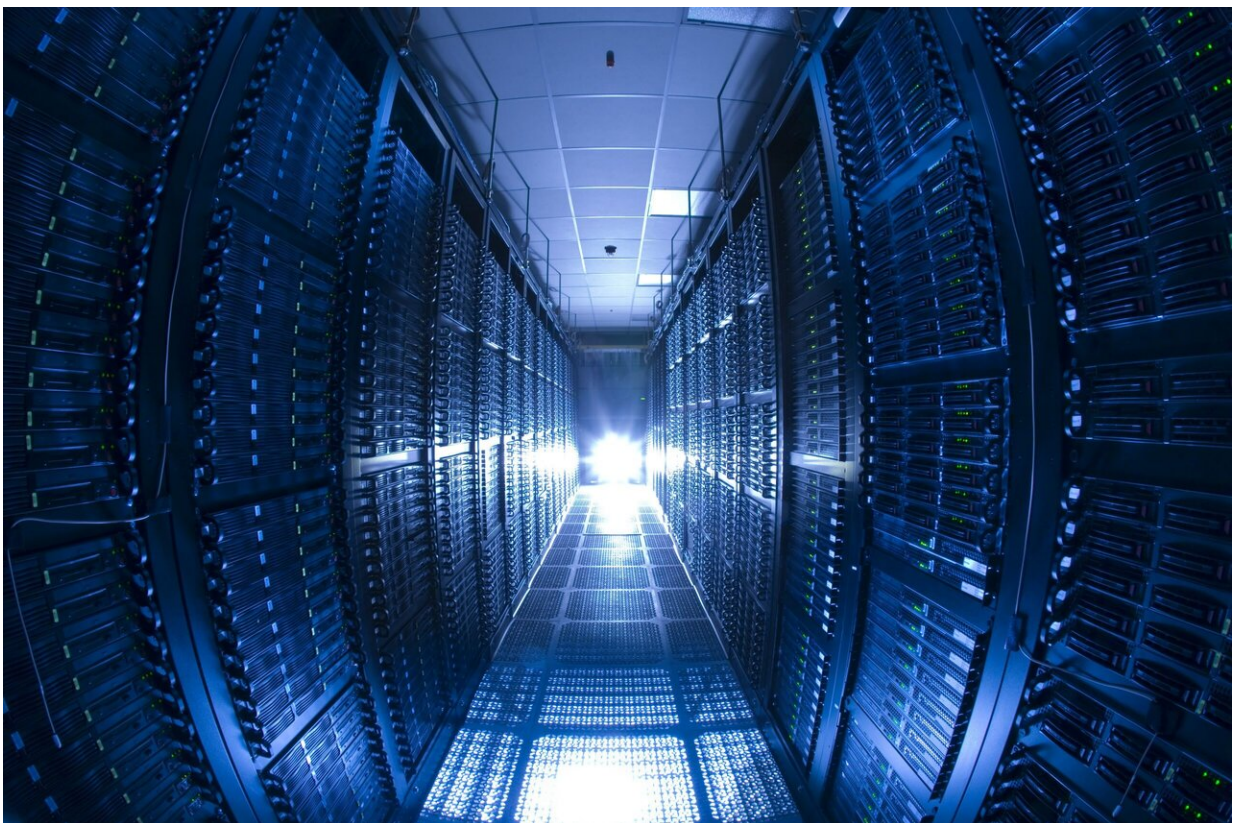
Most machine learning platforms are built to understand how to classify images, but not physics-specific images. Physicists have to teach them the physics part, such as recognizing tracks created by the Higgs boson or searching for hints of dark matter.

Researchers at Fermilab, CERN, MIT, the University of Washington and other collaborators trained Azure ML to identify pictures of top quarks—a short-lived elementary particle that is about 180 times heavier than a proton—from simulated CMS data. Specifically, Azure was to

look for images of top quark jets, clouds of particles pulled out of the vacuum by a single top quark zinging away from the collision.

"We sent it the images, training it on physics data," said Fermilab scientist Burt Holzman, a lead on the project. "And it exhibited state-of-the-art performance. It was very fast. That means we can pipeline a large number of these things. In general, these techniques are pretty good."

One of the techniques behind inference acceleration is to combine traditional with specialized processors, a marriage known as heterogeneous computing architecture.



Data from particle physics experiments are stored on computing farms like this one, the Grid Computing Center at Fermilab. Outside organizations offer their computing farms as a service to particle physics experiments, making more

space available on the experiments' servers. Credit: Reidar Hahn

Different platforms use different architectures. The traditional processors are CPUs (central processing units). The best known specialized processors are GPUs (graphics processing units) and FPGAs (field programmable gate arrays). Azure ML combines CPUs and FPGAs.

"The reason that these processes need to be accelerated is that these are big computations. You're talking about 25 billion operations," Tran said. "Fitting that onto an FPGA, mapping that on, and doing it in a reasonable amount of time is a real achievement."

And it's starting to be offered as a service, too. The test was the first time anyone has demonstrated how this kind of heterogeneous, as-a-service architecture can be used for fundamental physics.

In the computing world, using something "as a service" has a specific meaning. An outside organization provides resources—machine learning or hardware—as a service, and users—scientists—draw on those resources when needed. It's similar to how your video streaming company provides hours of binge-watching TV as a service. You don't need to own your own DVDs and DVD player. You use their library and interface instead.

Data from the Large Hadron Collider is typically stored and processed on computer servers at CERN and partner institutions such as Fermilab. With machine learning offered up as easily as any other web service might be, intensive computations can be carried out anywhere the service is offered—including off site. This bolsters the labs' capabilities with additional computing power and resources while sparing them from

having to furnish their own servers.

"The idea of doing accelerated computing has been around decades, but the traditional model was to buy a computer cluster with GPUs and install it locally at the lab," Holzman said. "The idea of offloading the work to a farm off site with specialized hardware, providing machine learning as a service—that worked as advertised."

The Azure ML farm is in Virginia. It takes only 100 milliseconds for computers at Fermilab near Chicago, Illinois, to send an image of a particle event to the Azure cloud, process it, and return it. That's a 2,500-kilometer, data-dense trip in the blink of an eye.

"The plumbing that goes with all of that is another achievement," Tran said. "The concept of abstracting that data as a thing you just send somewhere else, and it just comes back, was the most pleasantly surprising thing about this project. We don't have to replace everything in our own computing center with a whole bunch of new stuff. We keep all of it, send the hard computations off and get it to come back later."

Scientists look forward to scaling the technology to tackle other big-data challenges at the LHC. They also plan to test other platforms, such as Amazon AWS, Google Cloud and IBM Cloud, as they explore what else can be accomplished through machine learning, which has seen rapid evolution over the past few years.

"The models that were state-of-the-art for 2015 are standard today," Tran said.

As a tool, machine learning continues to give particle physics new ways of glimpsing the universe. It's also impressive in its own right.

"That we can take something that's trained to discriminate between



pictures of animals and people, do some modest amount computation, and have it tell me the difference between a top quark jet and background?" Holzman said. "That's something that blows my mind."

Provided by Fermi National Accelerator Laboratory

Citation: A glimpse into the future: Accelerated computing for accelerated particles (2019, August 16) retrieved 20 April 2024 from <https://phys.org/news/2019-08-glimpse-future-particles.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.