

High throughput computing helps LIGO confirm Einstein's last unproven theory

March 9 2016, by Brian Mattmiller



A few years ago, a global team of scientists parlayed decades of research into the discovery of the Higgs boson, a subatomic particle considered a building block of the universe. A humble software program called HTCondor churned away in the background, helping analyze data gathered from billions of particle collisions.

Cut to 2016, and HTCondor is on to a new collision: helping scientists detect gravitational waves caused 1.3 billion years ago by a collision between two black holes 30 times as massive as our sun.

From revealing the Higgs boson, among the smallest particles known to science, to detecting the impossibly massive astrophysics of black holes, the [HTCondor High Throughput Computing](#) (HTC) software has proven indispensable to processing the vast and complex data produced by big international science. Computer scientists at the University of Wisconsin—Madison pioneered these distributed high throughput [computing technologies](#) over the last three decades.

The announcement in February that scientists from the Laser Interferometer Gravitational-Wave Observatory (LIGO) unlocked the final door to Albert Einstein's Theory of Relativity—proof that gravitational waves produce ripples through space and time—has a rich back story involving HTCondor. Since 2004, HTCondor has been a core part of the [data analysis](#) effort of the project that includes more than 1,000 scientists from 80 institutions across 15 countries.

By the numbers, more than 700 LIGO scientists have used HTCondor over the past 12 years to run complex data analysis workflows on computing resources scattered throughout the U.S. and Europe. About 50 million core-hours managed by HTCondor in the past six months alone supported the data analysis that led to the detection reported in the February 2016 papers.

The path of HTCondor to LIGO was paved by a collaboration that started more than a decade ago between two University of Wisconsin teams—the 29-member LIGO team at UW-Milwaukee and the HTCondor team at UW–Madison. This interdisciplinary collaboration helped transform the computational model of LIGO and advance the state of the art of HTC.

The HTCondor team is led by Miron Livny, UW–Madison professor of computer sciences and chief technology officer for the Morgridge Institute for Research and the Wisconsin Institute for Discovery. The HTCondor software implements innovative high-throughput computing technologies that harness the power of tens of thousands of networked computers to run large ensembles of computational tasks.

"What we have is the expertise of two UW System schools coming together to tackle a complex data analysis problem," says Thomas Downes, a UW-Milwaukee senior scientist of physics and LIGO investigator. "The problem was, how do you manage thousands upon thousands of interrelated data analysis jobs in a way that scientists can effectively use? And it was much more easily solved because Milwaukee and Madison are right down the street from each other."

The UW-Milwaukee team began using HTCondor in the early 2000s as part of an NSF Information Technology Research (ITR) project. Its then-lead scientist, Bruce Allen, landed a job as director of the Albert Einstein Institute for Gravitational Physics in Hannover, Germany, one of the leading hubs in the LIGO project. Duncan Brown, then a UW-Milwaukee physics Ph.D. candidate, became a professor of physics at Syracuse University, leading that university's LIGO efforts.

Allen, Brown and others worked hard at Milwaukee to demonstrate the value of the HTCondor approach to the mission of LIGO, eventually leading to its adoption at other LIGO sites. HTCondor soon became the go-to technology for the core LIGO data groups at UW-Milwaukee, Syracuse University, the Albert Einstein Institute, the California Institute of Technology (Caltech), and Cardiff University in the UK.

Peter Couvares has a 360-degree view of HTCondor's relationship to LIGO. He worked on the HTCondor team for 10 years at UW–Madison, and managed the relationship between LIGO and HTCondor for about

five years after joining the LIGO team led by Brown at Syracuse. Today he is a senior scientist at Caltech managing the LIGO data analysis computing team.

Why is the HTCondor software such a boon to big science endeavors like LIGO?

"We know it will work—that's the killer feature of HTCondor," says Couvares.

It works, he adds, because HTCondor takes seriously the core challenge of distributed computing: It's impossible to assume that a network of thousands of individual computers will not have local failures. HTCondor bakes that assumption into its core software.

"The HTCondor team always asks people to think ahead to the issues that are going to come up in real production environments, and they're good about not letting HTCondor users take shortcuts or make bad assumptions," Couvares adds.

In a project like LIGO, that approach is especially important. The steady stream of data from LIGO detectors is a mix of gravitational information and noise such as seismic activity, wind, temperature and light—all of which helps define and differentiate both good and bad data.

"In the absence of noise, this would have been a very easy search," Couvares says. "But the trick is in picking a needle out of a haystack of noise. The biggest trick of all the data analysis in LIGO is to come up with a better signal-to-noise ratio."

Stuart Anderson, LIGO senior staff scientist at Caltech, has been supporting the use of HTCondor within LIGO for more than a decade.

The reason HTCondor succeeds is less about technology than it is about the human element, he says.

"The HTCondor team provides a level of long-term collaboration and support in cyber-infrastructure that I have not seen anywhere else," he says. "The team has provided the highest quality of technical expertise, communication skills and collaborative problem-solving that I have had the privilege of working with."

Adds Todd Tannenbaum, the current HTCondor technical lead who works closely with Anderson and Couvares: "Our relationship with LIGO is mutually profitable. The improvements made on behalf of our relationship with LIGO have greatly benefited HTCondor and the wider high throughput computing community."

Ewa Deelman, research associate professor and research director at the University of Southern California Information Sciences Institute (ISI), became involved with HTCondor in 2001 as she launched Pegasus, a system that automates the workflow for scientists using systems like HTCondor. Together, Pegasus and HTCondor help lower the technological barriers for scientists.

"I think the automation and the reliability provided by Pegasus and HTCondor are key to enabling scientists to focus on their science, rather than the details of the underlying cyber-infrastructure and its inevitable failures," she says.

The future of LIGO is tremendously exciting, and the underpinning high-throughput computing technologies of HTCondor will change with the science and the computing technologies. Most agree that with the initial observation of gravitational waves, a potential tsunami of data awaits from the universe.

"The field of gravitational wave astronomy has just begun," says Couvares. "This was a physics and engineering experiment. Now it's astronomy, where we're seeing things. For 20 years, LIGO was trying to find a needle in a haystack. Now we're going to build a needle detection factory."

Adds Livny: "What started 15 years ago as a local Madison-Milwaukee collaboration turned into a computational framework for a new field of astronomy. We are ready and eager to address the evolving HTC needs of this new field. By collaborating with scientists from other international efforts that use technologies ranging from a neutrino detector in the South Pole (IceCube) to a telescope floating in space (Hubble) to collect data about our universe, HTC will continue to support scientific discovery."

Provided by University of Wisconsin-Madison

Citation: High throughput computing helps LIGO confirm Einstein's last unproven theory (2016, March 9) retrieved 7 May 2024 from <https://phys.org/news/2016-03-high-throughput-ligo-einstein-unproven.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.