Lincoln Laboratory Supercomputing Center

5th Anniversary

Lincoln Laboratory Supercomputing Center

The Lincoln Laboratory Supercomputing Center (LLSC) celebrated a significant anniversary this year, marking five years of the center's mission to enhance the computing power available to the Laboratory, MIT, and other researchers.

The LLSC was established on 1 April 2016 to build on the Laboratory's past groundbreaking work in computing. In the late 1940s and early 1950s, prior to the establishment of the Laboratory, MIT designed and built the Whirlwind 1 computer using vacuum tubes for the U.S. Navy. Whirlwind was unique for its ability to output information in real time and perform calculations in parallel. Later, in 1955, the first fully transistor-based computer TX-0 was designed by staff at the now-established Lincoln Laboratory.

These and other early computing accomplishments eventually led to the establishment in 2003 of LLGrid, a Laboratory supercomputing system. In 2008, the Laboratory demonstrated the largest problem ever run on a computer using the TX-2500 supercomputer, a computer that was a part of LLGrid.

The LLSC was established with a focus on interactive supercomputing and high-performance data analysis. The staff who started the LLSC decided to model the center on the partnership model found at federal laboratories and universities around the world.

The LLSC recently deployed TX-GAIA in 2019 and was rated by Top500.org as the 42nd fastest supercomputer in the world. It has served a critical role in AI research, including research conducted by the MIT-Air Force AI Accelerator and research related to COVID-19.

The LLSC connects to a data center situated 90 miles away in Holyoke, a Massachusetts city on the Connecticut River. This data center, the Massachusetts Green High Performance Computing Center, is a computing infrastructure available to researchers from a consortium of technology industries and Massachusetts schools, including MIT, Harvard University, the University of Massachusetts, Boston University, and Northeastern University. The center is powered by a combination of green energy sources.

LLSC CELEBRATES 5TH ANNIVERSARY

It is an honor to be a part of the long tradition of innovative computing at Lincoln Laboratory. The LLSC is privileged to work with the best researchers in the world to address the Nations' most challenging problems while advancing human knowledge. The existence of the LLSC is a result of the vision of Lincoln Laboratory leadership and the LLSC mission remains what it was at inception:

- Address supercomputing needs across all Laboratory missions
- Develop new supercomputing capabilities and technologies
- Maintain close collaborations with MIT supercomputing initiatives

This 5th anniversary compendium of articles written by the dedicated staff at the Lincoln Laboratory Bulletin, MIT News, and other writers is a testament to the continuing impact and potential of supercomputing to advance science and engineering. Finally, on a personal note, I would like to express my profound gratitude to my LLSC colleagues; your commitment to each other and to our team is truly special.

Jeremy Kepner
Laboratory Fellow

CONTENTS

CONTENTS (Continued)

Supporting Research

The LLSC is an interactive, on-demand parallel computing system that uses large computing clusters to enable Laboratory researchers to augment the power of desktop systems to process large sets of sensor data, create high-fidelity simulations, and develop entirely new algorithms.

LLSC develops and deploys unique, energy-efficient supercomputing that:

- Provides data centers, hardware, software, user support, and pioneering research
- Is 100x more productive than standard supercomputing[1]
- Delivers 100x more performance than standard cloud[2]

[1]Interactive Grid Computing at Lincoln Laboratory, Bliss et al., LL Journal, 2006

[2]Scalability! But at what COST?, McSherry et al., HotOS XV, 2015

Hydra-zen Framework Makes Scientific Computing Easier for Researchers

Director's Office

MIT Lincoln Laboratory

August 13, 2021

Hydra-zen, an open-source framework developed at the Laboratory, aims to make scientific computing, like the computing done with the Laboratory's supercomputing resources, easier by simplifying the configuration and documentation process for experiments.

Using computers to perform complex experiments has become a crucial part of research across the world. To verify these experiments, and to perform the same experiments again, researchers must keep track of the many variables and settings that are being used to run them.

To document these variables and settings, many researchers have relied on hand-taken notes in something like a lab notebook. However, with the number of variables and settings involved and other circumstances like someone being unavailable, this can cause issues in documentation reliability and has the potential to taint or even ruin the experiments being performed.

To solve this issue, a team at the Laboratory has developed Hydra-zen, a framework based on Python that allows researchers to better document and implement their changes to variables and settings and save them alongside the resulting calculations, ensuring that the experiment can be duplicated. Hydra-zen aims to simplify and automate the scientific computing process.

"Using Hydra-zen, researchers can automatically and dynamically generate configurations for their experiments," said Dr. Ryan Soklaski, Artificial Intelligence Technology, Group 01. "It makes accessible, through automations, standardized bestpractices for configuring and running scientific computing experiments."

Hydra-zen is based on Facebook's open-source project Hydra, which is designed for configuring traditional software applications like websites. Hydra-zen builds upon Hydra to make it more suitable for these scientific computing projects, which are often more dynamic and require on-thefly adaptations as the experiment is conducted.

Development on Hydra-zen began as a way to speed up the configuration work required for another project. Once the team saw the massive improvements in their workflow and the potential for further improvements, they decided to make Hydra-zen their focus instead.

Researchers use Hydra-zen by generating a composable configuration for each part of their experiment's code. This enables the researchers to design their experiment in a modular way and to have that modularity be reflected in its configuration, meaning that modifying the experiment will not require a redesign of its configuration.

While executing an experiment, Hydra-zen automatically generates and saves a configuration file that is saved with the experiment results, making it easier to reproduce the results later. This is crucial for continuing experiments that may be older, when key factors of the experiment could be easily lost or forgotten. It is also important for verifying someone else's work.

"Over time, people forget how to run their own software-based experiments or interpret the results," said Soklaski. "Many Ph.D. theses are unsalvageable within a couple years of their completion because the author did not leave enough documentation to make the work reproducible. It is a really big problem in scientific computing research."

Hydra-zen is a part of the Laboratory's Responsible AI Initiative, which supports the development of AI technologies that are reliable and trustworthy. Hydra-zen is well-suited to this initiative because it can help researchers ensure they are producing reliable experiments that can be reproduced and verified. It is also a part of the Department of the Air Force-MIT AI Accelerator Program.

Soklaski hopes that Hydra-zen can have a big impact in machine learning work, and is working to improve the documentation to encourage new researchers to use it. The team behind Hydra-zen also hopes to present on the project at various conferences to receive feedback and suggestions.

"The machine learning field is astoundingly, dizzyingly active," said Soklaski. "Many researchers are constantly reinventing the wheel when it comes to configuring their code and experiments, and they typically don't have strong backgrounds in programming, so their solutions can be quite crude. Hydra-zen stands to help this field significantly by providing a simple tool for everyone to configure their code in a high-quality and elegant way."

As a part of an analysis on bias in facial recognition systems, Megan Frisella, a student intern from Artificial Intelligence Technology and Systems, Group 52, utilized Hydra-zen to perform comparison experiments that involved applying a face identification model to a dataset of face images. While normally a researcher would have to write a new script for each experiment's different models and datasets, Hydra-zen simplified this process by allowing them to easily specify which dataset, model, and pre-processing they wanted to use.

"I would recommend Hydra-zen to others. Its compatibility with LLGrid was very useful," said Dr. Pooya Khorrami, Group 52, who also worked on the facial recognition project. "I would especially recommend Hydrazen to machine learning researchers, as it provides a very elegant way to organize experiments and do hyperparameter tuning."

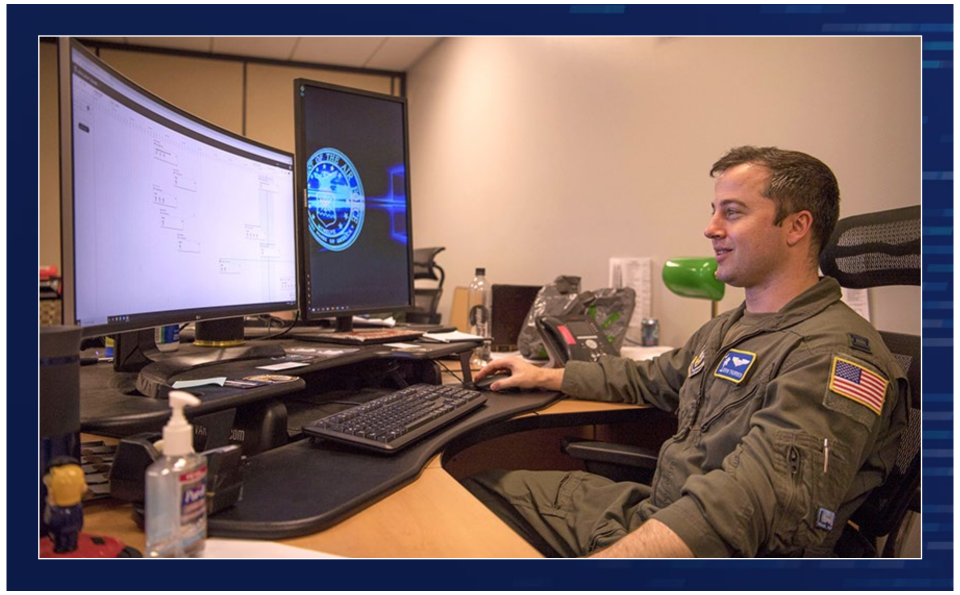
Hydra-zen is available as open-source software on GitHub. Documentation is also available.

Air Force Pilots Get an AI-Assist with Scheduling Aircrews

The Bulletin

July 16, 2021

From left to right: Capt. Ronisha Carter, Group 01, Dr. Amy Alexander, Group 43, Capt. Lindsey McEvoy, Group 43, Kendrick Cancio, Group 43, Michael Snyder, Group 41, Capt. Kyle McAlpin, Group 01, and 2nd Lt Matt Koch are pictured inside the cockpit of a C-17 aircraft at Hickam Air Force Base.



Captain Kevin Thurber of the 315th Airlift Wing at Joint Base Charleston, S.C., uses the Puckboard software program to schedule C-17 aircrews.

Take it from Captain Kyle McAlpin when he says that scheduling C-17 aircraft crews is complicated. An Artificial Intelligence (AI) Research Flight Commander for the Department of Air Force–MIT AI Accelerator Program, McAlpin is also an experienced C-17 pilot. “You could have a mission change and spend the next 12 hours of your life rebuilding a schedule that works,” he said.

It’s a pain point for Airmen of 52 squadrons who operate C-17s—the military cargo aircraft that transport troops and supplies globally. This year, the Air Force marked four million flight hours for its C-17 fleet, which comprises 275 U.S. and allied aircraft. Each flight requires scheduling a crew of six on average, though crew requirements vary depending on the mission.

“Being a scheduler is an additional duty on top of an Airman’s main job, such as being a pilot,” said Captain Ronisha Carter, who is a military liaison in Group 01, a Cyberspace Operations Officer, and the primary Airman on a research team spanning the Laboratory, the Department of the Air Force (DAF), and the MIT Department of Aeronautics and Astronautics. “What we want is for a scheduler to click a button, and an optimal schedule is created.”

Collaborating with their Air Force sponsor organization, Tron, the team has developed an AI–enabled plugin for the existing C-17 scheduling tool to fulfill that vision. The software plugin automates C-17 aircrew scheduling and optimizes crew resources. It was developed as part of the DAF–MIT AI Accelerator partnership.

Nearly 7,600 Airmen are poised to use the technology once it is rolled out this summer. It is being integrated into the scheduling software, called Puckboard, that C-17 Airmen currently use to build schedules two weeks in advance. Prior to Puckboard’s development in 2019, the squadrons had been using whiteboards and spreadsheets to manually plan out schedules. While Puckboard was a major improvement to the paper-and-pen, it didn’t have the “brains of optimization algorithms” to help schedulers avoid the mentally draining aspects of the task, said Michael Snyder, AI Software Architectures and Algorithms, Group 41, who is a software engineer and team lead.

The Airmen have many factors to consider as they build the schedules. When is air space available? Who is available to fly given rest requirements, deployments, and vacations? From that subset of available pilots, who is qualified? Some pilots, for example, may not be certified for night flying or air refueling. It’s also up to the scheduler to book training flights to keep pilots qualified in these areas.

“You have a lot of data and factors and the information is so spread out. It’s not something a human being can do in an efficient manner and the decisions they come to might not make the most efficient use of resources. That’s where AI plays a role,” said Hamsa Balakrishnan, who is the William E. Leonhard (1940) Professor of Aeronautics and Astronautics at MIT and principal investigator of the program.

The team’s approach to solving this scheduling problem fuses two techniques. The first is integer programming. In this approach, the algorithm solves an optimization problem by using binary (yes or no) decisions to decide whether or not to assign a pilot to an event. An optimal solution maximizes the values assigned to the desired characteristics of a “good” schedule. Examples of desired characteristics include increasing the rate at which pilots make progress towards satisfying their training requirements, or not unnecessarily assigning personnel to a flight who are significantly overqualified for the task at hand.

Candidate schedules with pilot assignments are then presented to an Airman (or an automatic agent), who can accept or reject a schedule. Each time a schedule is accepted, the algorithm is rewarded for its choices, which allows it to recognize successful patterns and improve its decisions over time. This process is called reinforcement learning.

Training the model has required feeding it a lot of historical C-17 aircrew and flight data. Accessing these data has been one of the greatest challenges, as old datasets were tossed out or housed in legacy systems that were hard to access and difficult to harmonize so that the model could pull from them all. “Once the data is connected, it’s then challenging to enumerate all of the constraints a scheduler considers,” said Matthew Koch, a graduate student in the MIT Operations Research Center.

For example, it’s straightforward to program the model around explicit constraints, such as the restrictions on how many hours a pilot can fly a day. Coding for implicit constraints is harder, or even impossible, and relies on the insight that an Airman brings to the desk, such as knowing that two pilots’ personalities don’t mesh or that the strengths of one pilot will complement another’s weaknesses to build the safest flight.

That’s where the research team’s relationship with the C-17 pilots has been essential. “There have been so many user interviews,” Carter said. In those interviews, the pilots and research team discuss the nuances of different scheduling outputs — what they liked and disliked, and what they would change about certain decisions that the algorithm made. “Every step of the way, it’s been a very integrated relationship and allows us to improve our algorithm,” she said.

By design, the technology is an assistant. It’s still up to the human to accept the schedule. This approach, the team hopes, will make the system trusted and accepted by users, some of whom have spent years building their own approach to the problem. “We’re figuring out what buttons or charts we can add to the interface so that our algorithms aren’t black boxes. We want to keep the scheduler in the loop,” said Snyder.

Showing fairness and equity in their algorithms is also important. “We want to enable a level of explainability of why someone was scheduled over someone else,” Koch added.

That goal is still aspirational. One technique to both improve the algorithm and to provide equity is to have the system present multiple schedules from which an Airman can choose. Understanding why a user chooses one over the other allows the researchers to tweak the model further.

Today, the team is continuing to integrate their plugin with Puckboard and explore how to measure its success. “It’s hard to say that there is one optimal solution; there could be several different, but very good schedules. It’s a bit of a trial-and-error process with users,” Koch said.

But, summing up the tool’s impact, McAlpin said, “It’s taking rocks out of rucksacks.”

The technology is particularly helpful under the realities of schedule disruptions. As McAlpin mentioned, an unexpected change can create a frustrating snowball effect, scrapping a two-week schedule that may have taken days to build. The algorithm easily accounts for sudden changes, and it can plan up to six months ahead. Changes are still inevitable, but the system allows Airmen to gain more predictability around their schedules.

The team is considering other applications of their research. Puckboard is used widely across the Air Force for other scheduling needs, though each optimization problem is unique. “It’s a whole different set of efficiencies and a new set of problems, but that’s the exciting thing with these. It’s a nice thing for a researcher. We want to solve real problems,” Balakrishnan said.

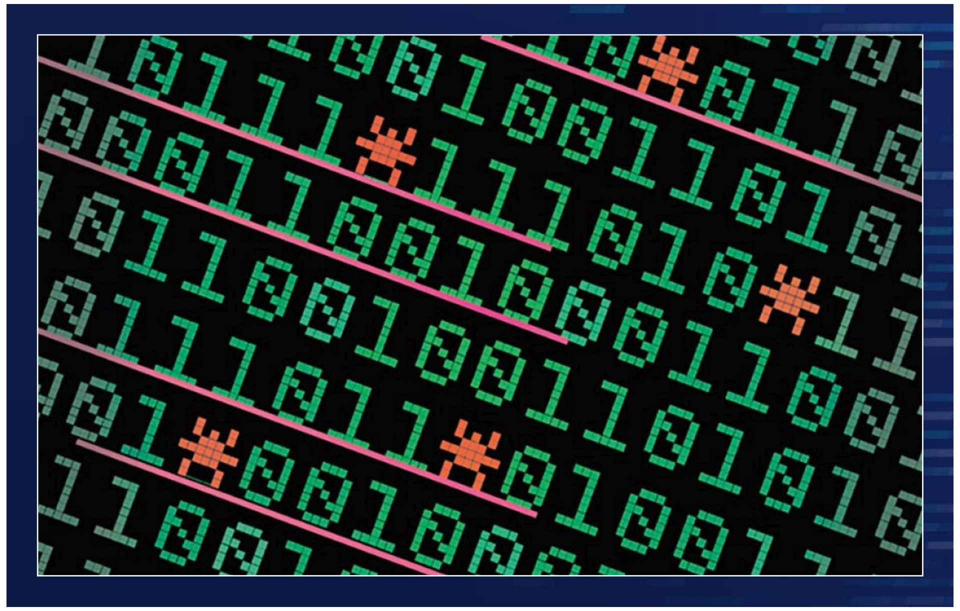
In May, Koch defended his thesis on this project to complete his master’s degree. Sharing the sentiments of his colleagues, Koch said that the seamless collaboration between all three partner organizations in the DAF–MIT AI Accelerator program was invaluable. He himself personifies all three institutes, as an MIT student, a Lincoln Laboratory Military Fellow, and a lieutenant in the Air Force.

“It’s very cool to see how many people care,” Koch said about the collaborators in the program. “With this program, I see the Air Force letting its guard down and letting others in to help us leverage AI and machine learning to make people’s lives better on a daily basis. As a member of the Air Force, I appreciate that.”

Bug-Injecting System Helps to Advance the State-of-the-Art in Debugging Software

The Bulletin

March 5, 2021



By automatically injecting thousands of bugs into program code, LAVA aims to improve vulnerability detection.

Since about the turn of the century, there has been rapid advancement in technological fields such as self-driving vehicles, image recognition, and human language technology. One of the key factors spurring this technology has been competition, and the Laboratory also is playing a key role in driving innovation in these fields.

One example of the Laboratory spurring progress is a system called Large-scale Automated Vulnerability Addition (LAVA), which enables evaluation of bug-finding systems. Bug finding systems are used after developers have written code to try to identify mistakes they have made. If these systems find a bug, they can be fixed easily before code is deployed. Unfortunately, these systems fail to find many bugs, which is one of the reasons why new vulnerabilities and crashes still exist in computer programs today.

"Finding bugs in software is like trying to find a tiny needle buried deep in a very large haystack," said Timothy Leek, Senior Staff, Cyber System Assessments, Group 59.

For many years, the cyber security community has sought to develop new techniques and tools that can more efficiently find bugs in software. However, in 2014, Leek recognized a flaw in these approaches; the scarce documentation of known bugs and how those bugs manifest in a program was making it impossible to measure the success of bug-finding tools. Computer scientists needed to measure the effectiveness of their bug-detectors to know when they were failing to detect a bug. "If a bug finder reports finding 42 bugs in a program, there is no way to know whether that's 1% or 99% of the total number of bugs actually in the program," said Professor Brendan Dolan-Gavitt, a collaborator on LAVA from New York University.

The LAVA system works by producing thousands of realistic bugs that are automatically injected into pre-existing program code. "It turns out that adding bugs to programs teaches us a lot about how bugs work and thus about how to find them," Leek explained. Once these bugs are injected, various vulnerability discovery techniques and software can be tested to see how many of the bugs they can find and how many they miss.

In a virtual seminar on 22 January, staff presented information about four of the eight Laboratory-developed technologies that were selected to receive a R&D 100 Award in 2020, and LAVA was one of these technologies. Speaking during that seminar, Andrew Fasano, Group 59, summarized the usefulness of having a system like LAVA.

"LAVA was the first system of its kind," said Fasano. "In the time since [our work was first published], we've seen a significant increase in the number of bug-finding systems being created, and the majority of those new systems have used our corpus in their evaluation."

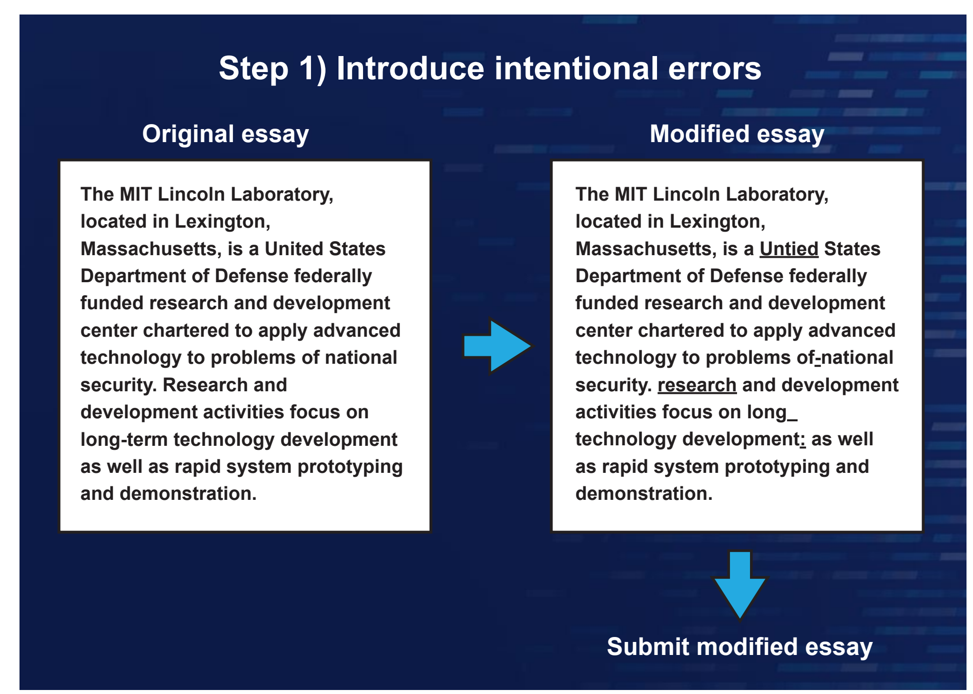
LAVA was developed in 2015–17 with Line funding, and leverages PANDA, another Laboratory-developed dynamic analysis platform and previous R&D 100 winner. In 2016, staff published a paper in IEEE Symposium on Security and Privacy about LAVA, and a few years later, LAVA code was released on GitHub. As a result, over the past five years, LAVA has become the first widely used benchmark for evaluation of bug-finding systems.

After developing and sharing their bug-injecting program, staff needed a way to further test its effectiveness while also disseminating it to the cyber security community. So, they created Rode0day—a bug-finding competition powered by LAVA. Each month, Laboratory researchers released a new, never-before-seen collection of buggy programs and then challenged state-of-the-art bug finding systems to find as many bugs as they can. More than 3,000 LAVA-bugs were injected into more than 60 programs as part of this competition, and the results are being used to develop novel bug-finding systems today. The LAVA team later released their findings from the competition in a paper entitled "The Rode0day to Less-Buggy Programs" that was published to IEEE.

Through the monthly challenges of Rode0day, the LAVA team discovered that it was helpful for the cyber community to have fresh data samples to evaluate and test the effectiveness of their bug-finding tools—however, the competition was insufficient to conduct rigorous evaluations of bug-finding systems, as the Laboratory team didn't know what resources and techniques competitors were using. So, the researchers decided to use the Lincoln Laboratory Supercomputing Center to test leading bug-finding systems.

"Using the Laboratory's supercomputing resources, we were able to run an experiment to test eight of the leading bug-finding systems with a subset of the Rode0day challenges that had nearly 1,000 bugs added," said Fasano, who further explained that this allowed staff to make comparisons using different systems, configurations, and durations. "Using a standard one-CPU system, it would have taken more than 83 years to collect all of the data."

The researchers will be presenting an analysis of their data at the ACM AsiaCCS conference in June, and plan to use this data to continue to advance LAVA's utility and to create more difficult benchmarks for the bug-finding community.



In the illustration above, spelling and grammar errors are introduced to an essay to see if a proofreader can spot those errors. LAVA does this in a similar manner with program code.

Award-Winning Tools Enable Agencies to Efficiently and Easily Analyze Surveillance Footage

MIT Lincoln Laboratory

Division 4: Homeland Protection and Air Traffic Control

Feburary 21, 2021

Closed-circuit television and other surveillance systems are commonplace in busy, high-traffic areas like airports, subway stations, and stadiums. However, when investigators need to use these systems to search through recorded footage, these systems' clunky interfaces and tools can create hours of work just in organizing footage.

The Forensic Video Exploitation and Analysis (FOVEA) suite of tools, developed at the Laboratory in response to the Boston Marathon bombing in 2013, is an easy to set up and use solution to this problem. FOVEA, which won a 2020 R&D 100 Award, can be integrated with surveillance systems to streamline many of the tasks associated with searching through video footage.

"We're helping make existing video surveillance systems more efficient and more effective by giving users tools that help them get through video faster," said Marianne DeAngelus, Homeland Sensors and Analytics, Group 45.

FOVEA's most prominent tool is a video player that makes it easier for researchers to jump between camera views and annotate people or events in the scene. It also includes a tool called Path Reconstruction, which lets a user tag an object and follow it through multiple camera views. Without this tool, researchers have typically had to manually match up video recordings by time, as well as figure out where one camera is in relation to another, to follow (or track) actions through them.

Another tool, called Jump Back, allows investigators to tag an object like a suspicious bag or a vehicle and see when that object first appeared on the recording. Lastly, the Video Summarization tool makes it easy to review hours of video in a matter of minutes.

"The FOVEA program allows us to identify a package that's been left behind, figure out who left it behind, and then start to track that person to determine whether or not this is a threatful situation and deploy proper resources to keep the people safe," said Shawn Doody, Manager, Digital Video Evidence Unit, Washington Metropolitan Area Transit Authority (WMATA) Police, in a video about FOVEA.

All of these tools allow investigators to focus on finding information rather than sorting through it, which can shorten investigations that would typically take days into just hours or even minutes.

"We leverage the computer with the things that it's good at," said DeAngelus. "The machine is very good at crunching data and doing things that require a lot of attention from a human otherwise, so really we want to use the machine as much as possible so the investigator can focus on the higher-level tasks."

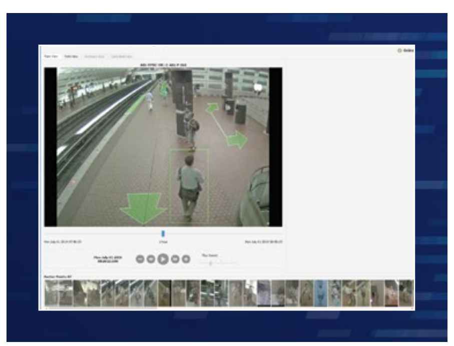
"We had an incident here at Amtrak that took us two full days to identify exactly what had happened," said John Carroll, Senior Director of Emergency Management and Corporate Security at Amtrak, in a video about FOVEA. "When we ran it through [FOVEA], it was able to do it in just a matter of hours. FOVEA acts as a force multiplier."

While these tools may sound like something that would require a powerful computer to run, the team focused on making sure FOVEA tools can run on a wide variety of computers, even those with less processing power. The team also had to make sure that FOVEA could integrate with a wide variety of systems.

"In the surveillance industry, everything is proprietary. It's very difficult to create a third-party plug-in for these systems," said DeAngelus. "One of the main challenges was the significant software engineering. Being able to integrate with the live systems was a challenge, but it had a big payoff in that the tool could then be used on any camera in the whole system."

FOVEA is currently in use by the WMATA, Amtrak, and other organizations. Funding for FOVEA was provided by the Department of Homeland Security's Science and Technology Directorate. The team hopes to expand the program's capability beyond local computer-based implementations and use server-based implementations, while also investigating commercialization of the program.

"The algorithms we originally wrote in MATLAB were very quickly used in the field by real investigators," said DeAngelus. "It was really satisfying to see our ideas and prototypes go all the way into operation."



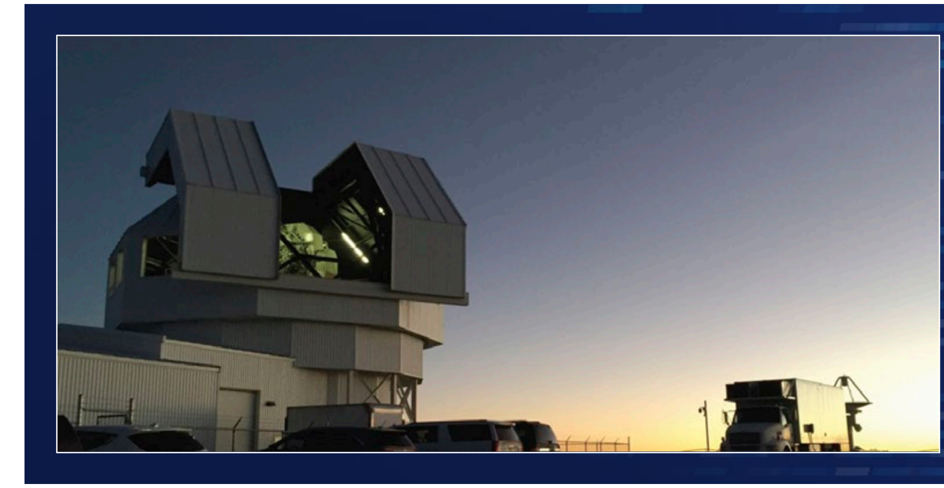
FOVEA tools (shown above) enhance existing video surveillance systems by making it easier for investigators to analyze video footage.

LINEAR Is On the Watch for Potentially Hazardous Asteroids

MIT Lincoln Laboratory

Space Systems and Technology

February 5, 2021



The Space Surveillance Telescope was located at North Oscura Peak on the White Sands Missile Range in New Mexico until 2017 and is currently being relocated to Australia.

NASA estimates that an asteroid the size of a car enters Earth's atmosphere about once a year, creating a great fireball while burning up before reaching Earth's surface; and roughly every 2,000 years, a football-stadium-sized meteoroid strikes Earth potentially causing significant damage. When will the next dangerous asteroid penetrate the atmosphere and seriously impact the Earth? Could that next asteroid be large enough to jeopardize civilization or the future of the human species?

To answer these questions, scientists first need to know what asteroids are orbiting the Sun. Then, their orbits can be calculated to estimate how close any particular asteroid will come to Earth. The Laboratory has been working since the late 1990s to help with the discovery and characterization of potentially hazardous asteroids. Laboratory researchers have found approximately one quarter of all known near-Earth objects (NEOs) that are at least 140 meters (460 feet) in size—large enough to have significant regional effects were they to impact the Earth.

In the mid-1990s, the Laboratory developed a charge-coupled-device (CCD) focal plane upgrade for the U.S. Air Force's Ground-based Electro-Optical Deep-Space Surveillance (GEODSS) system; this upgrade replaced the aging GEODSS cameras that used vacuum tube detectors called Ebsicons. The new back-illuminated, large-format CCDs, built at the Laboratory, and prototype cameras were tested on two developmental GEODSS telescopes at the Laboratory's Experimental Test Site (ETS) at the White Sands Missile Range located near Socorro, New Mexico. These new focal planes both improved the GEODDS system sensitivity and sped up the readout of the collected images. After validating that the new focal planes were capable of wide-area, highly sensitive detection and tracking of satellites, technology transfer of the GEODSS upgrade was initiated, and the prototype CCDs and cameras on the ETS GEODSS telescopes became available for use on other efforts.

In 1996, the Laboratory research team, under the leadership of Dr. Grant Stokes, Division Head, Space Systems and Technology, Division 9, experimented with using the GEODSS telescopes to detect asteroids (or as astronomers say, minor planets) and comets. In 1998, the Laboratory initiated the Lincoln Near-Earth Asteroid Research (LINEAR) program in collaboration with the Air Force and under NASA sponsorship, and LINEAR began using the GEODSS test systems equipped with the large-format CCDs for wide-area search for asteroids and comets. In the first month during which the LINEAR program operated this advanced system, more than 151,000 observations were produced and submitted. All discoveries and observations of asteroids and comets are reported to the International Astronomical Union's Minor Planet Center (MPC) at the Smithsonian Astrophysical Observatory, which is responsible for the verification, designation, orbit computation, and official archive for all known asteroids and comets, as well as for other natural objects orbiting the major planets. At the time, the rest of the asteroid search systems operating on the planet produced and submitted approximately 10,000 observations per month in total.

Also in 1998, NASA established a major program to discover how many natural objects come within the region termed "near-Earth" and thus could represent collision threats. This Near-Earth Object (NEO) program, now the Near Earth Object Observations (NEOO) program, began with a goal of finding and cataloging 90 percent of NEOs larger than 1 kilometer (about 0.6 miles) in diameter—a size at which scientists believed an impact would have worldwide effects.

"As part of the NASA effort to find 1 km and larger Near-Earth Objects, LINEAR was incredibly successful," said Stokes, the principal investigator of the LINEAR program. "In fact LINEAR discovered more of the 1 km and larger NEOs than any other program, and since most of them have now been discovered, that record will stand."

In 2005, as the 1 km and larger goal was nearing completion and search technology was improving, the search goal was extended to discovering 90 percent of NEOs greater than 140 meters in diameter. That goal, now directed to NASA by Congress, was recommended by a Laboratory-led study conducted for NASA.

In 2013, the LINEAR program began transitioning asteroid search operations to the Space Surveillance Telescope (SST), then located at White Sands Missile Range. The 3.5-meter SST was developed by the Laboratory under sponsorship from the Defense Advanced Research Projects Agency for advanced space surveillance applications. The SST is a highly capable telescope whose innovative curved focal plane enables sensitive, wide-area searches of the night sky to detect small objects in geosynchronous orbits. These capabilities are also well suited to searching for asteroids and comets.

Since the SST began observations in January 2014, the LINEAR program has reported more than 14 million observations of asteroids and comets to the MPC and has found 6,001 new objects, including 142 previously undiscovered NEOs and eight new comets. In 2015, the LINEAR team submitted 7.2 million observations to the MPC, making SST the most productive asteroid search instrument ever in terms of number of observations submitted in a single calendar year.

The SST was recently relocated from New Mexico to the Naval Communication Station Harold E. Holt on the North West Cape in Western Australia, where, as part of the U.S. Space Surveillance Network, it will be operated jointly by the U.S. Space Force and the Royal Australian Air Force. Discussions between NASA and the U.S. Space Force are in progress for extending the current LINEAR asteroid search effort to SST, with the goal of SST continuing to contribute to NASA's asteroid search and discovery mission.

This meteor crater in Arizona, about 3,397 feet wide and 558 feet deep, was created by the impact 50,000 years ago of an asteroid estimated at 130 to 164 feet in diameter, about 35 percent of the size of the asteroids the NEOO program focuses on. (Photo Credit: NASA)

The LINEAR program has been a reliable, prolific discoverer of NEOs, comets, unusual asteroids, and main-belt asteroids. It achieved its success by using advanced telescopes, detectors, and detection algorithms and by aggressively scheduling the telescopes for as many nights as possible each lunar month—thus covering the entire available sky at least once each month, weather permitting.

"All together, the LINEAR program is responsible for discovering more than 24 percent of all currently known potentially hazardous asteroids," said Stokes.

Researchers working in the LINEAR program have also contributed to the scientific characterization of the NEO population through their analysis of the number, orbital properties, albedo properties (i.e., reflection of solar radiation), sizes, and impact hazards of NEOs. The archive of LINEAR images, generated from the 1-meter GEODSS telescopes located at ETS (amounting to more than 6 million images), has been used by scientists to accomplish numerous astronomical investigations.

Originally, the discovery of new asteroids was driven by astronomers' curiosity. Now an important motivation is scientists' desire to reliably discover any asteroids that may eventually be a collision threat to Earth. With knowledge of the number, composition, and behavior of the vast population of asteroids, scientists gain a better understanding of the likelihood and the potential for such collisions. One goal of NASA's continuing NEOO program is to provide warning and other data that could help nations prepare for impending collisions and thus to reduce the casualties and other damages caused by an asteroid impacting Earth's surface.

Mission-ready Reinforcement Learning

MIT Lincoln Laboratory

The Bulletin

August 28, 2020

The MeRLin program is looking to solve complex planning and coordination problems across a range of Department of Defense mission areas. Shown above is a simulation environment known as army GroundForce, developed for training artificial intelligence systems for short-range air defense missions.

Reinforcement learning is an exciting emergent technology with strong applicability across a range of Department of Defense (DoD) mission areas. It is a form of machine learning where an agent is trained via rewards and punishments based on its performance. Deep reinforcement learning (DRL) has shown the capacity to find effective strategies in complex and highdimensional environments, such as learning to play strategy games like StarCraft or Defense of the Ancients. As a result, the DoD is investing in simulation environments for training autonomous agents to solve challenging DoD problems. For example, the Laboratory has helped develop U.S. Army environments for training artificial intelligence agents to defend against unmanned aerial systems, rockets, artillery, and mortar threats. Likewise, the Defense Advanced Research Projects Agency has developed the AlphaDogfight environment to train fully autonomous pilots in aerial combat.

While the development of simulation environments provides an exciting opportunity for training machine learning approaches to DoD mission areas, there are several limitations to current state-of-the-art methods that will likely restrict their integration in the near future.

These limitations fall into three related categories: sample efficiency, adaptability, and interpretability. For most environments, DRL methods may require years of virtual simulation time to master a specific task.

Additionally, while DRL methods may find a solution to a challenging task, the learned solution is often brittle and not adaptable to changing objectives and perturbations. Finally, the learned solutions are often black box models that are difficult to interpret in any meaningful way, making it a challenge to verify their safety and performance and incorporate them into any composite system or teaming environment. One application which stresses these limitations in particular is human-machine teaming, where agents do not have sufficient samples to train against specific human tendencies, must adapt to changing objectives, and must act in a reasonable or interpretable way to effectively coordinate with their human counterparts.

The Line-funded program Mission-ready reinforcement learning (MeRLin) is focusing on developing and training DRL algorithms capable of maintaining performance on complex tasks with human allies. To achieve this goal and to help overcome the limitations described above, the program team is researching cross-cutting algorithm solutions. The program staff's initial focus has been to examine fully autonomous teams in complex coordination tasks, where each agent of a team independently controls itself. AI technologies and support their deployment to DoD missions.

Staff Build Robust Algorithms to Strengthen Machine Learning Methods

MIT Lincoln Laboratory

Cyber Security and Information Sciences

June 26, 2020

The goal of the Robust Artificial Intelligence Development Environment project is to design and train machine learning models to excel in the face of unexpected data.

Machine learning is a process where algorithms are trained on data to make predictions. Recommendations on Netflix are one example—the Netflix algorithm is programmed to keep track of users' behavior on the platform and use that data to recommend shows and movies. Weather forecasts are another example—those algorithms are trained to parse past and current weather conditions to predict what might happen next.

Beginning in the 1990s, a series of breakthroughs in artificial intelligence capabilities has brought machine learning to the forefront of applications across industry and government. However, there are vulnerabilities inherent in these systems that have yet to be addressed, which is concerning because these techniques are being adopted by the U.S. Air Force and used for other applications in national security. Dealing with data uncertainty is one such issue.

"Despite a lot of recent successes for many different tasks, [machine learning algorithms] have also been shown to be brittle," said Olivia Brown, Artificial Intelligence Technology, Group 01. "They can easily fail when the data or conditions that they're being tested against start to change from the data and conditions that they encountered during training."

Brown is part of a research team that is addressing this issue of data uncertainty through the Robust Artificial Intelligence Development Environment (RAIDEN) project. Their goal is to find ways to design and train machine learning models to be robust, meaning that they won't break down in the face of unexpected data.

Justin Goodwin, Ballistic Missile Defense System Integration, Group 36, who is also part of the project, explained that robust machine learning helps machine learning models be able to handle "small shifts in data between the distribution of data used for training and the data seen once deployed." These shifts can be caused by natural variations in physical characteristics or signal processing parameters, or even intentionally by an adversary.

To build their robust machine learning system, the team is using robust optimization, which involves training algorithms on data from worst-case scenarios. Additionally, they are exploring how to leverage expert knowledge and physics-based constraints as the machine learning model is trained.

As an example of worst-case scenario data, imagine a self-driving car approaching a stop sign where part of the sign has been obscured. Although the car's algorithm has been trained on how to identify stop signs, it may not be able to recognize a sign that has been altered in such a way. In order to ensure that the algorithm would not misclassify a partially obscured stop sign, it would need to be trained using examples of obscured stop signs.

In terms of leveraging expert knowledge and physics constraints, Victoria Helus, Group 36, who is also part of the research team, explained: "We know that a vehicle can only be as wide as one lane on the road. Therefore, an object that a sensor detects that is larger than one lane ought to be classified as 'not a car.' We would want this 'rule' to be incorporated into the training process so that the model's predictions are constrained within this rule."

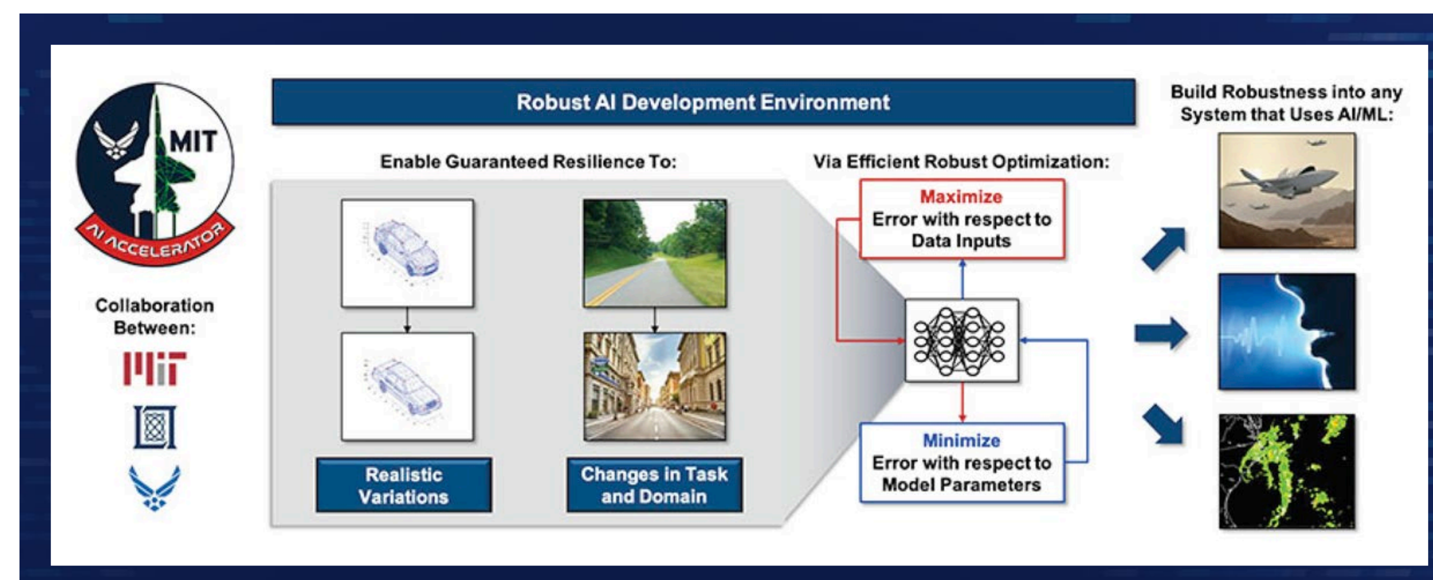
Another goal of the RAIDEN project is to find ways to speed up the process of optimizing and training a robust system, which currently can take up to 10 times longer than training a standard, non-robust system. This long training time is part of why robust algorithms are not often developed for use in real-world systems.

"Another reason robust learning is not commonly used for real systems is that increased robustness often comes with a drop in accuracy on the original—uncorrupted—test data," said Brown. "Robustness is more difficult to define and measure compared to accuracy, and current robust learning techniques are designed to address specific and narrow definitions of robustness. The RAIDEN project is focused on defining robustness metrics and creating solutions for a broader set of robustness objectives, which include both worst-case scenarios and more common data corruptions. Additionally, we are pursuing solutions that achieve a balance between robustness and accuracy."

Now, the team is building a software package that will help researchers experiment with new robust learning techniques as well as understand why and how robust algorithms differ from their nonrobust counterparts. They hope this kind of tool will facilitate more research into robust learning and provide insight into how to debug faulty algorithms and identify their vulnerabilities.

The RAIDEN project is funded through the MIT-Air Force AI Accelerator program, and is a joint effort among the Laboratory, MIT campus, and the U.S. Air Force. Brown said that the Laboratory team meets regularly with the other team members to talk about different ideas and research directions. "As [the other teams] develop new techniques, we can pull promising ones into this software package where we can demonstrate them on U.S. Air Force or Laboratory-relevant data sets," she said.

"Our ultimate goal is to provide the researchers and stakeholders of our community with a set of robustness tools, techniques, and best practices so that they can embrace the great promise of machine learning technology with the confidence that they can meet the safety and security demands that are specific to the national security domain," said Stephen Relyea, Group 36, who is also part of the RAIDEN project team.



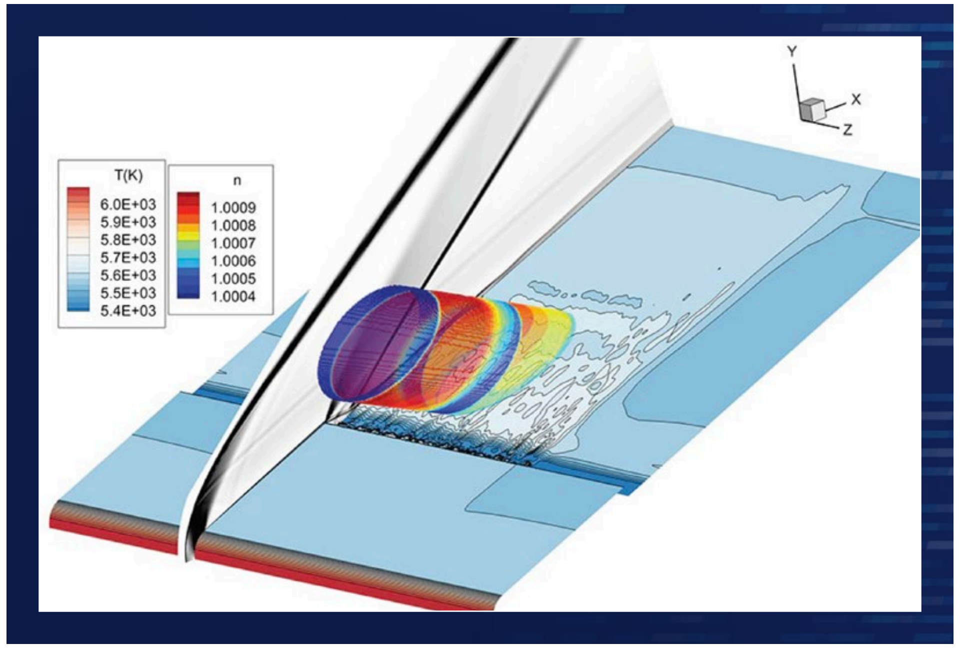
The RAIDEN research team is developing a framework that leverages robust optimization to build machine learning algorithms that are robust to realistic variations and unexpected changes in data inputs for a variety of real-world applications.

High-Fidelity Multi-Physics Simulations for Hypersonics

MIT Lincoln Laboratory

Technology Office

June 12, 2020



This image shows an example simulation of a side-mounted seeker window in a hypersonic flow. A film of Nitrogen gas is used to cool the window, and the impact of this film cooling is included in the aero-optic calculation.

Predicting the performance of hypersonic vehicles and their sub-systems is an inherently multidisciplinary problem. The complex aero-thermal-chemical environment around a vehicle is challenging to simulate and highly coupled. This can impact performance and survivability for both the vehicle as well as any onboard sensors (such as infrared or radio-frequency seekers). Pure fluid dynamic simulations of hypersonic flow are already challenging due to high speeds, thin boundary layers, and chemical reactivity. Adding additional physics, such as structural-thermal coupling or aero-optics, increases the complexity and computational cost. This leads vehicle designers to use lower-fidelity tool sets, which often produce erroneous results due to simplification of complex and critical phenomena.

The goal of this work, supported with Engineering Research allocated funding, is to develop a highfidelity, hypersonic multi-physics solver by coupling a state-of-the-art hypersonic fluid dynamics code (US3D) with the Lincoln Laboratory Integrated Modeling and Analysis Software (LLIMAS) tool. US3D is a widely used and validated hypersonic computational fluid dynamics code developed by Professor Graham Candler at the University of Minnesota. LLIMAS is a simulation framework developed by staff in Engineering, Division 7, to enable data exchange as well as supercomputing and design exploration across multiple disciplines, including academic, commercial, and government codes, and often on terabyte-class models. LLIMAS has been used extensively for low-speed aero-optics modeling but has yet to be expanded to hypersonic flows.

For this work, US3D will be integrated into the LLIMAS framework and the tool set will be developed to enable aero-optical simulations as a first step to wider multi-physics problems. The results of these simulations will be validated against historical aero-optics data from hypersonic ground testing facilities. After validation, additional capabilities will be included, such as thermal-structural coupling, to get a more complete picture of the vehicle's performance.

New Algorithm Uses Supercomputing to Combat Cyber Attacks

MIT Lincoln Laboratory

Cyber Security and Information Sciences

May 15, 2020

This work detects and classifies different types of cyber network attacks. The approach leverages easy-to-collect network packets and uses machine learning to characterize anomalous behavior found in these network packets.

Sophisticated cyber attacks are on the rise. Early techniques for cyber attacks, such as guessing passwords manually, have evolved throughout the years—from session hijacking to ransomware and beyond. In the United States, cyber attacks were estimated to have caused more than $2 trillion worth of damage in 2019, and 350,000 new malware programs continue to be registered every day. The versatility of these attacks compounds the issue because each type of attack interacts with a network in a different way, making it difficult for analysts to identify the attacks and react accordingly.

"Cyber operators across the country are drowning in a sea of false positives," said Dr. Vijay Gadepally, Senior Staff, Lincoln Laboratory Supercomputing Center (LLSC). "Alerts are going off, but they're not necessarily helping operators determine what is going on and how to take appropriate action." For example, existing tools may indicate that something suspicious is going on in the network but fail to characterize what is happening.

In order to manage this tricky cyber attack landscape, analysts need automated tools that can accurately detect and classify threats. Gadepally is working with Emily Do, a former graduate student in MIT's Electrical Engineering and Computer Science department, to apply the power of supercomputing to this very problem. The team is using machine learning to characterize anomalous behavior within a cyber network.

A network packet is a unit of data that is sent from a source to a destination. A series of packets that is sent between a single source and destination is called a flow. "Our hypothesis is that if there is a network anomaly such as a cyber attack, there will be a change in the way flows occur between sources and destinations," said Gadepally.

In order to test this hypothesis, the team first gathered data from the MAWI Working Group Traffic Archive—an open source collection of continually updated raw internet traffic data—and then converted this data into network flow. This aspect of the project necessitated a huge amount of computing power, which is why the LLSC is involved.

"These data sets are massive—a 15-minute window of packet [format] data corresponds to nearly 20GB of data," said Gadepally. "While we end up converting everything to the flow format, which is much smaller, this conversion is a computationally heavy task."

After converting the network packets into flow data, these network flows were bucketed into flows within 10-second time windows. The team then calculated the entropy (the amount of change) of each feature, such as IP addresses, within each time window. The idea is that a change in entropy would be a good indicator of anomalous network activity.

As an example, consider a Distributed Denial of Service attack. In one form of this attack, an attacker disrupts a victim network by bombarding it with traffic from multiple compromised systems. An analyst observing this attack would expect to see a significant increase in the entropy of source IP flow—the increase is due to a big change in the number of source IP addresses within the time window in which the attack occurs.

The last step of the process was to feed the entropy results for each feature into a neural network. This is the step that allows the algorithm to accurately classify a network attack. In the end, the research team found that their system could detect and identify incoming attacks that affected as little as 5 percent of the total traffic flow within a network.

"This work quantifies the sensitivity of our method of detection," said Do. "The lower the number, the more sensitive the detection. [This] means that if your network has a usual total traffic of 10,000 packets per second, the method can detect and classify an attack at 500 packets per second or more with high accuracy."

So far, the team has has been using synthetic data to simulate network attacks for training and testing their algorithm. Now that they have successfully demonstrated the utility of their method, their next step is to use real data to use to test the system further. The end goal is to build an effective, working system that they can make publicly available for all cyber analysts who wish to use it.

MIT Joins White House Supercomputing Effort to Speed Up Search for Covid-19 Solutions

Jennifer Chu

MIT News Office

March 23, 2020

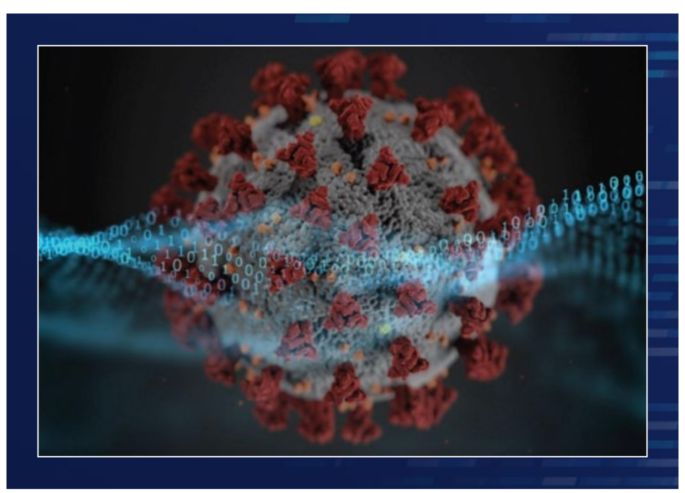
The White House has announced the launch of the Covid-19 High Performance Computing Consortium, a collaboration among various industry, government, and academic institutions which will aim to make their supercomputing resources available to the wider research community, in an effort to speed up the search for solutions to the evolving Covid-19 pandemic.

MIT has joined the consortium, which is led by the U.S. Department of Energy, the National Science Foundation, and NASA.

MIT News spoke with Christopher Hill, principal research scientist in MIT's Department of Earth, Atmospheric, and Planetary Sciences, who is serving on the new consortium's steering committee, about how MIT's computing power will aid in the fight against Covid-19.

Q: How did MIT become a part of this consortium?

A: IBM, which has longstanding computing relationships with both the government and MIT, approached the Institute late last week about joining. The Department of Energy owns IBM's Summit supercomputer, located at Oak Ridge National Laboratory, which was already working on finding pharmaceutical compounds that might be effective against this coronavirus.



MIT joins a consortium of supercomputing resources to help speed the search for Covid-19 solutions. Image: CDC, MIT News

In addition to its close working relationship with MIT, IBM also had donated the Satori supercomputer as part of the launch of the MIT Schwarzman College of Computing.

We obviously want to do everything we can to help combat this pandemic, so we jumped at the chance to be part of a larger effort.

Q: What is MIT bringing to the consortium?

A: We're primarily bringing two systems to the effort: Satori and Supercloud, which is an unclassified system run by Lincoln Laboratory.

Both systems have very large numbers of the computing units — known as GPUs — that enable the machines to process information far more quickly, and they also have extra large memory. That makes the systems slightly different from other machines in the consortium in ways that may be helpful for some types of problems.

For example, MIT's two systems seem to be especially helpful at examining images from cryoelectron microscopy, which entails use of an electron microscope on materials at ultralow temperatures.

Ultralow temperatures slow the motion of atoms, making the images clearer. In addition to the hardware, MIT faculty and staff have already expressed interest in assisting outside researchers who are using MIT equipment.

Q: How will MIT operate as part of the consortium?

A: The consortium will receive proposals through a single portal being run in conjunction with the NSF. A steering committee will decide which proposals are accepted and where to route them. The steering committee will be relying on guidance from a larger technical review committee, which will include the steering committee members and additional experts. Both committees are made of researchers from the participating institutions. I will serve on both committees for MIT, and we'll be appointing a second person to serve on the technical review committee.

Four individuals at MIT — Ben Forget, Nick Roy, Jeremy Kepner (Lincoln Lab), and myself — will oversee the work at the Institute.

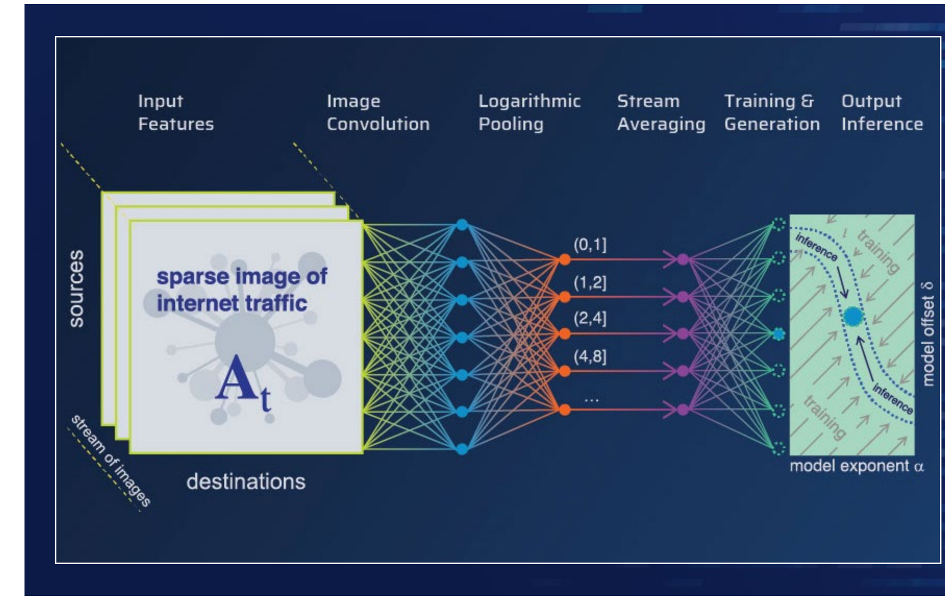
The goal of the consortium is to focus on projects where computing is likely to produce relevant advances in one week to three months — though some projects, like those related to vaccines — may take longer.

Cybersecurity Phenomenology Exploration and Reasoning

MIT Lincoln Laboratory

Technology Office

March 20, 2020



The approach used in Cybersecurity Phenomenology Exploration and Reasoning for analyzing sparse traffic matrices is depicted in this figure, from streaming snapshots of network traffic to model fitting.

Most cyber defense tools identify potential malicious events based on known signatures. Methods that are not signature-based are used to identify new, unknown attack vectors. These methods typically include establishing models to baseline normal traffic behavior and then using those models to find anomalous traffic patterns. The accuracy of these models is dependent on the amount of data used to baseline the network activity. The analysis pipeline to construct these models requires the support of high-performance computing in order to run computationally demanding analytics over large quantities of network packet data.

As part of a Technology Office Seedling project, Laboratory researchers developed a neural network pipeline and utilized the Lincoln Laboratory Supercomputing Center to analyze large, publicly available internet traffic datasets. These data sets comprise 50 billion data packets collected at different locations and at different time points over a period of several years. As part of this work, a low-dimension distribution model was found to accurately fit a range of network characteristics. The parameters of the distribution model were able to delineate different network traffic topologies, which therefore provided a method for characterizing normal traffic behaviors.

The Line-funded Cybersecurity Phenomenology Exploration and Reasoning (CyPhER) project will expand on this prior work by further optimizing the data processing and analysis pipeline, which will enable using larger network traffic matrices to build models. This effort will include integrating GraphBlas, which is a suite of tools for analyzing sparse matrices. The low-dimension distribution models examined as part of the Seedling project will also be further validated using data from the Lincoln Research Network Operations Center. These data will be used to examine the stability of the distribution models over time. With validation, the network traffic baseline models will be utilized to develop anomaly detection algorithms that are not signature-based. The techniques and tools developed as part of this project will improve the efficiency of analyzing traffic matrices and the characterization of network traffic at a large scale.

How Supercomputers Are Helping to Fight COVID-19

Ryan F. Mandelbaum

Gizmodo

March 20, 2020

Summit supercomputer at Oak Ridge National Lab in Tennessee.
Photo: Carlos Jones/ORNL

A host of companies, including IBM, Microsoft, and Google, along with universities and national labs have teamed up to form the COVID-19 High Performance Computing (HPC) Consortium. This new partnership is designed to provide scientists with supercomputing resources as they figure out how to combat the coronavirus-caused disease known as COVID-19.

Faced with a rapidly spreading illness, scientists can create thousands of models on supercomputers in order to better understand the epidemic, characterize the virus, and devise potential vaccines and drug treatments. The organizers of the new consortium will provide 16 supercomputing systems to researchers, as well as a community to engage in the fight together.

"The benefit of having the consortium is to speed up and accelerate the scientific discovery that has to happen in order to develop a vaccine, understand the virus, and eventually kill it," Michael Rosenfeld, vice president of Data Centric Solutions at IBM, told Gizmodo. He said that high-performance supercomputers might be able to do in minutes or hours what regular computers do in days, months, or years.

The consortium currently represents supercomputers from companies including IBM, Amazon, Google, and Microsoft; universities including Massachusetts Institute of Technology and Rensselaer Polytechnic Institute; and Department of Energy National Laboratories including Lawrence Livermore, Oak Ridge, and Los Alamos, as well as NASA and the National Science Foundation. The consortium is encouraging COVID-19 researchers to submit proposals through a central portal, which a steering committee will review in order to connect researchers with the right supercomputing resources.

Supercomputing centers have always supplied discretionary computing time for emergencies, such as during hurricane response, said Kelly Gaither, Texas Advanced Computing Center's director of health analytics. She told Gizmodo it was a no-brainer to devote time to fighting the coronavirus.

As for what scientists will actually do with supercomputers during this pandemic, many are trying to understand the structure of the virus and its "spike" protein, as well as how it differs from other coronaviruses, like the virus behind SARS. Supercomputers have already shown their worth in fighting the disease on this front; the Summit supercomputer at Oak Ridge National Laboratory allowed researchers to whittle 8,000 potential virus-fighting molecules down to just 77, for example. Others are using the computers to generate simulations of how the pandemic could play out, when the peak will occur, how long it will last depending on what measures are in place, and what locations will be in most need of supplies.

Researchers have already been submitting proposals to research the virus using U.S. supercomputing resources. The National Science Foundation (NSF) issued a call for proposals relating to COVID-19 earlier this month and has already funded 10 quick-turnaround grants totaling $1,592,789, an NSF spokesperson told Gizmodo.

Rosenfeld told Gizmodo that the consortium offers researchers an opportunity to collaborate in ways they might not have done before, such as by helping one another get their code up and running more quickly on the processors. Gaither said that this encourages scientists from disparate specialties to link up and solve problems in new ways and to think creatively about how to incorporate supercomputers into their research.

While it's impossible to predict how long this pandemic will last, we can only hope that new scientific advances will help us beat it sooner and increase our defenses against future pandemics.

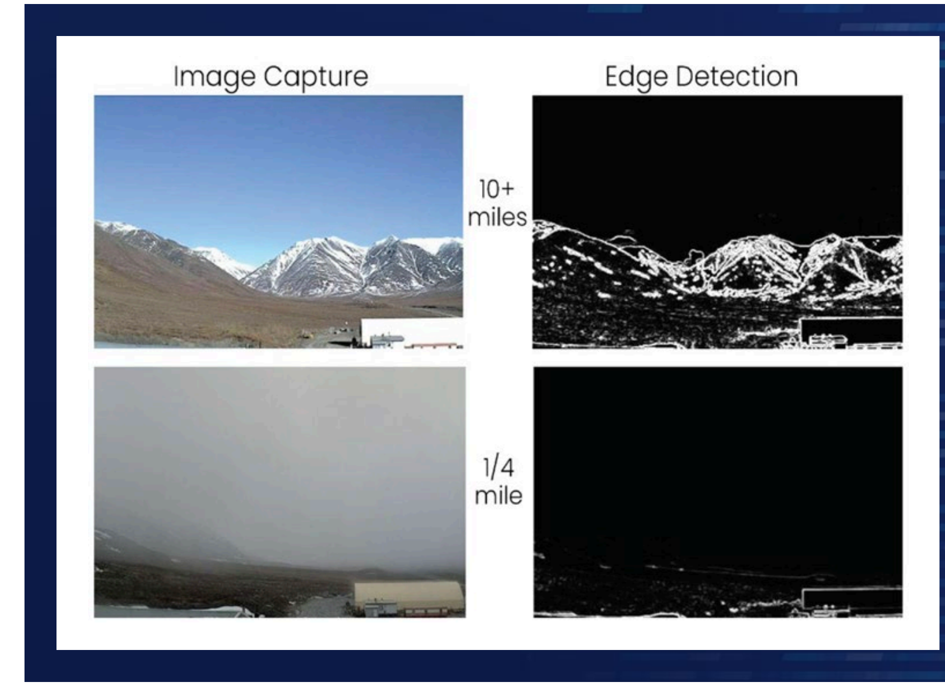
Source: Gizmodo

An Algorithm with an Eye for Visibility Helps Pilots in Alaska

Kylie Foy

MIT Lincoln Laboratory

December 27, 2019



An FAA web camera in Chandler Shelf, Alaska, captured these clear-day and low-visibility images. A Sobel filter highlights the edges in both photos, and the VEIA algorithm uses the number and strength of the edges to estimate the visibility in miles.

More than three-quarters of Alaskan communities have no access to highways or roads. In these remote regions, small aircraft are a town's bus, ambulance, food delivery — the only means of getting people and things in and out.

As routine as daily flight may be, it can be dangerous. These small (or general aviation) aircraft are typically flown visually, by a pilot looking out of the cockpit windows. If sudden storms or fog appear, a pilot might not be able to see a runway, nearby aircraft, or rising terrain. In 2018, the Federal Aviation Administration (FAA) reported 95 aviation accidents in Alaska, including several fatal crashes that occurred in remote regions where poor visibility may have played a role.

"General aviation pilots in Alaska need to be aware of the forecasted conditions during preflight planning, but also of any rapidly changing conditions during flight," said Michael Matthews, a meteorologist at Lincoln Laboratory. "There are certain rules, like you can't fly with less than three miles of visibility. If it is worse, pilots need to fly on instruments, but they need to be certified for that."

Pilots check current or forecasted weather conditions before they fly, but a lack of automated weather observation stations throughout the Alaskan bush makes it hard to know exactly what to expect. To help, the FAA recently installed 221 web cameras near runways and mountain passes. Pilots can look at the image feeds online to plan their route. Still, it's difficult to go through what could be hundreds of images and estimate just how far one can see.

So, Matthews has been working with the FAA to turn these web cameras into visibility sensors. He has developed an algorithm, called Visibility Estimation through Image Analytics (VEIA), that uses a camera's image feed to automatically determine the area's visibility. These estimates can then be shared among forecasters and with pilots online in real time.

In concept, the VEIA algorithm determines visibility the same way humans do. It looks for stationary "edges." For human observers, these edges are landmarks of known distances from an airfield, such as a tower or mountain top. They're trained to interpret how well they can see each marker compared to on a clear, sunny day.

Likewise, the algorithm is first taught what edges look like in clear conditions. The system looks at the past 10 days' worth of imagery, an optimal timeframe because any shorter timeframe could be skewed by bad weather and any longer could be affected by seasonal changes, according to Matthews. Using these 10-day images, the system creates a composite "clear" image. This image becomes the reference to which a current image is compared.

To run a comparison, an edge-detection algorithm (called a Sobel filter) is applied to both the reference and current image. This algorithm identifies edges that are persistent — the horizon, buildings, mountain sides — and removes fleeting edges like cars and clouds. Then, the system compares the overall edge strengths and generates a ratio. The ratio is converted into visibility in miles.

Developing an algorithm that works well across images taken from any web camera was challenging, Matthews said. Based on where they are placed, some cameras might have a view of 100 miles and others just 100 feet. Other problems stemmed from permanent objects that were very close to the camera and dominated the view, such as a large antenna. The algorithm had to be designed to look past these near objects.

"If you're an observer on Mount Washington, you have a trained eye to look for very specific things to get a visibility estimate. Say, the ski lifts on Attitash Mountain, and so on. We didn't want to make an algorithm that is trained so specifically; we wanted this same algorithm to apply anywhere and across all types of edges," Matthews said.

To validate its estimates, the VEIA algorithm was tested against data from Automated Surface Observing Stations (ASOS). These stations, of which there are close to 50 in Alaska, are outfitted with sensors that can estimate visibility each hour. The VEIA algorithm, which provides estimates every 10 minutes, was more than 90% accurate in detecting low visibility conditions when compared to data from co-located ASOS systems.

The FAA plans to test the VEIA algorithm in summer 2020 on an experimental website. During the test period, pilots can visit the experimental website to see realtime visibility estimates alongside the camera imagery itself.

"Furthermore, the VEIA estimates can be ingested into weather prediction models to improve the forecasts," said Jenny Colavito, who is the Ceiling and Visibility Research Project Lead at the FAA. "All of this leads to keeping pilots better informed of weather conditions so that they can avoid flying into hazards."

The FAA is looking into using weather cameras in other regions, starting in Hawaii. "Like Alaska, Hawaii has extreme terrain and weather conditions that can change rapidly. I anticipate that the VEIA algorithm will be utilized along with the weather cameras in Hawaii to provide as much information to pilots as possible," Colavito added. One of the key advantages of VEIA is that it requires no specialized sensors to do its job, just the image feed from the web cameras.

Matthews recently accepted an R&D 100 Award for the algorithm, named one of the world's 100 most innovative products developed in 2019. As a researcher in air traffic management for 28 years, he is thrilled to have achieved this honor.

"Some mountain passes in Alaska are like highways, especially in the summertime, with the number of people flying. You can find countless stories of terrible crashes, people just doing everyday things — a family on their way to a volleyball game," Matthews reflected. "I hope that VEIA might help people go about their lives safer."

Robert Hallowell, left, and Michael Matthews, right, developed the Visibility Estimation through Image Analytics algorithm. Photo: Nicole Fandel

Enabling the Foundations of AI: Data, Computation, and Algorithms

MIT Technology Office

November 1, 2019

Laboratory staff in the Lincln Laboratory Supercomputing Center are developing tools to address challenges in data management and algorithmic techniques for novel neural network architectures to enable rapid prototyping of AI.

A critical component of artificial intelligence (AI) initiatives across the government is access to computing and infrastructure that can be used by the wider national security community. While AI holds significant potential to improve Department of Defense (DoD) operations, there are a number of critical challenges that are not directly solved by commercial or academic solutions. These challenges include (1) managing DoD-specific data sets that are vast and diverse, (2) optimizing the performance of enormous inference and training calculations on heterogeneous computing hardware, and (3) developing algorithmic techniques for novel neural network architectures.

To enable AI research for a variety of Laboratory and sponsor missions, the Line-funded project Enabling the Foundations of AI is developing and operationalizing technologies within the Lincoln Laboratory Supercomputing Center (LLSC) to address these challenges. First, since data management is a critical and time-consuming part of an AI workflow, this project is focused on developing extensible data hubs that will allow collaborators to integrate federated datasets for use in AI solutions designed for DoD problems. Currently, this data hub technology is being considered for transition to the Navy. Second, training AI models on large datasets can require several days of compute time. To enable faster training and inference on new and legacy AI hardware deployed by the DoD, Laboratory staff are collaborating with Dr. Tao Schardl and Professor Charles Leiserson at MIT CSAIL to embed parallelism into compilers for AI frameworks. By optimizing low-level parallel computations, this approach has shown a speedup of 30–100 percent across different CPU architectures. Laboratory staff are also developing capabilities to enable massively parallel, distributed neural network training on LLSC's TX-GAIA cluster, which has a peak AI performance of 100 petaflops. Finally, on the algorithm front, this project is exploring the mathematical foundation of sparse AI to help make the case for using sparse processing hardware and hypersparse graphs. For demonstration purposes, this research is being applied to the example problem of detecting cyber anomalies in internet traffic.

Through developments in data management, compute optimization, and algorithms, this project has significantly impacted missions at the Laboratory and has been a critical enabler of the recently announced MIT-Air Force AI Innovation Accelerator. Any technologies developed through this program that are mature enough will be made available to all LLSC users.

Video and Imagery Dataset to Drive Public Safety Capabilities

Homeland Protection and Air Traffic Control

MIT Lincoln Laboratory

August 23, 2019

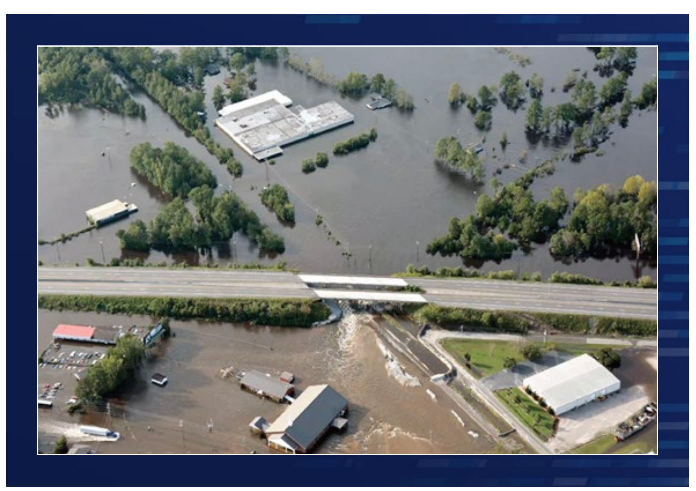
Imagine the following scenario: It is nighttime. A police officer exits their patrol vehicle, approaches another person, and then begins running—indicating that a foot pursuit has ensued.

The officer is equipped with a body-worn camera that monitors the scene, makes observations about its surroundings, and—in the case of an incident—uses machine-learning analytics to determine what action is needed.

"Computer vision capabilities have rapidly been advancing and are expected to become an important component to incident and disaster response," said Andrew Weinert, Surveillance Systems, Group 42. "However, a majority of video analytics are not meeting public safety's needs due to the lack of appropriate training data and requirements."

Weinert and about a dozen staff from Groups 42, 44, and 45 have been developing a computer vision dataset of operational and representative public safety scenarios. This dataset will enable technology development tailored to public safety scenarios, and includes operational images and videos from several organizations.

In the case of the law enforcement scenario, the researchers' dataset would effectively allow these body-worn cameras to function as a second set of eyes for police dispatchers. The cameras would actively apply a "20 Questions" approach to analyze what they are seeing and assess the officer's risk level.



Laboratory staff have built a labeled imagery and video dataset that can recognize elements of a scene. For example, with existing commercial software, the highway shown above was incorrectly classified as a runway or heliport, due to the surrounding flooding. With the researchers' additional training data, the highway will be correctly labeled.

When a certain amount of these triggers are met [e.g. Is it nighttime? Is the officer near someone else? Are they running?], the analytics software in the camera could recognize that the officer is in a high-risk scenario, automatically trigger the camera's recording function, and then notify the local agency of the incident. After receiving the notification, the agency might choose to take further action, such as launching a drone to follow the suspect and gather additional data.

Another potential use for this video and imagery dataset involves labeling low-altitude airborne imagery to support search and rescue operations. To build their dataset, the researchers have tapped into a database of more than 500,000 low-altitude airborne images, 80,000 of which were collected over Puerto Rico as part of the Hurricane Maria response. They have labeled the images so that their machine learning algorithms can recognize a wide range of relevant public safety features in different environments.

"The information within these images could improve various aspects of a response and recovery effort, such as damage assessment," explained Weinert. "Our dataset will enable the development of machinelearned analytics to prioritize and characterize images."

Both the law enforcement bodyworn cameras and the low-altitude airborne imagery applications are common scenarios nationwide that would maximize the utility of the researchers' dataset.

The researchers are now focused on transitioning the dataset and software to their sponsor, the New Jersey Office of Homeland Security and Preparedness, and to the public domain. In the future, they would like to develop decision support capabilities using the collection of low-altitude imagery, as well as address computer vision challenges using the dataset.

"This dataset will serve as risk reduction to enable academia and industry to tailor applications for public safety," said Weinert. "It will also improve existing commercial and open-source capabilities to be more robust to incident and disaster imagery."

Record-breaking DNA Comparisons Drive Fast Forensics

IdPrism and its award-winning algorithms provide rapid analysis for complex forensic DNA samples.

Anne McGovern

MIT Lincoln Laboratory

June 17, 2019

New algorithms encode genetic markers as bits to allow for fast DNA comparisons in forensics. Illustration: Thomas Sheehan.

Forensic investigators arrive at the scene of a crime to search for clues. There are no known suspects, and every second that passes means more time for the trail to run cold. A DNA sample is discovered, collected, and then sent to a nearby forensics laboratory. There, it is sequenced and fed into a program that compares its genetic contents against DNA profiles stored in the FBI's National DNA Index System (NDIS) — a database containing profiles of 18 million people that have passed through the criminal justice system. The hope is that the crime scene sample will match a profile from the database, pointing the way to a suspect. The sample can also be used for kinship analysis through which the sample is linked to blood relatives, as was done in April 2018 to catch the infamous Golden State Killer.

DNA forensics is a powerful tool, yet it presents a computational scaling problem when it is improved and expanded for complex samples (those containing DNA from more than one individual) and kinship analysis. Consider the volume of data that the FBI must handle for the nation. "If you think of all the police stations across the country, all operating each week, it's a lot of data to keep track of and organize," said Darrell Ricke from the Bioengineering Systems and Technologies Group. To put this into perspective, if each state compares 2,000 crime scene samples weekly, that's 100,000 samples to compare against 18 million profiles per week.

Ricke is part of a team at the Laboratory that developed an integrated web-based platform called IdPrism that provides expanded comparison capabilities without compromising speed or functionality. IdPrism allows identification of more than 10 individuals in a complex DNA sample along with extended kinship results. At its heart are two algorithms that Ricke developed, FastID and TachysSTR, which encode genetic markers as bits (0 or 1) and operate quickly and smoothly. These algorithms recently won a 2018 R&D 100 Award, which is given annually by R&D Magazine to the 100 most significant inventions of the year.

These markers are two types of variations in DNA called short tandem repeats (STR) and single nucleotide polymorphisms (SNP). They are considered to be a kind of DNA fingerprint that can be used to identify individuals as well as their relatives. Each person has a unique combination of SNP or STR variations — one person's combination presents in a specific pattern while another person's presents in a different pattern. When analysts run a crime scene DNA sample against a profile in the NDIS database, their finding a matching combination of these STRs shows a high chance that the DNA belongs to the same person.

The FBI currently uses software algorithms that must pass through a complex set of calculations to reveal if a sample matches a profile. Ricke's algorithms assign a bit value to normal (0) or rare (1) versions of SNPs or a bit for each different STR marker. The normal label indicates that the SNP or STR is common in many people and is thus not a unique marker that can be used to identify an individual. With this digital DNA encoding for both identity comparisons and complex mixtures, analysis can be done with just three hardware bit instructions: exclusive OR, logical AND, and population count.

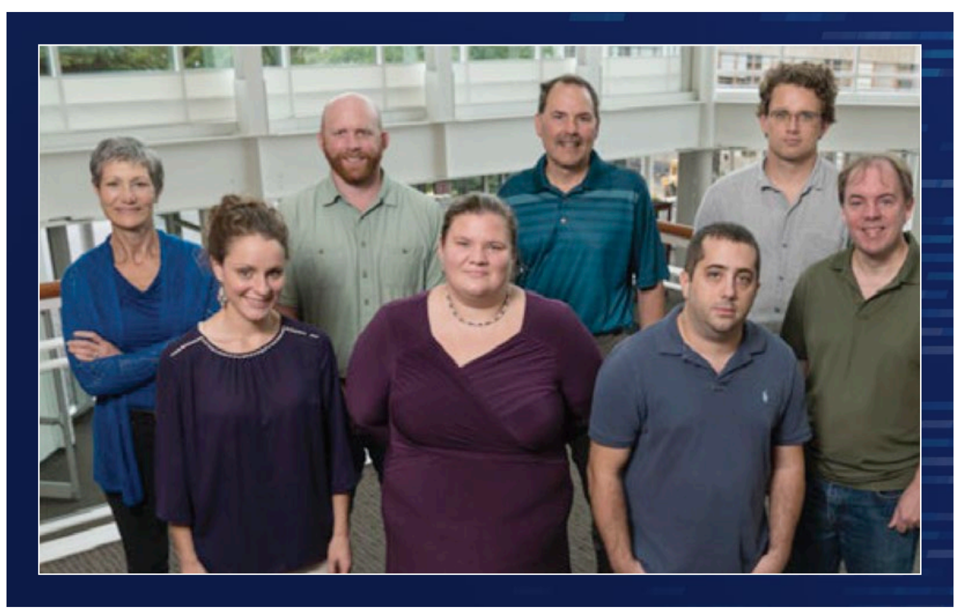
An exclusive OR instruction allows for a comparison of whether two DNA profiles are the same or different. For the forensic comparisons, this instruction will output a 0 when an SNP or STR in a sample matches that in a profile, and it will output a 1 when they don't match. This technique works well when the crime scene sample contains DNA from only one individual, but if there are more contributors, a matching result could be hidden among mismatches from the other people in the same sample. This issue is addressed by adding a logical AND with the database profile to the results of the exclusive OR. This step, in a sense, gets rid of the mismatch noise to reveal whether the database profile has matched against an individual in the sample. The final step is population count, which sums up all of the 1s. In the end, a match is represented by mostly 0s and a mismatch will have a high number of 1s.

Using these three hardware bit instructions, the FastID algorithm can compare 5,000 SNPs in a crime scene DNA sample against 20 million reference profiles in under 12 seconds. Alternative methods would take hours to do so on this scale. Similarly, TachysSTR can compare STRs in 1 million samples in 1.8 seconds whereas current algorithms take 10 minutes to do the same.

The results are displayed inside the IdPrism system in which investigators can run, view, query, and store their DNA comparison data. In addition to being fast and convenient, the system has improved the accuracy of forensics by including a panel of 2,650 SNP markers that are used for complex sample and kinship analysis.

Last November, the system was transitioned to users outside of the Laboratory. "Although getting IdPrism to a transition-ready product was challenging, it is awesome to think that our technology is being used," said Philip Fremont-Smith, who is also from the Bioengineering Systems and Technologies Group and was involved in the bioinformatics side of the project.

"When Hollywood finds out about this, they're going to change their scripts," Ricke said. "The capabilities are so different from what's out there."



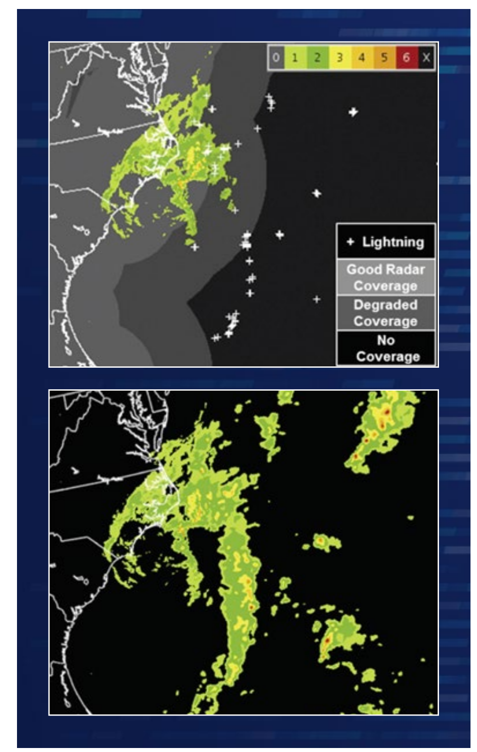
Ricke worked with the bioinformatics team to develop the IdPrism system: back row, Martha Petrovick, Philip Fremont-Smith, Darrell Ricke, and James Watkins; front row, Mallory Clark, Sara Stankiewicz, Adam Michaleas, and Thomas White. Photo: Glen Cooper

Creating Synthetic Radar Imagery Using Convolutional Neural Networks

Homeland Protection and Air Traffic Control

MIT Lincoln Laboratory

February 15, 2019



The figure on the top shows precipitation intensity, with lightning flashes as white plus symbols in the area outside the coverage of weather radar. By comparison, on the bottom are results from the Offshore Precipitation Capability's synthetic radar, which was created using convolutional neural networks and the methods described in this year's Laboratory Best Paper.

Weather radar can track the location and intensity of storms and is useful for managing transportation around hazardous weather. Air traffic controllers, for example, rely on weather radar to track storms that could impact aircraft and flight schedules. Although land-based radar is sufficient to cover most continental air space, many offshore and oceanic controllers do not have sufficient access to the weather information that they need for proper air traffic management.

For this reason, a few years ago, staff from Air Traffic Control Systems, Group 43, developed the Offshore Precipitation Capability (OPC), a system that creates a radar-like depiction of precipitation—known as synthetic radar—by combining data from multiple nonradar sources. These nonradar sources include cloud-to-ground lightning flashes, geostationary satellite imagery, and numerical weather prediction models. This provides coverage in areas where there is no weather radar coverage.

Now, staff are taking this work one step further by implementing convolutional neural networks into the OPC system. Convolutional neural networks are useful in areas such as image recognition and classification. For OPC, they serve as a means of combining the nonradar data sources to create synthetic radar.

"Convolutional neural networks do not require hand-engineered weather features to train," said Christopher Mattioli, Group 43, regarding the benefit that convolutional neural networks bring to OPC. "They possess a unique ability to pick up on particular aspects of the weather scene that even the human eye can miss. On top of that, their ability to view disparate inputs as a single entity, all of which exist at different resolutions, is a key strength. This allows for a fusion of information, which ultimately informs the creation of the synthetic radar."

Mattioli and fellow staff recently published their paper "Creating Synthetic Radar Imagery Using Convolutional Neural Networks" February 15, 2019 in the Journal of Oceanic and Atmospheric Technology in December, and then received this year's Best Paper Award from the Laboratory. The Best Paper Award recognizes the authors of the most outstanding paper appearing in a peer-reviewed journal or peer-selected conference publication during an approximately one-year period preceding the award announcement.

"The capability described in this paper fills this gap [of coverage outside the range of land-based radar] by providing real-time radar-like analysis of current weather conditions in offshore regions using advanced machine learning methods," said Dr. Mark Veillette, Group 43, and lead author of the paper. "It was quickly adopted by the Federal Aviation Administration and the air traffic community, and is now in the process of being acquired by the U.S. Air Force for global operations."

During the summer of 2017, a research prototype of the system was transitioned and demonstrated in five key air traffic control centers, including Miami, San Juan, Houston, New York City, and the Air Traffic Control System Command Center in Warrenton, Virginia. After the destruction of the weather radar on Puerto Rico by Hurricane Maria, OPC was used extensively by the National Weather Service office in San Juan. Thus far, the work has focused on generating synthetic radar over the Gulf of Mexico and the Caribbean Sea, but future steps will involve extending this capability to regions of the world without weather radar. Moving forward, advanced sensors and newly acquired training data will also lead to better model performance and greater coverage.

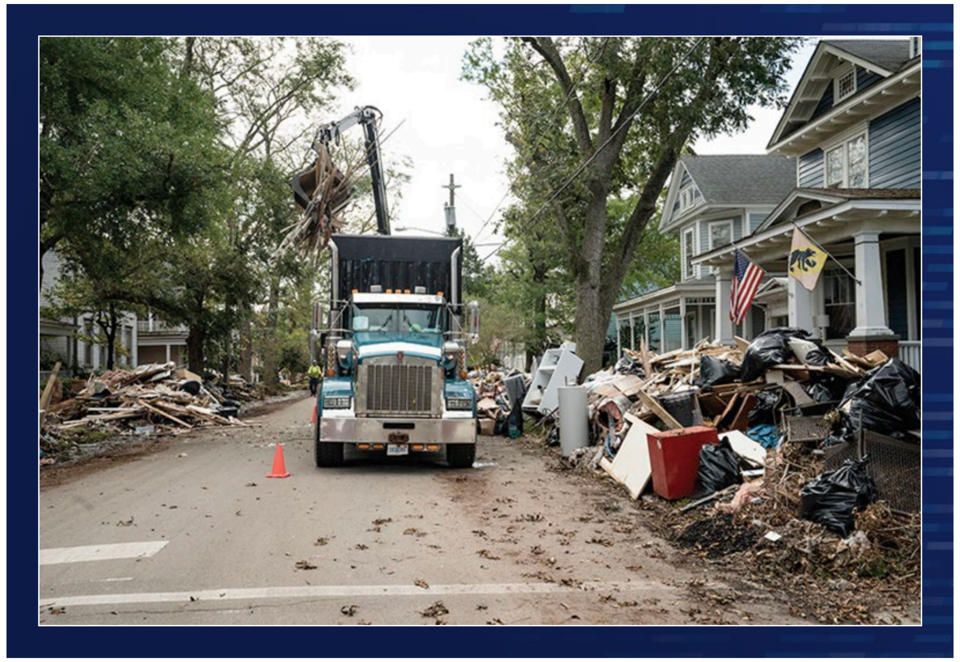
Eric Hassey, Group 43, said that the team behind the paper are grateful for the support they received from Laboratory groups, such as the Lincoln Laboratory Supercomputing Center, which helped make the analysis in the paper possible. "The team is thrilled and honored to receive the Laboratory's prestigious Best Paper Award," said Hassey. "Thank you to the selection committee for recognizing this work."

Lidar Scans Over the Carolinas Accelerate Hurricane Recovery

Homeland Protection
and Air Traffic Control

MIT Lincoln Laboratory

January 11, 2019



A street in New Bern, North Carolina, is lined with debris, much of it dragged out from water-damaged homes after Hurricane Florence. Lidar data can tell FEMA the volume of debris piles, helping the agency coordinate their removal.

Hurricane Florence's slow trot over North and South Carolina in September led to inundating rain, record storm surges, and another major disaster for the Federal Emergency Management Agency (FEMA) to contend with. Facing damage over hundreds of square miles, FEMA called upon the Laboratory to use its state-of-the-art light detection and ranging (lidar) system to image the destruction in the region.

Installed onto an airplane and flown nightly over a disaster site, the lidar sends out pulses of laser light that bounce off the land and structures below and are collected again by the instrument. The timing of each light pulse's return to the instrument is used to build a "point-cloud map": a high-resolution 3D model of the scanned area that depicts the heights of structures and landscape features. Laboratory analysts can then process this point-cloud data to glean information that helps FEMA focus their recovery efforts—for example, by estimating the number of collapsed houses in an area, the volume of debris piles, and the reach of flood waters.

But quickly sending the nearly two terabytes of data from a single night's scan, or sortie, to the Laboratory for processing is a challenge. After a storm, local internet connections may be gone or spotty. When the Laboratory used this same lidar platform after Hurricane Maria in Puerto Rico, downed networks meant having to physically ship a hard drive back to Massachusetts—a more than two-day delay in getting the data into analysts' hands. When the team started the campaign in the Carolinas in mid-September, they faced the same obstacle.

This time, the obstacle was hurdled thanks to the Microelectronics Center of North Carolina (MCNC). The nonprofit organization is based out of Research Triangle Park near Durham, North Carolina, which was not directly affected by Hurricane Florence. MCNC gave the Laboratory free access to their North Carolina Research and Education Network (NCREN).

"Our state was hit hard by Hurricane Florence," said Tommy Jacobson, MCNC chief operating officer and vice president. "For MCNC's leadership, it was a quick and easy decision to enable MIT, who was in the state to assist FEMA, with access to our network resources to help however we could in making sure relief got to those that needed it."

The NCREN is North Carolina's broadband backbone, connecting more than 750 institutions including all public school districts in the state, universities and colleges, public safety locations, and other community anchor institutions. Access for the Laboratory meant rack space for equipment inside the MCNC data center. From there, MCNC provisioned a ten-gigabit IP connection from the NCREN to Internet2: an ultrafast, advanced network that connects research centers around the world. This connection gave the team the ability to upload large volumes of data daily from their equipment inside the data center back to a computing center on MIT campus that is also connected to Internet2.

Then, another 10-gigabit connection bounced the data from campus to the Lincoln Laboratory Supercomputing Center in Holyoke, Massachusetts, where the data were processed.

"The 10-gig uplink from MCNC has allowed us to transmit the data at such a higher speed that some of our uploads are done in about six to seven hours," said Daniel Ribeirinha-Braga, Humanitarian Assistance and Disaster Relief Systems, Group 44, a member of the Laboratory's data management team in this hurricane effort. "Keep in mind that this is lidar data, which we get about 1.5 to 1.9 terabytes a night of, that needs to be copied to multiple places (such as other hard drives), organized to a single solid-state drive, and then uploaded to the Laboratory from MCNC."

The collaboration between the Laboratory and MCNC came about through Matt Daggett, Group 44. He had worked at MCNC more than a decade ago and contacted the organization. "I was aware of the NCREN backbone and the data center on the MCNC campus. When it became clear that our flight operations would be based out of the Raleigh–Durham International airport, I knew MCNC would be the perfect place to get our data onto the Internet2," Daggett said.

"We were grateful that MIT sought us out to provide that help," Jacobson added.

With the data processing underway, the Laboratory has begun delivering reports to FEMA. The lidar imagery reveals things that would be impossible for FEMA to know from looking only at satellite images. "The most important difference between a satellite image and the lidar image is that you can do 3D measurements on it," said Dr. Anthony Lapadula, Group 44. "So, because it's 3D data, we can do things like tell you how big a hole in a road is, or tell you how big an elevation drop is as a result of a landslide."

One of the greatest advantages of the lidar work has been the time saved for FEMA. When someone reports damage at a specific location, FEMA can assess the damage quickly by asking Laboratory analysts to virtually "visit" the location in the point-cloud map and report what they find.

For instance, FEMA asked them to zoom in on a small town on the Lumber River in North Carolina. The town was inundated with flood waters. Analysis of the data told FEMA the extent of the flood inundation, the volume of debris piles in the town, and changes in the river's path. There were also environment questions to be answered, such as what the volume was of a coal ash pile to determine how much, if any, washed away with flood waters. They could also check in on public infrastructure, like a lock and dam along the Cape Fear River that the data showed to be completely flooded.

Completing the 40 sorties requested over the Carolinas took the team several weeks to complete, right up until Thanksgiving. The sorties covered areas down the coastline from the Cape Lookout National Seashore to Myrtle Beach and through bands stretching inland. These hundreds of miles of lidar data were processed to a resolution of approximately 25 centimeters. To put that resolution into perspective, Lapadula says that if the scanned areas were covered completely with basketballs, they would be able to precisely measure the location of each ball.

But with only one of these extremely advanced systems available for use, Laboratory staff are limited in how much area they can cover and how many disasters they can respond to with the technology. The lidar system was originally developed by Active Optical Systems, Group 106, who has been assisting Group 44 with data collection, processing, and algorithm development.

While the Laboratory has been involved in disaster recovery since the 2010 Haiti earthquake, it has never been so active in these efforts as in the past year since Hurricanes Harvey, Irma, and Maria in 2017.

"We went a decade without a major hurricane hitting the continental United States," Lapadula reflected. "Now, it's like they just keep coming."

Winds Forecast Rapid Prototype

Technology Office

MIT Lincoln Laboratory

June 8, 2018

Wind causes significant problems for aviation, yet the Federal Aviation Administration has limited tools for forecasting its impact.

Strong winds are a leading cause of air traffic delay due to their impact on aircraft spacing and runway use, yet most airports lack high-quality real-time wind forecast information and wind decision support. In response, the Laboratory is developing technology to generate improved wind forecasts through the use of data fusion and machine learning. The goal of this project is to develop an initial platform where new wind forecast products can be developed, assessed, and demonstrated. Beyond manned aircraft, this technology could benefit unmanned aircraft systems, bio/chemical dispersion modeling, wind energy, fire and rescue operations, and Department of Defense (DoD) applications.

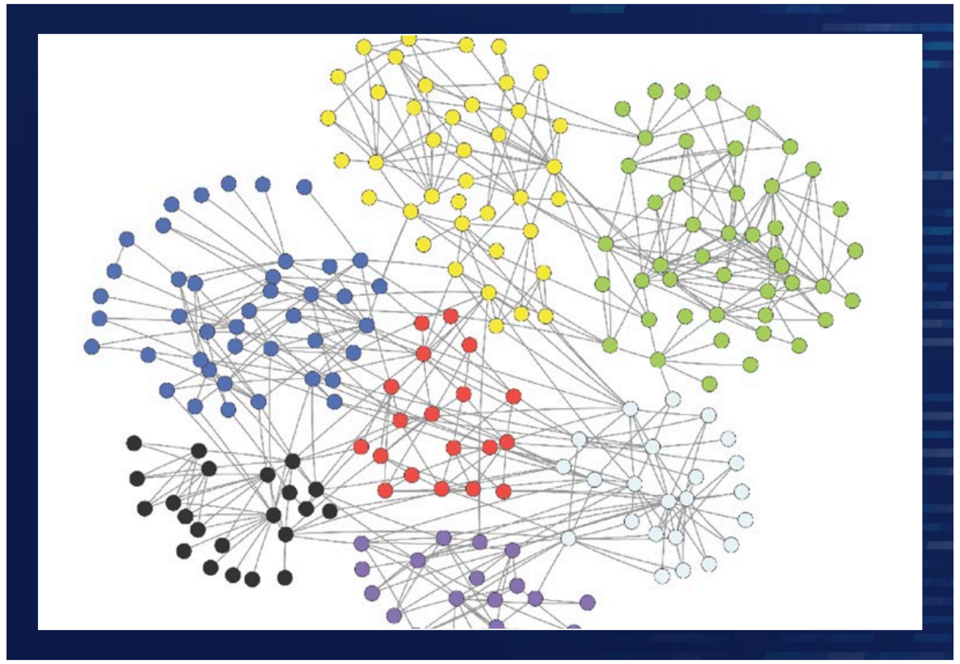
The Winds Forecast Rapid Prototype, an allocated-funded project, focuses on the highest-priority needs for the Federal Aviation Administration (FAA): runway surface winds and low altitude terminal approach. Using the Lincoln Laboratory Supercomputing Center, researchers are applying machine learning algorithms, such as convolutional neural networks and Gaussian processes, to develop forecasts that leverage the Laboratory's rich weather data sources, including aircraft-derived wind observations. These algorithms add value to less timely numerical model-based wind forecasts by intelligently fusing wind measurements in real time. This allows for more timely wind forecasts that also include measures of forecast confidence. The project will demonstrate its capabilities to potential stakeholders (e.g., FAA, Department of Homeland Security, DoD) to identify areas for further development and application. Objectives for fiscal year 2018 include building the project database and testbed, developing an initial wind forecast algorithm with confidence metrics, and performing product evaluation.

DataSToRM: Data Science and Technology Research Environment

Technology Office

MIT Lincoln Laboratory

February 26, 2018



The above is an illustration of a graph that has been analyzed by a community detection algorithm in order to reveal groups of similar vertices—here, circles with the same color belong to the same community. As an example, in social network graphs, these communities may represent groups of people sharing similar tastes in music or a common set of friends.

Information such as protein interactions or social networks can be represented with a data structure called a graph. By analyzing these graphs using specialized algorithms, complex relationships and deeper insight can be extracted from the raw information. Analyzing customer buying habits for targeted recommendations and finding new medical applications for existing drugs are two of many applications. In the last several years, the Laboratory has developed the Graph Processor, which has a unique hardware architecture that provides 100 to 1000 times better processing performance for analyzing large graph datasets. To put this performance improvement in perspective, even basic analysis, which can take several hours or days to compute in a super-computing center, can be performed in seconds or minutes.

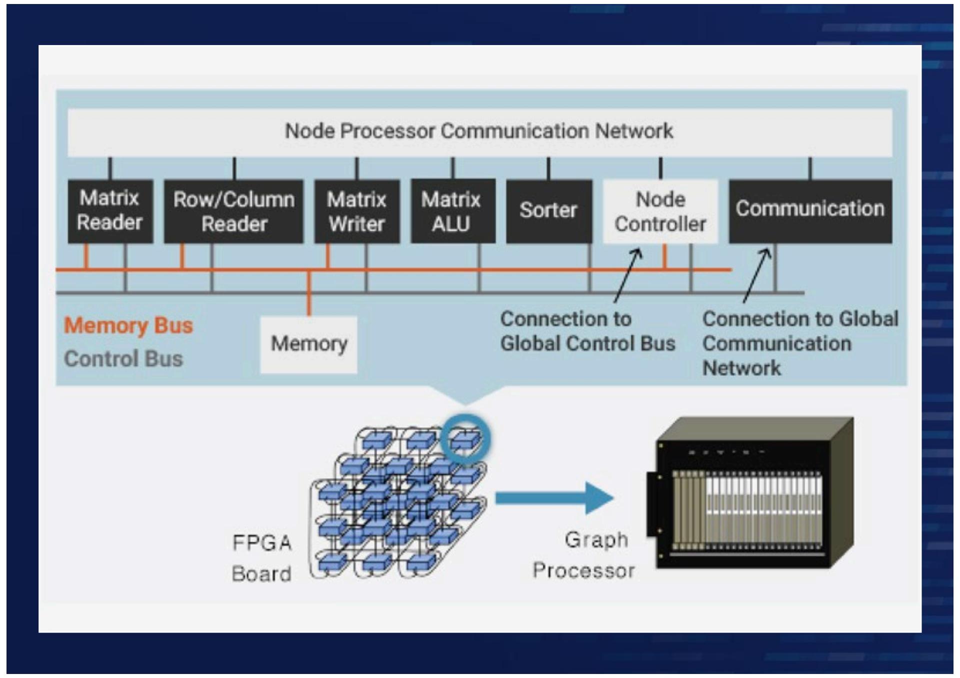
The Data Science and Technology Research Environment (DataSToRM) program is developing a software environment and algorithms to take advantage of the Graph Processor's capabilities. The environment will be closely integrated with the Lincoln Laboratory Supercomputing Center infrastructure, which will enable development of Department of Defense applications that can leverage both traditional computing resources, such as central processing units or graphics processing units, and the specialized graph hardware. The initial focus of the effort is to develop basic capabilities that will become the foundation for a full suite of graph analysis tools. Utilizing this environment, the program is looking to enable research and development of new high-performance algorithms as well as continued hardware innovation for accelerating graph applications.

Graph Processor Prototype

Technology Office

MIT Lincoln Laboratory

February 10, 2017



Advanced graph processor prototype development.

Financial transactions, social networking, and Internet traffic can all be represented as sparse, highly interconnected graphs. Algorithms can analyze these graphs to detect cyber network intrusions, terrorist social networks, and traffic patterns that indicate a hostile environment. While graph algorithms work well with moderate-size databases, even the fastest modern processors have difficulty providing sufficient throughput for very large databases (billions to trillions of nodes). The difficulty is random memory access and computation dealing with indices, vertices, and edges of the large graphs that soak up a large fraction of the processor resources. Consequently, graph computation can be slowed by several orders of magnitude.

In order to achieve significantly better graph computation performance, an advanced multiprocessor architecture has been developed that is optimized for analysis of large databases. Most graph operations can be implemented as sparse matrix operations. This processor, which is optimized for sparse matrix algebra operations, can be hundreds of times more efficient than conventional processors. To address the memory access and inter-processor communication challenges, the new graph processor architecture does not use cache memory, but rather a specialized network architecture to connect thousands of processing nodes. Each processor node is designed to be capable of a communication rate of over one terabit per second. A specialized systolic sorter module handles most of the overhead computation associated with graph indices, vertices, and edges.

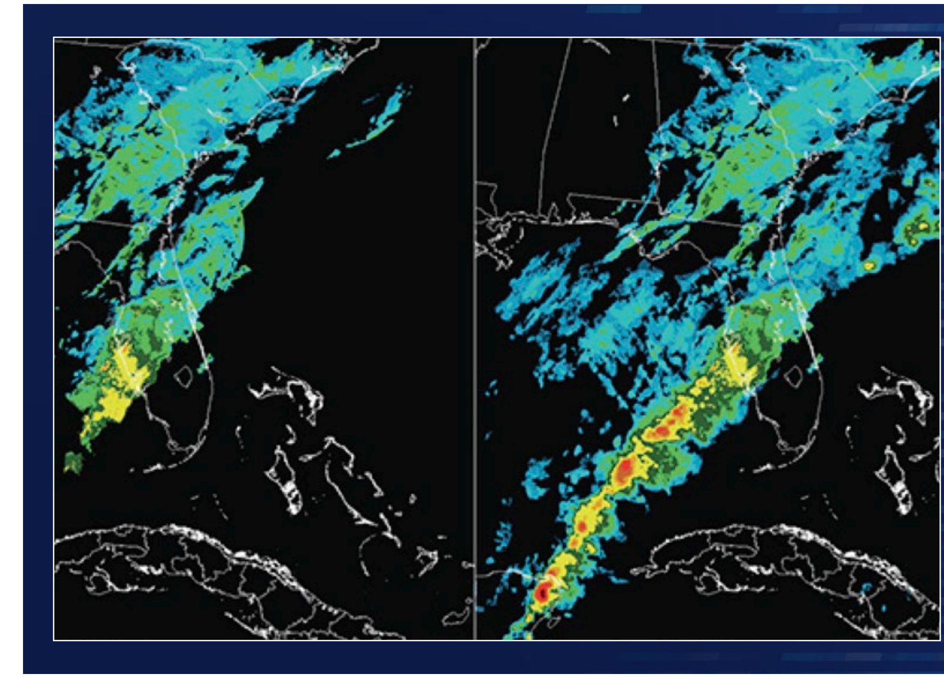
An 8-node graph processor prototype using commercial field-programmable gate array (FPGA) boards demonstrated a factor of ten times better performance than commercial processors. The effort is under way to demonstrate even higher performance with 64-node prototype systems using custom FPGA boards, as the performance advantage increases with system size.

Expanding Air Traffic Controllers' View of Offshore Weather

Homeland Protection and Air Traffic Control

MIT Lincoln Laboratory

October 14, 2016



Current land-based weather radar coverage (left) compared to the improved accuracy of the Offshore Precipitation Capability (right). Both images depict storm height according to the color bar at the far right.

Air traffic controllers depend on weather radar for information that enables them to safely reroute aircraft around storms that may cause hail, turbulence, icing, and other hazardous conditions. While land-based radar systems provide reliable weather data covering the continental United States, their effectiveness in capturing data about offshore and over-the-ocean weather systems is limited or nonexistent. Controllers who manage the large number of aircraft traversing oceanic sectors of the National Airspace System (NAS) are faced with an imperfect weather picture that can threaten the safety of air travel and can lead to inefficient routing of aircraft in the NAS.

Researchers at the Laboratory, working with the Federal Aviation Administration, have designed the Offshore Precipitation Capability (OPC) to address this lack of airspace situational awareness. The OPC provides radar-like depictions of precipitation intensity and storm height for regions where radar data are inadequate or unavailable. OPC creates the depictions by fusing together various non-radar data sources—lightning detections, information from geostationary satellites, and outputs from numerical weather prediction models—available within these offshore regions.

"Lightning is a strong indicator of convective weather, that is, weather capable of generating the powerful updrafts that can cause severe turbulence and other safety hazards like damaging hail," explained Dr. Mark Veillette, Air Traffic Control Systems, Group 43, team lead. "In OPC, we use a global lightning dataset that covers areas far offshore and over oceans."

Geostationary satellites positioned around the globe provide visible and infrared imagery of much of the Earth, while numerical weather prediction models simulate many meteorological measurements, including environmental temperature, pressure, and humidity, and other relevant parameters, such as radar reflectivity, rain rate, and convective cloud-top height.

Data from these non-radar sources are combined by using a supervised machine learning framework. The OPC model is "trained" to generate products normally derived from actual weather radar measurements. Results from OPC are blended with outputs from existing radar-based systems to create seamless mosaics of weather systems that extend into offshore and oceanic regions.

"OPC fills in gaps in radar coverage on controllers' displays of offshore regions beyond radar range. The output data from OPC are presented in a style and format that matches what air traffic controllers are used to seeing on existing weather processing systems," said Veillette.

OPC also applies a motion-tracking algorithm to estimate storm motion. This estimated motion is used to spatially shift the features obtained from satellite imagery and numerical models so that OPC can generate up-to-date mosaics in between satellite data updates.

To evaluate the capability's performance, the research team compared OPC-generated displays of weather to over-land radar displays produced by Next Generation Radar and to offshore and oceanic radar displays produced by the spaceborne radar on board NASA's Global Precipitation Measurement Mission's Core Observatory Satellite. "The comparisons showed that over land OPC can provide a proxy for radar by accurately identifying storm location, intensity, and height and that OPC accurately depicts precipitation intensity and storm height in offshore regions," said Christopher Mattioli, Group 43, a member of the OPC team.

Many other systems designed to provide air traffic controllers with a view of maritime weather conditions are based only on lightning flash densities or solely on infrared satellite imagery. "The key innovation of OPC is its data-driven modeling framework that allows the system to combine multiple heterogeneous meteorological datasets into a single, accurate, and familiar-looking product," said Veillette. "In addition, by including high-resolution inputs, such as visible satellite imagery and lightning density, OPC is able to resolve some finer-resolution storm characteristics that are missed by systems that rely only on infrared satellite or lightning data."

Both the NOAA and FAA have begun assessments of the OPC system's accuracy and suitability for deployment to air traffic control centers. While developed primarily for use by U.S. air traffic controllers working in Oceanic Air Route Traffic Control facilities, work is also being done to extend OPC coverage for broader applications such as hurricane tracking, disaster response, and global weather forecasting to monitor all areas lacking radar coverage.

Building Enabling Technologies

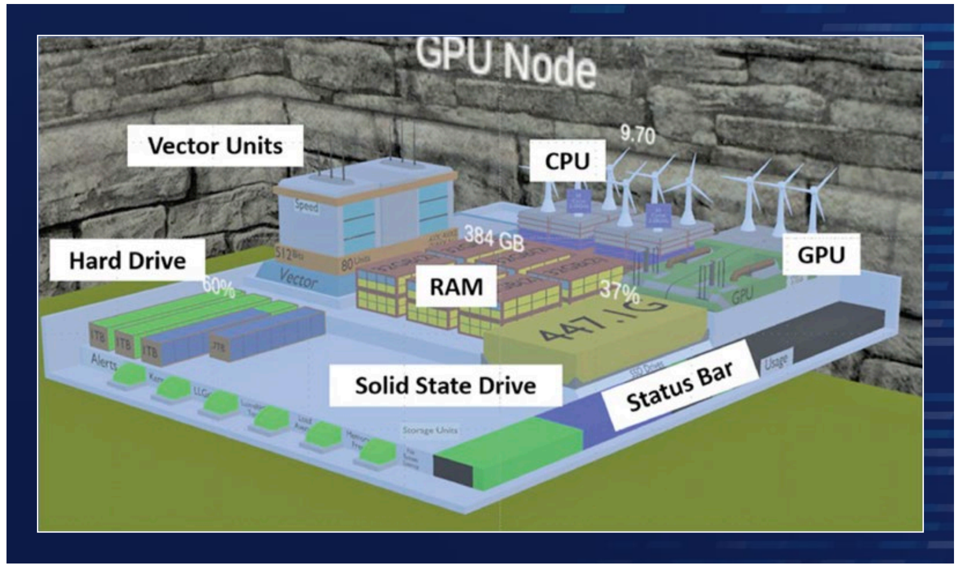
The Lincoln Laboratory Supercomputing Center (LLSC) staff are advancing the capabilities of our supercomputing system by developing new technologies to improve the system's performance. The center provides interactive, on-demand parallel computing that allows researchers from across the Laboratory to augment the processing power of their desktop systems in order to process large sets of sensor data, create high-fidelity simulations, and develop new algorithms. We are also collaborating with researchers from MIT on several supercomputing initiatives.

MIT Researchers Turn to Unity 3D Game Engine for Supercomputer Diagnostics

Division 5: Cyber Security and Information Sciences

Lincoln Laboratory

October 29, 2021



The above figure is a detailed representative view of a supercomputer node, with labels that show components and system load.

Rakeesearchers from MIT—including more than a dozen members of the Lincoln Laboratory Supercomputing Center (LLSC)—aim to speed up the process of monitoring, diagnosing, and fixing problems with multi-billion-dollar supercomputers. They plan to utilize Unity, a 3D game engine, to visualize the hardware.

Supercomputers are an extremely complex collection of hardware, as they operate with thousands of interconnected systems. However, there can be bottlenecks within the system that can result in reduced performance, as well as lost time while diagnosing those bottlenecks.

The average supercomputer has many components in the system. Each part of the system is called a node, and each node contains a specific set of hardware components. As a basic explanation, some nodes are designed for storing data while other nodes are for computing. The compute nodes typically contain processors and main system memory.

Engineers continuously test the machine during the installation process, encountering problems along the way. There could be storage, processor, and even networking problems in the system, and diagnosing the root cause can be difficult with such large-scale systems. For example, the upcoming Frontier supercomputer should have around 100 racks containing tens of nodes each, resulting in thousands of nodes to diagnose and monitor.

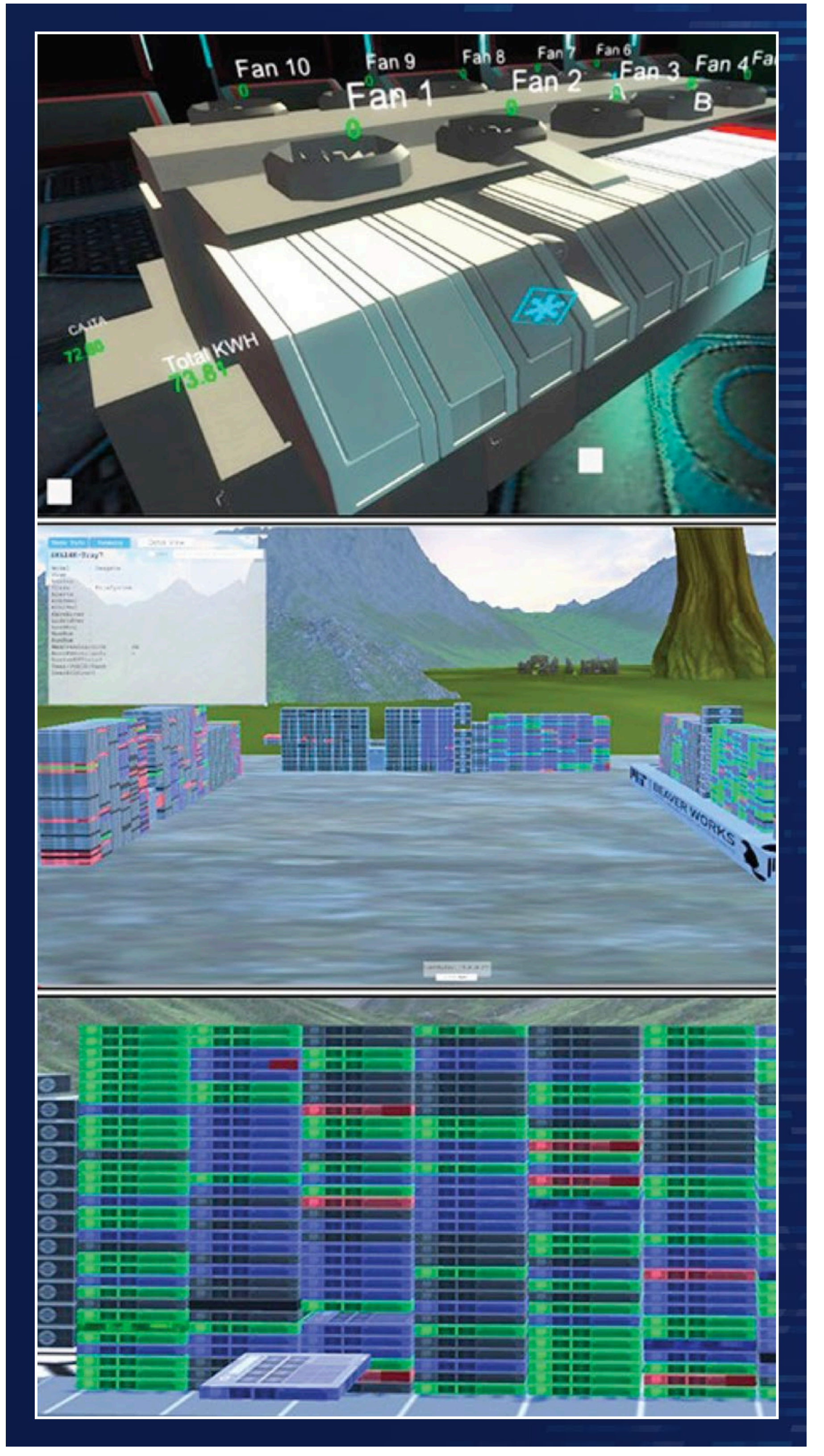
To help streamline these types of efforts, researchers from the Laboratory and MIT have developed a new technology to visualize node monitoring, offering real-time system reporting in the Unity 3D game engine found in many video games. Called the MM3D, it is a part of Data Center Infrastructure Management (DCIM) tools developed by the MIT SuperCloud division.

"The combination of supercomputing analytics and 3D gaming visualization enables real-time processing and visual data display of massive amounts of information that humans can process quickly with little training," said Dr. Jeremy Kepner, Laboratory Fellow, LLSC, while summarizing a paper that the researchers recently published to nonprofit arXiv. "Our system fully utilizes the capabilities of modern 3D gaming environments to create novel representations of computing hardware, which intuitively represent the physical attributes of the supercomputer while displaying real-time alerts and component utilization."

This means that this 3D engine can display component utilization and any alerts from the system in real time. For instance, if an alert pops up, a specific node could be overheating, and the system administrator would be alerted instantly in the 3D engine application.

While this is not a commercial application yet, this academic project could represent a step forward in the supercomputer monitoring department that helps ease system administration. Given that academic institutions share their work with other entities, it may be only a matter of time before we see a similar solution in the wild.

This story was originally authored by Aleksandar Kostovic and was published online at www.tomshardware.com. It has been adapted for the Bulletin.



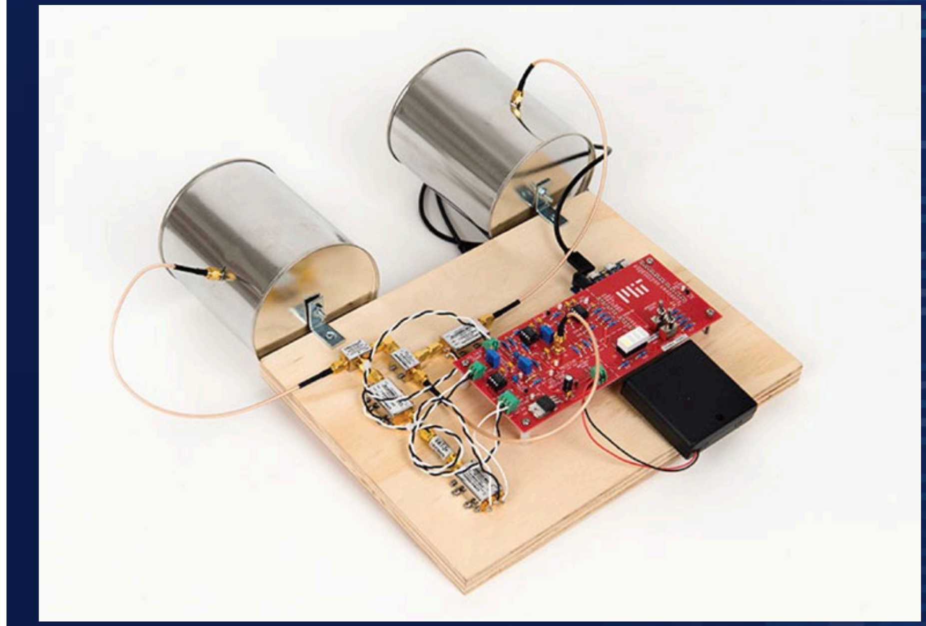
The above figure shows in-game views (from top to bottom) of the Hewlett Packard Enterprise EcoPod, node layout, and a display that provides situational awareness of high performance computing components in real time.

LLx Project Seeks to Improve Online Hands-On Learning

Division 5: Cyber Security and Information Sciences

MIT Lincoln Laboratory

December 6, 2019



The Build a Radar course, one of the classes offered on the Laboratory's education course platform (called LLx), aims to give students hands-on experience in an online setting.

Online learning has been around for decades, but educators still struggle to adapt certain hands-on subjects to the web. For example, electrical engineering requires both theoretical and practical knowledge because understanding how to build a system is not the same as actually building one. When students first attempt to assemble a physical system, they often need access to instructors or experts who can answer questions and help troubleshoot issues. In an online environment, where instructors may not be readily available, guiding students as they translate their theoretical knowledge into practice is a difficult task.

At the Laboratory, a team of researchers has been working on ways to close this practical learning gap through the Lincoln Laboratory Online Courses (LLx) project. The project was born from a desire to adapt the Laboratory's internal course offerings for a wider audience. Its goal is to identify best practices for online hands-on learning in order to provide unique Laboratory classes to sponsors, students, and the general public as self-paced, massively open online courses (MOOCs).

The program currently has courses that are offered internally and externally, including a course where students learn how to build a radar. Additionally, the project supports the online portion of the Beaver Works Summer Institute by providing material that prepares rising seniors for the four-week program.

Dr. Julia Mullen, Supercomputing Center (LLSC), leads the LLx program and believes that helping students build intuition about a subject is essential for successful learning: "If learning is defined as being able to transfer knowledge and apply theory to new applications, then building intuition about a subject matter is key to understanding how to transfer knowledge to new applications and cross-disciplinary projects."

In academia, intuition is built within a laboratory setting, where instructors can demonstrate troubleshooting techniques and guide students through the engineering process by asking a series of questions. The challenge in an online environment, particularly in self-paced courses, is to replicate this one-on-one help. Traditionally, online courses have provided theoretical material, while practical applications have been left for in-person sessions. The Laboratory is unique in developing techniques for bridging the physical-digital learning gap for online, self-paced courses.

Following standard best practices for online learning, the courses are designed to present theoretical content in modular blocks that break down difficult topics into small pieces and incorporate exercises and checks for understanding. To help develop intuition for the hands-on portion of the courses, the team has designed the build, test, and experiments sections using a collection of high-resolution, clearly annotated images of the final system. Students can use these while building their own system.

To help students get an idea of the type of results they should be seeing, the team has included simulations of experiments that include good and bad results. "At every stage, there's a check. Does it work? Great, go to the next step. If doesn't work, then here are things to try," said Mullen. "The goal is help learners develop troubleshooting skills, learn how to design good tests for their system or program, and develop confidence in their ability to track down and rectify issues as they arise."

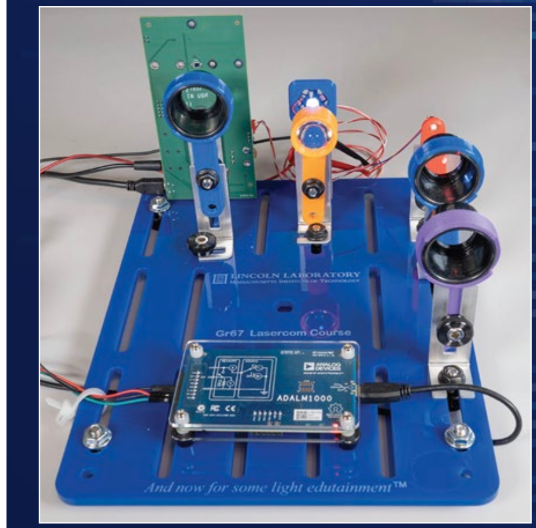
The first course developed in the LLx program was the High Performance Computing class for users of the LLSC systems—offered both internally and externally in MOOC format. The success of this course prompted the team to consider other Laboratory courses that would interest a wider audience. A prime candidate was the Build a Radar course.

"We have been running various versions of the Build a Radar course since 2011, and since it has gained a lot of interest, it has consumed a lot of staff hours teaching it to different student and sponsor groups," said Kenneth Kolodziej, Radio Frequency Technology, Group 86, who is the current lead of the Build a Radar course. "The MOOC version helps us drastically reduce the amount of staff time for each course we offer, and it has even opened the opportunity for others to take it remotely with only the online instruction."

The LLx team is currently working with Laboratory staff to develop a new Build a Lasercom System course, which will be launched on campus for students next month and internally in the spring.

Robert Schulein, Optical Communications Technology, Group 67, who is leading development of the Build a Lasercom System course, said that he hopes the course will help onboard new staff and deepen sponsor interactions. "One of the goals is to hold the course for folks with less background in the technology area. After introducing the technology through the course, a relationship is established that can translate over to demonstrating real Laboratory technology."

The team is also adapting the LLSC Practical Machine Learning Course and using it to explore methods of personalizing course content and activities. The end goal is to allow students to begin courses with whatever module best fits their experience level or pick and choose which modules to complete based on their experience and goals.



As government, business, and personal users increasingly rely on the internet task for researchers. Supercomputing is an untapped resource that can help them in this endeavour.

MIT Lincoln Lab Offers Advice for Delivering On-Demand HPC

John Russell

HPC wire

March 11, 2019

The 40,000-core interactive parallel computing system provides high-performance, on-demand computing resources to the Laboratory community.

Looking for advice on how to deliver HPC to a diverse science user community? MIT's Lincoln Laboratory has just posted a new paper – Lessons Learned from a Decade of Providing Interactive, On-Demand High Performance Computing to Scientists and Engineers – intended to fill the bill.

"For over a decade, MIT Lincoln Laboratory has been supporting interactive, on-demand high performance computing by seamlessly integrating familiar high productivity tools to provide users with an increased number of design turns, rapid prototyping capability, and faster time to insight. In this paper, we discuss the lessons learned while supporting interactive, on-demand high performance computing from the perspectives of the users and the team supporting the users and the system," write the authors.

The experiences outlined are derived from Lincoln Lab's development and operation of "SuperCloud" (specs not given).

It is described as "a fusion of the four large computing ecosystems: supercomputing, enterprise computing, big data and traditional databases into a coherent, unified platform. The MIT SuperCloud has spurred the development of a number of cross-ecosystem innovations in high performance databases; database management; data protection; database federation; data analytics; dynamic virtual machines and system monitoring."

Pointedly, the advice is aimed not at supercomputer centers fixed on maximizing machine utilization but at organizations seeking an ROI that blends ease of use with efficiency in an on-demand environment.

"Virtually all HPC centers report the percentage of system utilization as their metric of success. This choice of metrics often leads to queuing systems and user behavior designed to feed the system with the type of jobs that yield high utilization. However, these utilization-based queuing practices are often at odds with rapid prototyping of algorithms and simulations, exploration of large datasets and real time steering of complicated multi-physics simulations," according to the report.

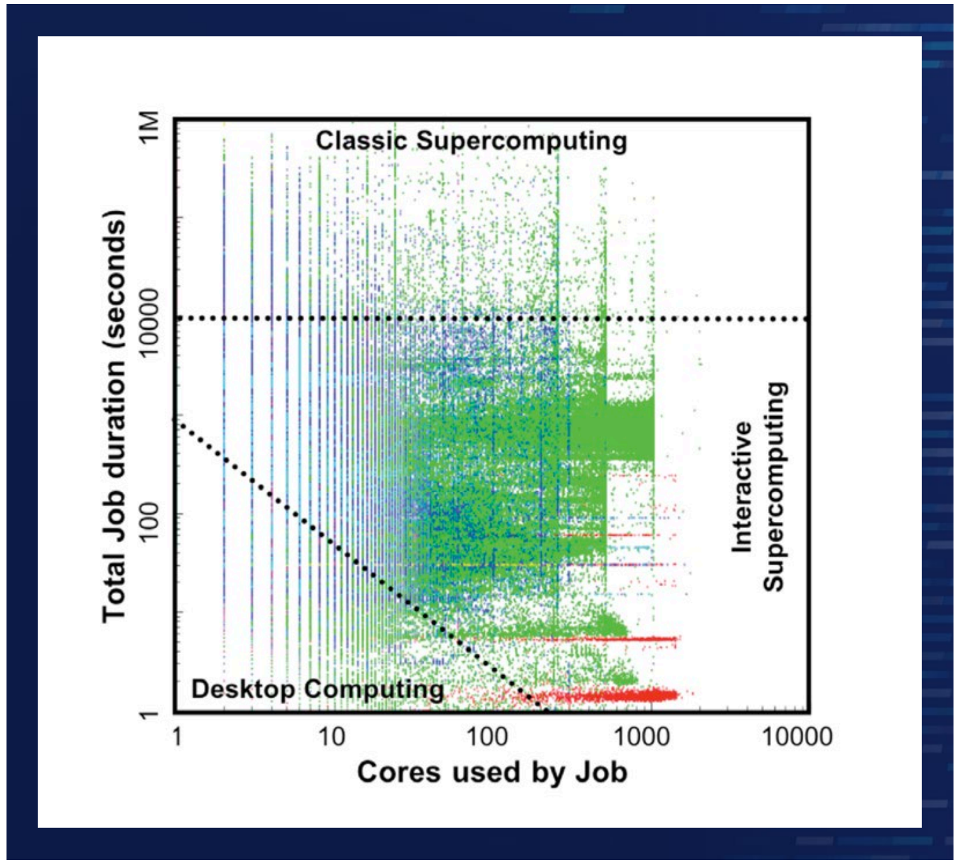
Lincoln Lab relied on "DARPA's High Productivity Computing System program where a productivity metric was developed as part of a larger analysis of HPC Return On Investment for a broad range of applications and research domains."

The report is an interesting blend of history and blueprint. Among the common use cases cited are: algorithm development for sensor signal processing; development of multiple program, multiple data (MPMD) real time signal processing systems; high throughput computing for aircraft collision avoidance system testing; and biomedical analytics to develop medical support techniques for personnel in remote areas.

The authors write, "Unlike traditional HPC applications, most of these capabilities involve prototyping efforts for multi-year, but not multidecade, mission-driven programs making it even more important that researchers are able to use familiar interactive tools and achieve a greater number of design cycles per day. To enable an interactive high performance development environment, our team turned the traditional HPC paradigm on its head.

"Rather than providing a batch system, training in MPI, and assistance porting serial code to a supercomputer, we developed the tools and training to bring HPC capabilities to the researchers' desktops and laptops. As common use cases and staff computational preparation change, we routinely update our tools so that we can provide relevant interactive research computing environments."

You get the general picture. The full report is on arXiv (https://arxiv.org/pdf/1903.01982.pdf)



A notional view of research computing regimes. The Desktop Computing region involves jobs that execute in less than five minutes, often for exploration and debugging. The Classic Supercomputing jobs are those that run for over three hours. Finally, the Interactive Supercomputing jobs run for between five minutes and three hours and are usually jobs that involve rapid turnarounds for insight.

[i] Lessons Learned from a Decade of Providing Interactive, On-Demand High Performance Computing to Scientists and Engineers; Julia Mullen, Albert Reuther, William Arcand, Bill Bergeron, David Bestor, Chansup Byun, Vijay Gadepally, Michael Houle, Matthew Hubbell, Michael Jones, Anna Klein, Peter Michaleas, Lauren Milechin, Andrew Prout, Antonio Rosa, Siddharth Samsi, Charles Yee, and Jeremy Kepner, https://arxiv.org/pdf/1903.01982.pdf

Supercomputers Can Spot Cyber Threats

Lincoln Laboratory researchers have developed a technique to compress hours of internet traffic into a bundle that can be analyzed for suspicious behavior.

Anne McGovern

MIT Lincoln Laboratory

February 26, 2019

Dr. Vijay Gadepally stands in the Laboratory's in-house supercomputing center. Gadepally is part of a team that leveraged supercomputing to address cyber security issues.

Identifying cyber security threats from raw internet data can be like locating a needle in a haystack. The amount of internet traffic data that is generated in a 48-hour period, for example, is too massive for one or even 100 laptops to process into something digestible for human analysts. Because of this issue, analysts rely on sampling to search for potential threats—they select small segments of data to look at in depth, hoping to find suspicious behavior. While this type of sampling may work for some tasks, such as identifying popular IP addresses, it is inadequate for finding subtler types of threatening trends.

"If you're trying to detect anomalous behavior, by definition that behavior is rare and unlikely," said Dr. Vijay Gadepally, Senior Staff, Supercomputing Center. "If you're sampling, it makes an already rare thing nearly impossible to find."

Gadepally is part of a research team at the Laboratory that believes supercomputing can offer a better method for identifying these subtle trends—one that grants analysts access to all pertinent data at once. In a recently published paper, the team successfully condensed 96 hours of raw, 1-gigabit network link internet traffic data into a query-ready bundle. They accomplished this by running 30,000 cores of processing (equal to about 1,000 laptops) at the Lincoln Laboratory Supercomputing Center (LLSC) located in Holyoke, Massachusetts. The data is stored in the MIT SuperCloud, where it can be accessed by those with an account.

"[Our research] showed that we could leverage supercomputing resources to bring in a massive quantity of data and put it in a position where a cyber security researcher can make use of it," Gadepally explained.

An example of the type of threatening activity that requires analysts to dig in to such a massive amount of data is that from command-and-control (C&C) servers. These servers issue commands to devices infected with malware in order to steal or manipulate data. Gadepally likens their pattern of behavior to that of spam phone callers—while a normal caller might make and receive an equal number of calls, a spammer would make millions of more calls than they receive. It's the same idea for a C&C server, and this pattern can be found only by looking at much more data over a long period of time.

"The current industry standard is to use small windows of data, where you toss out 99.99%," Gadepally said. "We were able to keep 100% of the data for this analysis."

The team plans to spread the word about their ability to compress such a large quantity of data. They hope that analysts will take advantage of this resource to take the next step in cracking down on threats that have so far been elusive.

They are also working on ways to better understand what "normal" internet behavior looks like as a whole, so that threats can be more easily identified.

"Detecting cyber threats can be greatly enhanced by having an accurate model of normal background network traffic," said Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center, who is spearheading this new research direction. Analysts could compare the internet traffic data they are investigating with these models in order to bring anomalous behavior to the surface more readily. "Using our processing pipeline, we are able to develop new techniques for computing these background models."

As government, business, and personal users increasingly rely on the internet for their daily operations, maintaining cyber security will remain an essential task for researchers. Supercomputing is an untapped resource that can help them in this endeavour.

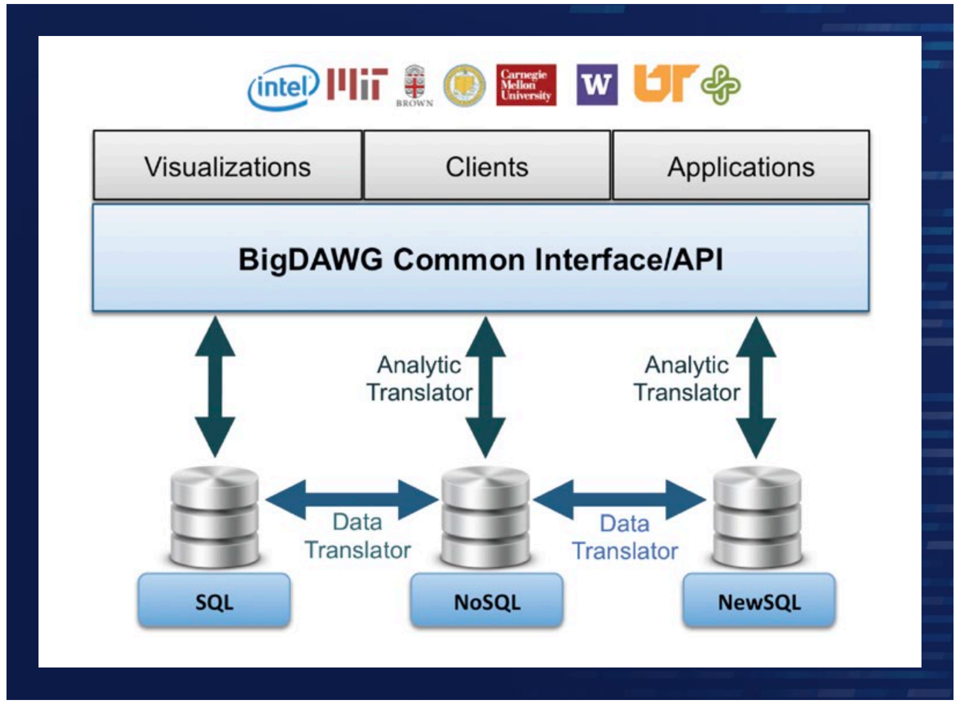
This research is sponsored by the Technology Office with Line funding.

Enabling Massive Computation and Resiliency in the Internet-of-Things Era

Technology Office

MIT Lincoln Laboratory

June 16, 2017



To address challenges in working with diverse datasets, Laboratory staff and university collaborators created the BigDAWG polystore system. BigDAWG is a new kind of database management system that allows users to integrate and run analytics on data that is spread across disparate database engines.

The Internet of Things (IoT), an ever-growing network of physical devices connected to the Internet, brings a unique set of challenges to the Department of Defense (DoD). These challenges include the billions of connected devices, the tremendous diversity of the data being generated by these devices, and the varied defenses required to protect the applications.

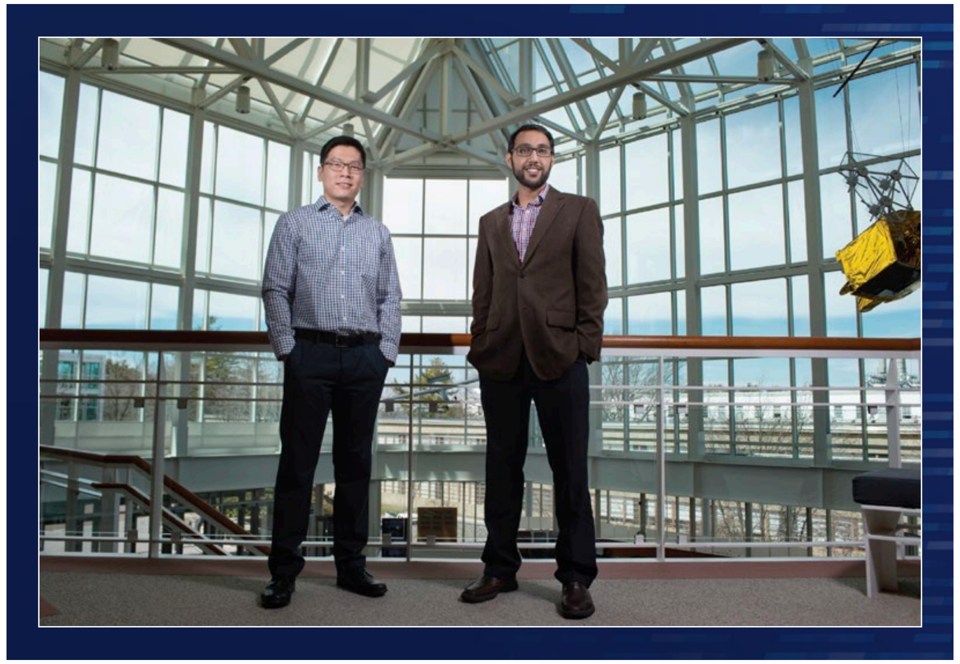
This Line-funded program has been developing fundamental technologies to address these challenges. For devices, the team has been working with several groups to develop hardware and software tools to process the high volume of data generated by these devices. These efforts include understanding how new processing paradigms—such as Group 102's graph processor, NVIDIA graphics processing units, and Intel's KNL processors—will play a role in the Lincoln Laboratory Supercomputing Center's future. To address the diversity challenge, the Laboratory team, in collaboration with university partners, developed the first-ever polystore database system called BigDAWG. BigDAWG greatly simplifies data analytics by allowing users to integrate data that is spread across disparate database engines. To address defense challenges, the team is developing algorithms resilient to certain cyber attacks. These algorithms are co-designed with existing algorithms in order to allow DoD sensors to operate in adversarial environments.

Staff Release Open Source Software for BigDAWG Polystore System

Cyber Security and Information Sciences

MIT Lincoln Laboratory

April 14, 2017



Kyle O'Brien (left), Group 104, and Dr. Vijay Gadepally, Supercomputing Center, are part of the team behind a new database management system called BigDAWG.

Working with a vast variety of data is often a challenge for organizations. Imagine, for example, all of the different types of medical data that hospitals record—such as vital sign measurements, results of laboratory tests, descriptions of patients' symptoms, etc. It would be difficult to organize such diverse data into a single database engine. But, the ability to connect multiple engines through a single user interface would greatly improve an organization's ability to access all the data, and help them analyze it without having to know which specific database to probe.

The Big Data Analytics Working Group (BigDAWG) polystore system, developed in part by Laboratory staff, is simplifying the integration and analytics of data. The system is a prototype of a "polystore" concept— a database management system that connects multiple, heterogeneous database engines behind a single programming interface. The interface allows users to enter queries, which BigDAWG automatically optimizes and responds to by moving data across database engines without user intervention. The team recently released the first open source version of the BigDAWG software.

"BigDAWG is beneficial to anyone seeking a simpler way to use data that spans multiple data models," said Dr. Vijay Gadepally, Supercomputing Center. He has led the BigDAWG effort for close to two years, collaborating with Kyle O'Brien, Intelligence and Decision Technologies, Group 104; Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center; and a team of researchers from a dozen universities around the country. The collaboration is fostered through the Intel Science and Technology Center for Big Data—a series of research collaborations that Intel is establishing with U.S. universities to prototype revolutionary technologies. "I'm very excited that we have essentially opened up a new research area with BigDAWG," Gadepally said.

Traditionally, if a user wants to access all of their data that is split into separate database engines, they have to code multiple connectors to link the data into a single data store. The BigDAWG architecture greatly reduces the need to write single connectors between each database the user wants to connect. The architecture is constructed in layers. At the top, the interface receives a user query, and passes it below to the appropriate "island" for execution. An island consists of a query language and a specific data model, and is associated with one or more database engines. A "shim" then connects an island to the engines below. "A shim basically serves as a translator that maps queries, expressed in terms of the operations defined by an island, into the native query language of a particular storage engine," Gadepally explained. Software components called "casts" move data across database engines as needed, and users are presented with the results of their query.

So far, the team has tested BigDAWG on two real-life use cases. The first was in collaboration with MIT's Chisholm Lab, whose datasets encompass ocean metagenomic data, including millions of bacterial DNA samples and associated data like the depth, salinity, and temperature of water the samples were collected from. The second was on an intensive-care unit dataset named MIMIC II, which comprises de-identified health data collected from about 40,000 critical care patients from Beth Israel Deaconess Hospital. In both cases, BigDAWG provided researchers with real-time support for streams of diverse data. "These two examples do a great job of simulating the volume, velocity, and variety of real applications. Working on these problems let us work through real issues such as messy and missing data, unknown or unquantifiable performance metrics, and large-scale data," Gadepally said.

The open source release allows users to download and begin testing the BigDAWG polystore system with real data. Scripts included in the release allow users to download publicly available parts of the MIMIC II dataset and load them into three database engines. "Patient history data is inserted into PostgreSQL, physiologic data is inserted to SciDB, and free-form text data is inserted into Accumulo. In a few minutes, users can have three databases running and issue cross-engine queries to them without having to install the databases permanently," O'Brien said.

The team hopes that research analysts, data scientists, and database administrators will try BigDAWG with MIMIC II and share feedback. Their goal is to continue developing the system to run as simply and automated as possible. "We recently organized a workshop at the IEEE Big Data conference, and it was amazing to see our work being followed by top researchers in three continents," Gadepally said. The team's broader goal is to further exploration and innovation in the field.

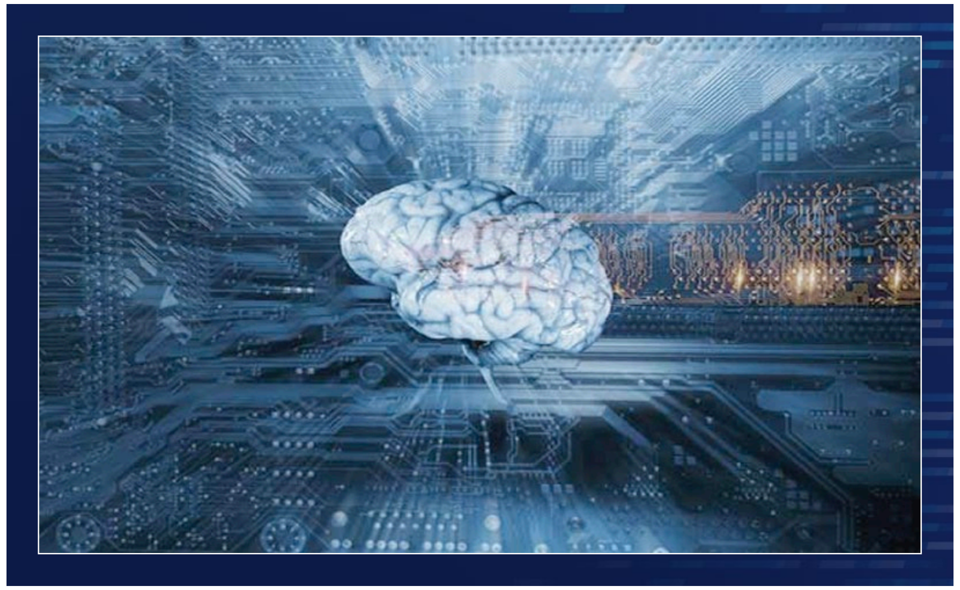
"This release was just our first and has opened up a number of fascinating problems," Gadepally said. "We hope to have many more releases in the future. We want to continue expanding the system based on user feedback, and apply our work to new and challenging problems." BigDAWG is available for download at bigdawg.mit.edu.

Capitalizing on Machine Learning—from Life Sciences to Financial Services

Dell

HPC wire

December 26, 2016



In the life sciences arena, researchers are leveraging machine learning in their work to drive groundbreaking discoveries that may help improve the health and wellbeing of people.

The promise of machine learning has a science fiction flavor to it: computer programs that learn from their experiences and get better and better at what they do. So is machine learning fact or fiction?

The global marketplace answers this question emphatically: Machine learning is not just real; it is a booming field of technology that is being applied in countless artificial intelligence (AI) applications, ranging from crop monitoring and drug development to fraud detection and autonomous vehicles. Collectively, the global AI market is expected to be worth more than $16 billion by 2022, according to the research firm MarketsandMarkets.[1]

In the life sciences arena, researchers are leveraging machine learning in their work to drive groundbreaking discoveries that may help improve the health and wellbeing of people. This research is taking place around the world.

In the United States, for example, researchers at the MIT Lincoln Laboratory Supercomputing Center (LLSC) are applying the power of machine learning algorithms and a new Dell EMC Top500 supercomputer to ferret out patterns in massive amounts of patient data collected from publicly available sources. These scientific investigations could potentially lead to faster personalized treatments and the discovery of cures.

In one such project, researchers affiliated with the LLSC used the new Top500 supercomputer to gain insights from an enormous amount of data collected from an intensive care unit over 10 years. “We did analytics and analysis on this data that was not possible before,” says LLSC researcher Vijay Gadepally. “We were able to reduce two to 10 times the amount of time taken to do analysis, such as finding patients who have similar waveforms.” Watch the video.

In China, Dell EMC is collaborating with the Chinese Academy of Sciences on a joint artificial intelligence and advanced computing laboratory. This lab focuses on research and applications of new computing architectures in the fields of brain information processing and artificial intelligence. Research conducted in the lab spans cognitive function simulation, deep learning, brain computer simulation, and related new computing systems. The lab also supports the development of brain science and intellect technology research, promoting Chinese innovation and breakthroughs at the forefront of science. In fact, Dell China was recently honored with an “Innovation Award of Artificial Intelligence in Technology & Practice” award in recognition of the collaboration. Read the blog.

In Europe, meanwhile, the University of Pisa is using deep learning technologies and systems from Dell EMC for DNA sequencing, encoding DNA as an image. The examples like these could go on and on, because the application of machine learning techniques in the life sciences has tremendous momentum in laboratories around the world.

So why does this matter? In short, because we need to gain insights from massive amounts of data, and this process requires systems that exceed human capabilities. Machine learning algorithms can dig through mountains of data to ferret patterns that might not otherwise be recognizable.

Moreover, machine learning algorithms get better over time, because they learn from their experiences.

In the healthcare arena, machine learning promises to drive life-saving advances in patient care. “While robots and computers will probably never completely replace doctors and nurses, machine learning/deep learning and AI are transforming the healthcare industry, improving outcomes, and changing the way doctors think about providing care,” notes author Bernard Marr, writing in Forbes. “Machine learning is improving diagnostics, predicting outcomes, and just beginning to scratch the surface of personalized care.”[2]

Machine learning is also making wide inroads in diverse industries and commercial applications. MasterCard, for example, is using machine learning to detect fraud, while Facebook is putting machine learning technologies to work via a facial recognition algorithm that continually improves its performance. Watch the video.

“Machine learning has become extremely popular,” says Jeremy Kepner, Laboratory Fellow and head of the MIT Lincoln Laboratory Supercomputing Center. “Computers can see now. That’s something that I could not say five years ago. That technology is now being applied everywhere. It’s so much easier when you can point a camera at something and it can then produce an output of all the things that were in that image.” Watch the video.

Here’s the bottom line: Machine learning is no longer the stuff of science fiction. It’s very real, it’s here today and it’s getting better all the time—in life sciences and fields beyond.

[1] MarketsandMarkets. “Artificial Intelligence Market by Technology (Deep Learning, Robotics, Digital Personal Assistant, Querying Method, Natural Language Processing, Context Aware Processing), Offering, End-User Industry, and Geography – Global Forecast to 2022.” November 2016.

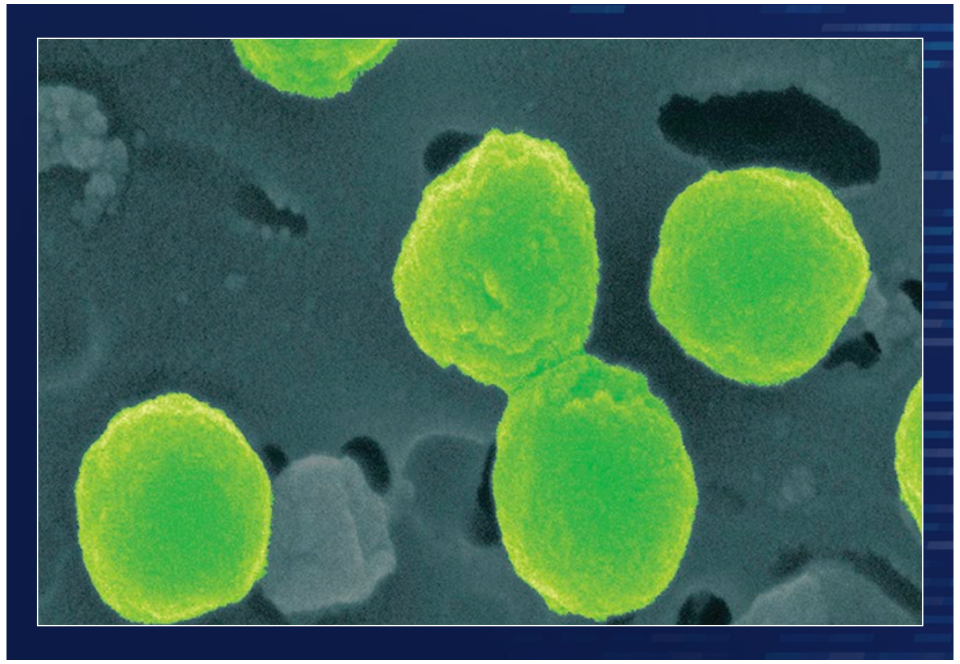
[2] Bernard Marr. “How Machine Learning, Big Data And AI Are Changing Healthcare Forever.” Forbes. Sept. 23, 2016.

A Tiny Organism with a Big Data Problem

Cyber Security and Information Sciences

MIT Lincoln Laboratory

September 9, 2016



Prochlorococcus is the smallest and most abundant photosynthetic organism on earth. (Photo by Anne Thompson, Chisholm Lab).

Among the rows of specimens in the Chisholm Lab at MIT lies a sea-dwelling bacterium named Prochlorococcus marinus. Despite its tiny size (less than a micrometer wide), it's an organism of global importance. A billion billion billion of these cells roam the ocean. Together, consuming carbon dioxide and releasing oxygen during photosynthesis, Prochlorococcus play an enormous role in the balance of these gases on earth.

Yet, since its discovery in the 1980s, there is still relatively little known about the organism and its relationship to the ocean's metabolism. Perhaps the most pressing questions on researchers' minds are how Prochlorococcus will respond to the warming oceans, and what the effects of climate change on these microbes means for the rest of life on the planet. These questions may be answered faster and sooner with the help of a new Big Data architecture, called "BigDAWG," under development by Intel Science and Technology Center (ISTC) for Big Data in collaboration with the Chisholm Lab.

In recent decades, researchers have skimmed the seas collecting Prochlorococcus samples, and have sequenced the organisms' genomes back at the lab. Meanwhile, advances in sequencing technologies have generated massive databases of ocean genomic data from around the world. So while there is rich data available about Prochlorococcus, mining the value of this Big Data is difficult because it requires simultaneously analyzing various types of complex information—billions of DNA sequences and metadata like location, date, depth, and chemical composition of the water samples—to find meaningful patterns.

"Working with the vast variety of data out there can be a huge challenge for organizations," said Dr. Vijay Gadepally, Lincoln Laboratory Supercomputing Center, and a member of the BigDAWG (Big Data Analytics Working Group) team. "Even basic data manipulation, exploration, and visualization are very difficult due to the volume, velocity, and variety of data being collected. We think BigDAWG will take away the data management problem researchers at the Chisholm Lab currently face."

The BigDAWG system has been a work in progress for the past year and a half, developed by researchers from universities across the nation. Last year, a hackathon held at Beaver Works led to the first prototype of the system. "We developed what we believe is the future of database architecture," said Gadepally. The BigDAWG architecture allows different kinds of data to operate together in a way that lets users get analytics without having to know which specific database or specific tools to use. "The BigDAWG system marks a new era in database technologies by giving users a single interface to all data."

For the past six months, the team has worked with the Chisholm Lab to develop applications within their BigDAWG architecture to fit the specific needs of the lab. In the "Data Explorer" application, scientists can run global queries on the data, which the system can then summarize and produce visualizations (such as graphs) of the search results. With "Genomic Processing," the system can analyze sequence data at scale both quickly and granularly, including finding potentially important outliers. BigDAWG also allows scientists to see relationships between data of four types of microorganisms, enabling the creation of 3D predictive models of microbial communities.

The BigDAWG architecture also makes collecting samples from the ocean more efficient. During expeditions, scientists pull water samples from different locations around the world, where the microbes in each sample are collected, frozen, and transported back to MIT to be sequenced. With the "Navigation" application, the ship's navigation can be adjusted in real time to find optimal seawater examples. "For example, the researchers can select a temperature and salinity level of which they hope to gather Prochlorococcus samples and put it in into the BigDAWG system, and the system will present an idea of where a lot of Prochlorococcus will be present based on the historical data," Gadepally explained.

The BigDAWG team hopes to release the first open source version of the architecture early next year. An important aspect of Gadepally's role on the team has been finding opportunities to apply the system to problems of interest to the larger scientific, industrial, and academic communities. Apart from its applications for the Chisholm Lab, the team also demonstrated the effectiveness of BigDAWG in an entirely different use—an intensive-care unit dataset from Beth Israel Deaconess Medical Center. The application served the needs of doctors by providing real-time support for streams of patient data.

"It's great to be involved with a project that has very real consequences," Gadepally said. As for the Chisholm Lab, researchers there will soon be able to use the BigDAWG applications. "While we have built the data management infrastructure and some initial applications, we want to continue working with the Chisholm Lab folks to now get into the good science. Even a basic understanding of how these organisms work can make a large impact."

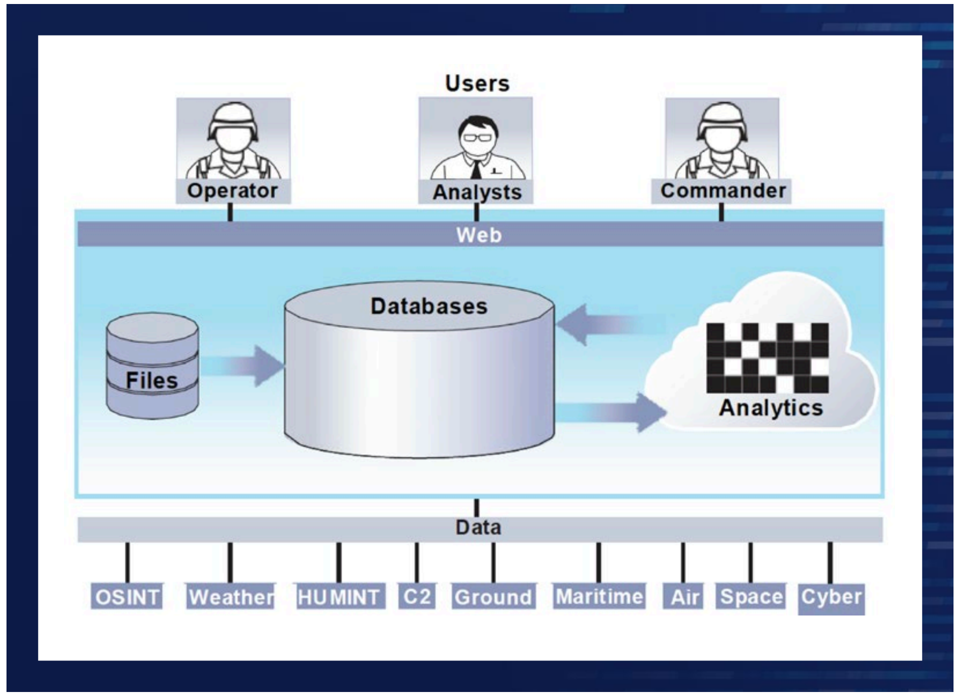
MIT SuperCloud

Line Project

The Technology Office

MIT Lincoln Laboratory

May 8, 2015



The MIT SuperCloud defines a common big data architecture that can be used to address the challenges faced by the volume, velocity, variety, and veracity of big data.

Traditional "dense" data sensors, such as video or radar, and new "sparse" data sensors, such as social media and computer network logs, are producing data faster than Department of Defense (DoD) analysts can exploit it. This "big data" problem is pervasive across the DoD and is most commonly characterized by the three "Vs" of big data: volume, velocity, and variety. Now, increasingly, a fourth "V", veracity (i.e., security), is becoming prominent. The MIT SuperCloud project has demonstrated significant quantitative impact on all of these areas. A key element of the MIT SuperCloud is its database management system, which allows unlimited instances of the National Security Agency (NSA)–developed Apache Accumulo database to be deployed on a project-by-project basis. This capability has moved Lincoln Laboratory to the forefront of the DoD big data community, as Accumulo becomes an increasingly widely used database for the U.S. intelligence community. Furthermore, the Laboratory was the first organization outside the NSA to publish benchmarks of Accumulo and has been an intellectual leader of the Accumulo community ever since. Data variety is perhaps the largest challenge in big data. Traditional SQL-style database schemas are too brittle to handle new and rapidly changing data. However, the SuperCloud's D4M Schema, widely used across the Accumulo community, has largely solved this problem. Dr. Jeremy Kepner, Senior Staff, and Dr. Vijay N. Gadepally, Secure Resilient Systems and Technology, Group 53, conduct this project. Current activities include demonstration of real-time integration of sparse social media data with dense sensor data.

LLSC in the News

On the pages that follow are some of the articles that have been written about the LLSC, the Laboratory programs that utilized LLSC resources, and supercomputing initiatives performed in collaboration with researchers from MIT over the LLSC's first five years.

These articles can also be found in the MIT DSpace Repository: *https://tinyurl.com/LLSC-news*

Publications from LLSC team members are available for viewing online and download from the Google Scholar website:

- To view sorted by year: *https://tinyurl.com/LLSC-pubs-year*
- To view sorted by cites: *https://tinyurl.com/LLSC-pubs-cites*

2020 IEEE Boston Section Distinguished Service Award Recipient

IEEE Boston

The Bulletin

August 23, 2021

Congratulations Dr. Albert Reuther!

The 2020 recipient of the IEEE Boston Section's "Distinguished Service Award" (DSA) is Dr. Albert Reuther. The DSA is awarded to an IEEE Boston Section volunteer who is active in our section, chapter, or conference efforts. Specifically, the DSA honors an IEEE Boston Section member who has made exceptional and distinguished contributions to the Boston IEEE Section. Dr. Reuther received the award for his extraordinary efforts on the section's High Performance Extreme Computing Conference (HPEC) Committee in 2020. His citation reads, "For Continuous and dedicated support of the section's HPEC Conference and extraordinary effort virtualizing HPEC 2020 in response to the global pandemic".

Dr. Albert Reuther is Senior Technical Staff Member in the MIT Lincoln Laboratory Supercomputing Center (LLSC). He brought supercomputing to Lincoln Laboratory through the establishment of LLGrid, founded the LLSC, and leads the LLSC Computational Science and Engineering team. He developed the gridMatlab high-performance computing (HPC) cluster toolbox for pMatlab and is the computer system architect of the MIT Supercloud and numerous interactive supercomputing clusters based on Supercloud, including those in the LLSC. He is the technical chair of the IEEE High Performance Extreme Computing Conference and has organized numerous workshops on interactive HPC, cloud HPC, economics of HPC, and HPC security. His areas of research include interactive HPC; computer architectures for machine learning, graph analytics, and parallel signal processing; and computational engineering. Dr. Reuther earned PhD degree in electrical and computer engineering in 2000 from Purdue University and an MBA degree from the Collège des Ingénieurs in Paris, France, and Stuttgart, Germany in 2001.

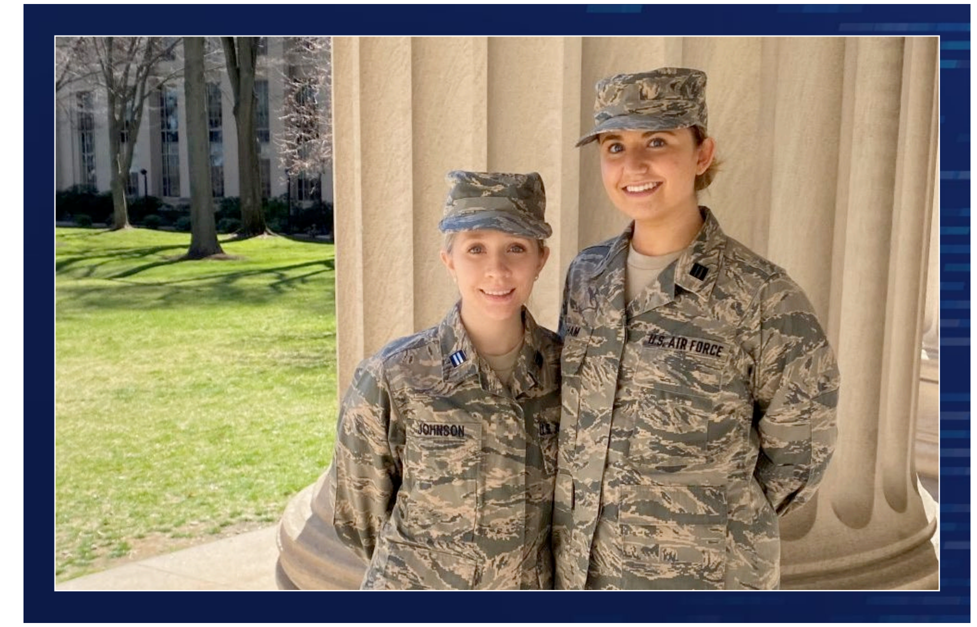
In addition to the DSA, the section has two other awards for which qualified section members can also be nominated. They are, "Student Achievement Award" – to recognize a college student who demonstrates the potential to become a distinguished leader and outstanding contributor in an IEEE field of interest, and the, "Distinguished Member Award" which recognizes outstanding longterm service (10-years or more) to the Boston Section of the IEEE and significant contributions in an IEEE field of interest.

Cadets Collaboratively Intern Through the Air Force-MIT Artificial Intelligence Accelerator

MIT Lincoln Laboratory

The Bulletin

May 14, 2021



Cadets, such as those shown above, provide a bridge between the research done at the Laboratory and MIT, and the U.S. Air Force.

Since the 1920s, when their Air Force Reserve Officer Training Course units were established, the U.S. Air Force has maintained great relationships with colleges and universities.

In recent years, the relationship between the Air Force and MIT has strengthened even more with the creation of the Department of the Air Force-MIT Artificial Intelligence Accelerator (DAF-MIT AIA). The program is a collaboration between the Air and Space Forces and MIT to create new technology that will help the Air Force better complete their mission.

With the demands of this new program came the need for interns, and the first place they looked was Detachment 365—which is based out of MIT, but also serves Harvard University, Tufts University, and Wellesley College.

"I think it was the January or February timeframe of 2020 that Dr. Jeremy Kepner from Lincoln Laboratory came to us and said, 'We have ROTC cadets with these resumes who would like to continue to work on these projects,'" said Capt. Victor Lopez, DAF-MIT AIA project lead for reconnaissance. "I think all of us thought about it and said this is a really great opportunity. We're already at this university."

There have already been several cadets that have interned with the program in various capacities. These cadets may only be undergraduate students, but their work is important to the program as a whole because they bridge a connection between the research at the Laboratory, MIT, and the Air Force.

"At the end of the day, part of our mission at the accelerator is not just to do research and create publications and find ways to bring technology into the gap," said Lopez. "It's also to be liaisons and to be advocates for what government service is going to look like, what the military's role in this new space is going to be and ensuring that it is in line with the democratic principles, ethics, and morals that we want our country to have."

This article first appeared on the USAF-MIT AI Accelerator website and has been adapted for The Bulletin.

Supercomputing Center Celebrates Fifth Anniversary

MIT Lincoln Laboratory

Division 5: Cyber Security and Information Sciences

May 14, 2021

The Lincoln Laboratory Supercomputing Center celebrates its fifth anniversary this year, and continues to act as a critical resource for Laboratory researchers.

The Lincoln Laboratory Supercomputing Center (LLSC) celebrated a significant anniversary this year, marking five years of the center's mission to enhance the computing power available to the Laboratory, MIT, and other researchers.

The LLSC was established on 1 April 2016 to build on the Laboratory's past groundbreaking work in computing. In the late 1940s and early 1950s, prior to the establishment of the Laboratory, MIT designed and built the Whirlwind 1 computer using vacuum tubes for the U.S. Navy. Whirlwind was unique for its ability to output information in real time and perform calculations in parallel. Later, in 1955, the first fully transistor-based computer TX-0 was designed by staff at the now-established Lincoln Laboratory.

These and other early computing accomplishments eventually led to the establishment in 2003 of LLGrid, a Laboratory supercomputing system. In 2008, the Laboratory demonstrated the largest problem ever run on a computer using the TX-2500 supercomputer, a computer that was a part of LLGrid.

The LLSC was established with a focus on interactive supercomputing and high-performance data analysis. The staff who started the LLSC decided to model the center on the partnership model found at federal laboratories and universities around the world.

"In any given year, over half of the Laboratory staff are conducting supercomputing-enabled research. Our hallmark is being the world leader in interactive supercomputing," said Dr. Jeremy Kepner, Laboratory Fellow, LLSC. "When we started, all processors had a single core, accelerators did not exist, and most of the work was signal processing in MATLAB. Today, all of our processors have at least 20 cores, we have hundreds of accelerators, and the work spans many domains using a wide range of programming environments."

The resources that LLSC provides have been critical to many award-winning technologies at the Laboratory. For example, a Laboratory-developed program called Large-scale Automated Vulnerability Addition (LAVA), which helps discover bugs in code, utilized the LLSC for a competition they hosted. The competition, called Rode0day, pitted programs based on LAVA against each other. LAVA won a 2020 R&D 100 Award.

"Using the Laboratory's supercomputing resources, we were able to run an experiment to test eight of the leading bug-finding systems," said Andrew Fasano, Cyber System Assessments, Group 59, who worked on LAVA. "Using a standard one-CPU system, it would have taken more than 83 years to collect all of the data."

Many of the staff who were there for the start of the LLSC are still at the Laboratory today.

"We continue to add new people to the group, but being able to take on this journey with this team has been particularly special," said Dr. Vijay Gadepally, Senior Staff, LLSC.

The LLSC recently deployed TX-GAIA in 2019, which is currently rated by Top500 as the 57th fastest supercomputer in the world. It has served a critical role in AI research, including research conducted by the MIT-Air Force AI Accelerator and research related to COVID-19.

In the future, the LLSC hopes to support even more research and collaborate with more researchers across the Laboratory and MIT.

"Collaboration with regional partners will be a key part of taking our next big step, and will allow us to do bigger science, engineering, and supercomputing together," said Kepner.

"Supercomputing plays a huge part in research going on across the Laboratory, and LLSC systems have provided the computational power for a number of Laboratory achievements," said Gadepally. "We've been working to ramp up the research output from the LLSC to become one of the nation's preeminent supercomputing research organizations, in addition to the cutting-edge infrastructure and systems we already provide."

The original LLSC team, when it was established in 2016, is shown above.

Staff Member Selected as 2021 SIAM Fellow

MIT Lincoln Laboratory

The Bulletin

April 23, 2021

Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center, was selected to join SIAM's class of 2021 Fellows.

The Society for Industrial and Applied Mathematics has selected Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center, as an esteemed member of its 2021 class of SIAM Fellows. These distinguished members were nominated for their exemplary research as well as outstanding service to the community. Through their contributions, SIAM Fellows have helped to advance the fields of applied mathematics and computational science.

The award recognizes Kepner for his "contributions to interactive parallel computing, matrix-based graph algorithms, green supercomputing, and big data."

"These individual honors are a recognition of the achievements of our entire Laboratory team, to whom I am eternally indebted," said Kepner.

Kepner has been a member of SIAM since he was a graduate student. He has published two of his books through SIAM, and collaborates with SIAM for both an annual conference and an MIT SIAM student section that he advices.

The goals of the SIAM Fellowship program are to honor SIAM members who are recognized by their peers as distinguished for their contributions to the discipline, help make outstanding SIAM members more competitive for awards and honors when they are being compared with colleagues from other disciplines, and support the advancement of SIAM members to leadership positions in their own institutions and in the broader society.

The Society for Industrial and Applied Mathematics (SIAM), headquartered in Philadelphia, Pennsylvania, is an international society of more than 14,500 individual, academic, and corporate members from more than 100 countries. SIAM helps build cooperation between mathematics and the worlds of science and technology to solve real-world problems through publications, conferences, and communities like chapters, sections and activity groups. Learn more at siam.org.

Kepner was nominated as a candidate by MIT professor Alan Edelman, and is the ninth SIAM Fellow from MIT and the first from the Laboratory.

Brainstorming energy-saving hacks on Satori, MIT's new supercomputer

Kim Martineau

MIT News Office

February 21, 2020

Several dozen students participated in the Green AI Hackathon, co-sponsored by the MIT Research Computing Project and MIT-IBM Watson AI Lab.

Mohammad Haft-Javaherian planned to spend an hour at the Green AI Hackathon — just long enough to get acquainted with MIT's new supercomputer, Satori. Three days later, he walked away with $1,000 for his winning strategy to shrink the carbon footprint of artificial intelligence models trained to detect heart disease.

"I never thought about the kilowatt-hours I was using," he says. "But this hackathon gave me a chance to look at my carbon footprint and find ways to trade a small amount of model accuracy for big energy savings."

Haft-Javaherian was among six teams to earn prizes at a hackathon co-sponsored by the MIT Research Computing Project and MIT-IBM Watson AI Lab Jan. 28-30. The event was meant to familiarize students with Satori, the computing cluster IBM donated to MIT last year, and to inspire new techniques for building energy-efficient AI models that put less planet-warming carbon dioxide into the air.

The event was also a celebration of Satori's green-computing credentials. With an architecture designed to minimize the transfer of data, among other energy-saving features, Satori recently earned fourth place on the Green500 list of supercomputers. Its location gives it additional credibility: It sits on a remediated brownfield site in Holyoke, Massachusetts, now the Massachusetts Green High Performance Computing Center, which runs largely on low-carbon hydro, wind and nuclear power.

A postdoc at MIT and Harvard Medical School, Haft-Javaherian came to the hackathon to learn more about Satori. He stayed for the challenge of trying to cut the energy intensity of his own work, focused on developing AI methods to screen the coronary arteries for disease. A new imaging method, optical coherence tomography, has given cardiologists a new tool for visualizing defects in the artery walls that can slow the flow of oxygenated blood to the heart. But even the experts can miss subtle patterns that computers excel at detecting.

At the hackathon, Haft-Javaherian ran a test on his model and saw that he could cut its energy use eight-fold by reducing the time Satori's graphics processors sat idle. He also experimented with adjusting the model's number of layers and features, trading varying degrees of accuracy for lower energy use.

A second team, Alex Andonian and Camilo Fosco, also won $1,000 by showing they could train a classification model nearly 10 times faster by optimizing their code and losing a small bit of accuracy. Graduate students in the Department of Electrical Engineering and Computer Science (EECS), Andonian and Fosco are currently training a classifier to tell legitimate videos from AI-manipulated fakes, to compete in Facebook's Deepfake Detection Challenge. Facebook launched the contest last fall to crowdsource ideas for stopping the spread of misinformation on its platform ahead of the 2020 presidential election.

If a technical solution to deepfakes is found, it will need to run on millions of machines at once, says Andonian. That makes energy efficiency key. "Every optimization we can find to train and run more efficient models will make a huge difference," he says.

To speed up the training process, they tried streamlining their code and lowering the resolution of their 100,000-video training set by eliminating some frames. They didn't expect a solution in three days, but Satori's size worked in their favor. "We were able to run 10 to 20 experiments at a time, which let us iterate on potential ideas and get results quickly," says Andonian.

As AI continues to improve at tasks like reading medical scans and interpreting video, models have grown bigger and more calculation-intensive, and thus, energy intensive. By one estimate, training a large language-processing model produces nearly as much carbon dioxide as the cradle-to-grave emissions from five American cars. The footprint of the typical model is modest by comparison, but as AI applications proliferate its environmental impact is growing.

One way to green AI, and tame the exponential growth in demand for training AI, is to build smaller models. That's the approach that a third hackathon competitor, EECS graduate student Jonathan Frankle, took. Frankle is looking for signals early in the training process that point to subnetworks within the larger, fully-trained network that can do the same job. The idea builds on his award-winning Lottery Ticket Hypothesis paper from last year that found a neural network could perform with 90 percent fewer connections if the right subnetwork was found early in training.

The hackathon competitors were judged by John Cohn, chief scientist at the MIT-IBM Watson AI Lab, Christopher Hill, director of MIT's Research Computing Project, and Lauren Milechin, a research software engineer at MIT.

The judges recognized four other teams: Department of Earth, Atmospheric and Planetary Sciences (EAPS) graduate students Ali Ramadhan, Suyash Bire, and James Schloss, for adapting the programming language Julia for Satori; MIT Lincoln Laboratory postdoc Andrew Kirby, for adapting code he wrote as a graduate student to Satori using a library designed for easy programming of computing architectures; and Department of Brain and Cognitive Sciences graduate students Jenelle Feather and Kelsey Allen, for applying a technique that drastically simplifies models by cutting their number of parameters.

IBM developers were on hand to answer questions and gather feedback. "We pushed the system — in a good way," says Cohn. "In the end, we improved the machine, the documentation, and the tools around it."

Going forward, Satori will be joined in Holyoke by TX-Gaia, Lincoln Laboratory's new supercomputer. Together, they will provide feedback on the energy use of their workloads. "We want to raise awareness and encourage users to find innovative ways to green-up all of their computing," says Hill.

The Laboratory's New AI Supercomputer is the Most Powerful at any University

MIT Lincoln Laboratory

Division 5: Cyber Security and Information Sciences

July 19, 2019

TX-GAIA is housed inside of a new EcoPOD deployable data center, manufactured by Hewlett Packard Enterprises, at the site of the Lincoln Laboratory Supercomputing Center in Holyoke, Massachusetts.

The new TX-GAIA computing system at the Lincoln Laboratory Supercomputing Center (LLSC) has been ranked as the most powerful artificial intelligence (AI) supercomputer at any university in the world. The ranking comes from TOP500, which publishes a list of the top supercomputers in various categories biannually. The system, which was built by Hewlett Packard Enterprises, combines traditional high-performance computing hardware (nearly 900 Intel processors) with hardware optimized for AI applications (900 Nvidia GPU accelerators).

"We are thrilled by the opportunity to enable researchers across the Laboratory and MIT to achieve incredible scientific and engineering breakthroughs," said Dr. Jeremy Kepner, Laboratory Fellow, LLSC. "TX-GAIA will play a large role in supporting AI, physical simulation, and data analysis across all Laboratory missions."

TOP500 rankings are based on the LINPACK Benchmark, which is a measure of a system's floating-point computing power, or how fast a computer solves a dense system of linear equations. TX-GAIA's TOP500 benchmark performance of 3.9 quadrillion floating-point operations per second, or petaflops, places it #1 in the Northeast, #20 in the United States, and #51 in the world for supercomputing power. The system's peak performance is more than 6 petaflops.

But more notably, TX-GAIA has a peak performance of 100 AI petaflops, which makes it #1 for AI flops at any university in the world. An AI flop is a measure of how fast a computer can perform deep neural network (DNN) operations. DNNs are a class of AI algorithms that learn to recognize patterns in huge amounts of data. This ability has given rise to "AI miracles," as Kepner put it, in speech recognition and computer vision; the technology is what allows Amazon's Alexa to understand questions and self-driving cars to recognize objects in their surroundings. As these DNNs grow more complex, it takes longer for them to process the massive datasets they learn from. TX-GAIA's Nvidia GPU accelerators are specially designed for performing these DNN operations quickly.

TX-GAIA is housed in a new modular data center, called an EcoPOD, at the site of the LLSC in Holyoke, Massachusetts. It joins the ranks of other powerful systems at the LLSC, such as the TX-E1, which supports collaborations with MIT campus and other institutions, and TX-Green, which is currently ranked #492 on the TOP500 list and is a zero-carbon system.

Kepner said that the system's integration into the LLSC will be completely transparent to users when it comes online this fall: "The only thing users should see is that many of their computations will be dramatically faster."

Among its AI applications, TX-GAIA will be tapped for training machine-learning algorithms, including those that use DNNs. It will more quickly crunch through terabytes of data—for example, hundreds of thousands of images or years' worth of speech samples to teach these algorithms to figure out solutions on their own. The system's compute power will also expedite simulations and data analysis. These capabilities will support projects across the Laboratory's R&D areas, such as improving weather forecasting, accelerating medical data analysis, building autonomous systems, designing synthetic DNA, and developing new materials and devices.

TX-GAIA, which is also ranked the #1 system in the Department of Defense, will also support the recently announced MIT-Air Force AI Accelerator. The partnership will combine the expertise and resources of MIT, including those at the LLSC and the Air Force, to conduct fundamental research directed at enabling rapid prototyping, scaling, and application of AI algorithms and systems.

MIT and U.S. Air Force Sign Agreement to Launch AI Accelerator

MIT Lincoln Laboratory

The Bulletin

July 12, 2019

MIT and the U.S. Air Force have signed an agreement to launch a new program designed to make fundamental advances in artificial intelligence that could improve Air Force operations while also addressing broader societal needs.

The effort, known as the MIT-Air Force AI Accelerator, will leverage the expertise and resources of MIT and the Air Force to conduct fundamental research directed at enabling rapid prototyping, scaling, and application of AI algorithms and systems. The Air Force plans to invest approximately $15 million per year as it builds upon its five-decade relationship with MIT.

The Beaver Works facility will house the program and the Laboratory will make available its specialized facilities and resources. Bob Bond, Chief Technology Officer, Technology Office, and Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center, will oversee the Laboratory's involvement.

"Laboratory researchers will be key contributors to all aspects of the AI Accelerator. This includes performing groundbreaking AI research, developing new AI challenges to engage the entire AI ecosystem, and transitioning AI innovations into Air Force missions," said Kepner.

The program is expected to support at least 10 MIT research projects addressing challenges that are important to both the Air Force and society more broadly, such as disaster response and medical readiness.

"This collaboration is very much in line with MIT's core value of service to the nation," said Maria Zuber, MIT's vice president for research and the E.A. Griswold Professor of Geophysics. "MIT researchers who choose to participate will bring state-of-the-art expertise in AI to advance Air Force mission areas and help train Air Force personnel in applications of AI."

Under the agreement, MIT will form interdisciplinary teams of researchers, faculty, and students whose work focuses on topics in artificial intelligence, control theory, formal methods, machine learning, robotics, and perception, among other fields. Teams will also include leaders in technology policy, history, and ethics from a range of departments, labs, and centers across the Institute. Members of the Air Force will join and lend expertise to each team.

"MIT is the leading institution for AI research, education, and application, making this a huge opportunity for the Air Force as we deepen and expand our scientific and technical enterprise," said Heather Wilson, former Secretary of the Air Force. "Drawing from one of the best of American research universities is vital."

The program will aim to develop new algorithms and systems to assist complex decision-making that might help the Air Force, for example, better focus its maintenance efforts—an expensive and critical part of its aircraft operations. This fundamental research also intends to develop AI to assist humans in aspects of planning, control, and other complex tasks. Finally, the work aims to enable rapid deployment of advanced algorithms and capabilities developed at MIT to foster AI innovation across the country.

In addition to disaster relief and medical readiness, other possible research areas may include data management, maintenance and logistics, vehicle safety, and cyber resiliency.

This article first appeared on the MIT News Office website, originally authored by Rob Matheson, and had been adapted for The Bulletin.

The new program will focus on rapid deployment of artificial intelligence innovations in operations, disaster response, and medical readiness. (Photo credit: Jake Belcher)

Wilkinson Prize Goes to Developers of Flexible Julia Programming Language

MIT Lincoln Laboratory

The Bulletin

March 29, 2019

Two recipients of the Wilkinson Prize—MIT researchers Viral Shah (left) and Jeff Bezanson (center)—are seen with founding sponsor of Julia, Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center. The third recipient, Stefan Karpinski, was unavailable for the photo. Credit: Jeremy Kepner.

Programmers developing applications for fields as diverse as astronomy, economics, artificial intelligence, energy optimization, and medicine often found themselves creating software with languages that offered slow computation. But in this era of big data, dynamic, flexible, and easy-to-implement code is required for programmers to efficiently build high-performance software tools needed for intensive data analysis. Enter Julia, an open-source language for advanced technical computing and data science.

Research on Julia began in MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) and its Department of Mathematics under the guidance of Alan Edelman, a professor of mathematics, and founding sponsor Dr. Jeremy Kepner, Laboratory Fellow, Supercomputing Center. Three researchers at MIT who were creating the initial version of Julia in 2009 were named the recipients of the 2019 James H. Wilkinson Prize for Numerical Software.

This highly competitive prize is awarded once every four years to recognize authors of "innovative software in scientific computing" and to encourage emerging research talent in computer science. The selection committee for the prize said that the Julia codevelopers have allowed researchers "to write high-level code in an intuitive syntax and produce code with the speed of production programming languages."

Julia was launched publicly in 2012 as an open-source product. Since then, Julia has been downloaded upwards of three million times and is used in more than 1,500 universities for scientific and numerical computing. The project lives online at https://julialang.org.

Labortory staff were some of the early adopters of the Julia language. Developers of an advanced collision avoidance system (ACAS X) for both manned and unmanned aircraft used Julia in creating the machine learning algorithms that determine the decisions and resulting warnings the system issues to pilots and air traffic managers. Currently, researchers across the Laboratory are using Julia in the refinement of ACAS X capabilities and in the development of various other systems, such as one to compare DNA signatures.

New Textbook Applies Mathematics to the Management of Big Data

MIT Lincoln Laboratory

The Bulletin

August 10, 2018

The rapid explosion of digitally available data in banking, commerce, science and engineering, transportation, healthcare, and social sciences has amplified the problems of storing and analyzing data. To organize huge datasets into constructs that they can manipulate with some ease, people have developed spreadsheets, databases, matrices, and graphs. Yet, the datasets are still often unmanageable, and analysis of the data consumes much time and computation. The new book Mathematics of Big Data: Spreadsheets, Databases, Matrices, and Graphs presents a mathematical approach to taming big data: using associative arrays to streamline data processing.

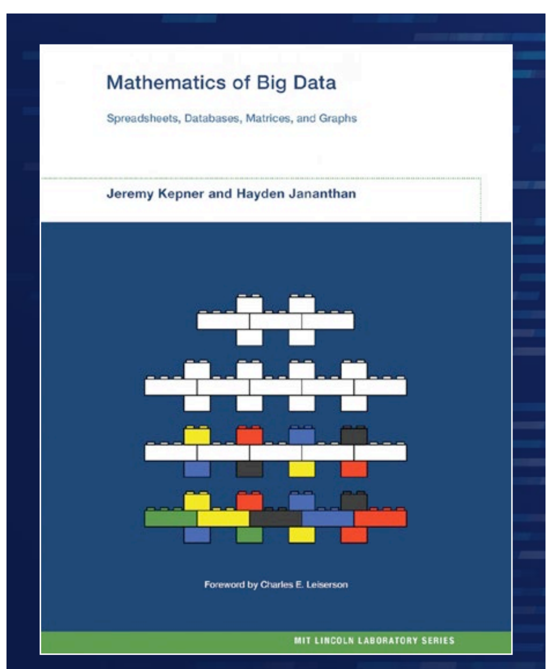
"Data provides insight into solutions for practical problems. The goal of this book is to provide readers with the concepts and techniques that will allow them to adapt to increasing data volume, velocity, and variety," said Kepner. "Nothing handles big like mathematics."

The book is divided into three sections. The first section, Applications and Practice, introduces the concept of the associative array and the application of associative arrays to graph analysis and machine learning systems. In the second part, Mathematical Foundations, the authors provide a mathematically rigorous definition of associative arrays. The third section, Linear Systems, discusses the extension of the mathematical concepts of linearity to associative arrays. Each chapter ends with a series of exercises that allow readers to test their understanding of the material.

In the foreword to the book, Charles Leiserson, professor of computer science and engineering at MIT and head of the Supertech Research Group in MIT's Computer Science and Artificial Intelligence Lab, said, "Jeremy Kepner and Hayden Jananthan will lead you on a voyage to relearn everything you know about matrices, graphs, databases, and spreadsheets, viewing their tight interrelationships through the lens of associative arrays. Or if you're new to these topics, they will lead you on a journey of discovery, teaching them to you without the arbitrary constraints of traditional educational narratives."

Kepner had previously offered a Technical Education course on the same topic.

Mathematics of Big Data is available for purchase in print or e-book at the MIT Press website.



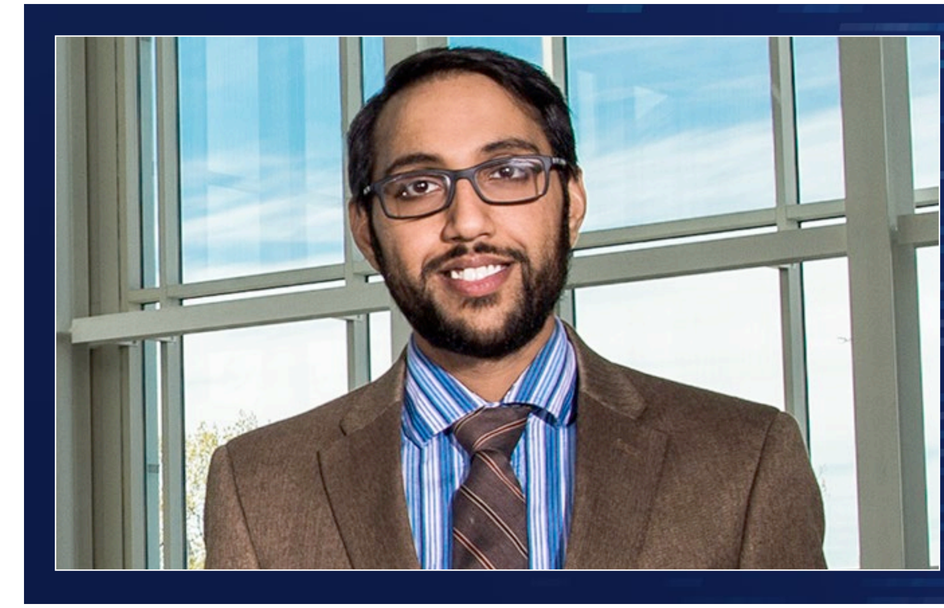
Mathematics of Big Data is the first book to present the common mathematical foundations of big data analysis across a range of applications and technologies.

Dr. Vijay Gadepally Receives AFCEA International's 40 Under 40 Award

MIT Lincoln Laboratory

The Bulletin

June 17, 2017



Dr. Vijay Gadepally has been instrumental in the establishment of the Lincoln Laboratory Supercomputing Center.

Dr. Vijay Gadepally was selected as a recipient of the Armed Forces Communications and Electronics Association (AFCEA) International's 40 Under 40 Award. Gadepally has been instrumental in the establishment and continued growth of the Lincoln Laboratory Supercomputing Center, which hosts New England's most powerful supercomputer.

"Vijay has demonstrated outstanding technical leadership and creativity in advancing high-performance computing at the Laboratory and throughout the academic computing community. He has made foundational contributions to database management and algorithm development for graph processing and machine learning," said David Martinez, Associate Division Head, Cyber Security and Information Sciences, Division 5.

Gadepally joined the Laboratory in 2013 after earning a doctorate in electrical and computer engineering at the Ohio State University, and was honored in 2016 with a Lincoln Laboratory Early Career Achievement Award.

MIT Lincoln Laboratory Takes the Mystery Out of Supercomputing

Dell

insideHPC

January 18, 2017

Lincoln Laboratory Supercomputer Center

The introduction of multicore and manycore processors capable of handling highly parallel workflows is changing the face of high performance computing (HPC)?

Many supercomputer users, like the big DOE labs, are implementing these next generation systems. They are now engaged in significant code modernization efforts to adapt their key present and future applications to the new processing paradigm, and to bring their internal and external users up to speed. For some in the HPC community, this creates unanticipated challenges along with great opportunities.

Here at MIT Lincoln Laboratory with its recently opened Supercomputer Center, the picture is a little different. We have been preparing for these specific changes over the past 15 years and conducting what we call “interactive supercomputing” since the 1950s. It has been clear from the start that data analysis is the major focus for most of our users; and that certain combinations of mathematics and processors will be beneficial to them.

We help our users to think about their problems with the right mathematics using a high level parallel programming environment with languages like Mattlab, Python, Java, R and Julia. These sophisticated production environments allow our users to become comfortable working with highly parallel workflows. They can try out different solutions using more complex algorithms that take full advantage of the manycore processors in our new 1+ petaflop system composed of the latest Intel Xeon Phi processors connected via the Intel Omni-Path network.

In 2008, Lincoln demonstrated the largest single problem ever run on a computer using its TX-2500 supercomputer, part of a system called the LLGrid. This in turn led to the establishment of the Lincoln Laboratory Supercomputer Center (LLSC) in April 2016.

The Center was developed to enhance computing power and accessibility to over 1000 researchers across the institute. Our engineers, scientists, faculty and students use the LLSC capabilities to conduct research in such diverse fields as space observation, robotic vehicles, communications, cyber security, sensor processing, electronic devices, bioinformatics and air traffic control to name just a few. Three areas of extreme interest that demand more computational capabilities include autonomous systems, device physics, and machine learning.

For example, if you have a large number of mobile, very smart systems navigating a limited space, you will want to simulate all the ways those systems can interact. They will need to learn how they can avoid negative situations – like bumping into each other – and realize positive results, such as figuring out what their mission is as individuals or part of a group, and carry out the steps necessary to realize that mission. All this requires extraordinary computational capabilities, which are now becoming available at LLSC.

In the realm of device physics, Lincoln has a long tradition of developing advanced device technology for both sensing and computation. As we design and prototype these devices, the use of leading–edge engineering practices have become the de facto standard. This includes extensive, computationally intensive simulation before a commitment is made to build and deploy.

In fact, we have an enormous amount of rapid prototyping underway – Lincoln Laboratory has been called the best rapid prototyping lab in the world[1]. Activities include mechanical system design and fabrication, electronics system design and fabrication, system engineering, integration into aircraft and other platforms, and laboratory and fieldtesting.

Our approach to interactive supercomputing makes fast turnaround possible not only for rapid prototyping, but for all projects designed for parallel processing on the new HPC platforms from Intel and others.

For the users, this means there are no queues to contend with – a complex scheduler runs transparently in the background to ensure that users with tight deadlines meet their goals without becoming bogged down in the complexities usually associated with HPC.

We also provide interfaces that make it easy for users to continue to use exactly the same desktop environment they have always used. Because they do not have to take the time to learn new ways to work, they can tackle bigger problems on a shorter time scale.

Guided by the principles of interactive supercomputing, Lincoln Laboratory was responsible for a lot of the early work on machine learning and neural networks. We now have a world-class group investigating speech and video processing as well as machine language topics including theoretical foundations, algorithms and applications.

In the process, we are changing the way we go about computing. Over the years we have tended to assign a specific systems to service a discrete market, audience or project. But today those once highly specialized systems are becoming increasingly heterogeneous. Users are interacting with computational resources that exhibit a high degree of autonomy. The system, not the user, decides on the computer hardware and software that will be used for the job.

This is not a new concept at Lincoln – using general purpose processors and accelerators we have already built our own systems for specific applications such as graph processing. However, with the Xeon Phi processor we are seeing the next step in an evolution that MIT has been part of for more than a decade. The basic idea of a large “mesh” architecture with many processors in a grid, each processor with a memory unit and a 3D-stack on top of it, was pioneered by MIT twenty years ago. So, for the past 15 years we have been preparing ourselves and our user base to use tools and techniques that we knew would work well when these highly parallel architectures became available as mainstream commodity processors.

We have been able to do this because the laws of physics give us a good idea of what those processors will look like in the future – 10 to 15 years out. So it is expected that the APIs and applications that we have been developing for the past decade will run well on these new systems.

This approach takes the mystery out of developing HPC systems for next generation applications. If you know how supercomputers will evolve over the next decade, then you will most likely know what kind of mathematics those systems will be running.

For example, we have worked with our user base over the years to incorporate matrix-oriented mathematics into their applications. This approach ensures that our user applications will work smoothly with the processors that are being produced now and in the near future. We anticipate that more than 70% of our user base will be able to use these new systems out of the box.

One of the milestones in the history of supercomputing at MIT was the establishment of the Lincoln Laboratory Supercomputer Center (LLSC) in April 2016 at a 9 acre location in Holyoke, Mass.

The LLSC was developed to enhance computing power and accessibility for over 1,000 researchers across the Institute. Its mission is to address supercomputing needs, develop new supercomputing capabilities and technologies, and to collaborate with MIT campus supercomputing initiatives.

The center is unique to MIT in that it is focused on interactive supercomputing for high performance data analysis. It allows our scientific and engineering users to tap into the vast resources that LLSC makes available – very large computing clusters and storage systems and advanced programming interfaces.

LLSC has selected Dell EMC to install a 648-node HPC system through the Dell EMC and Intel early access program for the Intel Xeon Phi processor. LLSC's new “TX-Green” system, one of the largest of its kind on the US East Coast, exceeds one petaflop and has provided the center with a 6X computing capacity boost

LLSC supports numerous programming languages and software libraries, including Python, C, C++, Fortran, Java, Julia, MPI, OpenMP, and parallel MATLAB®. LLSC collaborates closely with the MIT Beaver Works Center for Engaging Supercomputing.

Dell EMC has been a great partner in enabling us to dramatically increase the capabilities of our supercomputing center. The Dell HPC team was very knowledgeable and responsive and able to deliver, install, and benchmark our Petaflop-scale system in less than a month. This was a great example of a well-coordinated and dedicated organization that was able to allocate the appropriate resources to exceed customer expectations.

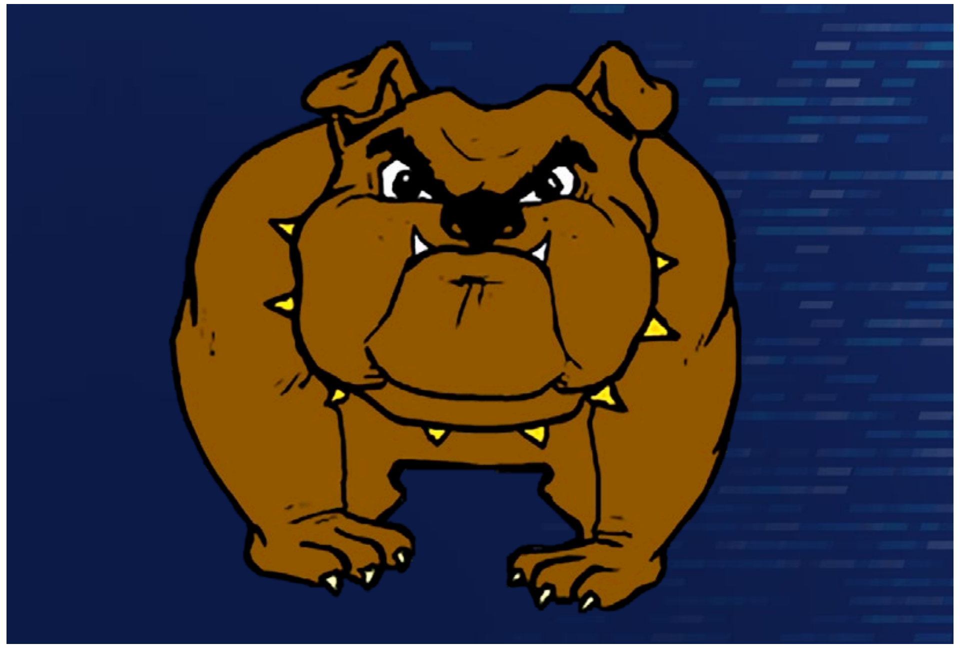
The end result is that not only do we have an extremely powerful computer resource supporting our research efforts, we also have created an “extremely green” computing center – our computers run 93% carbon free, something we are really proud of.

[1] MIT President shares vision for the future of innovation, Sep 30, 2016.

Supercharging Big Data Research with New England's Fastest Supercomputer

istc-bigdata.org

January 11, 2017



https://bigdawg.mit.edu

MIT recently announced that the MIT Lincoln Laboratory Supercomputing Center (LLSC) system, housed in Holyoke, Massachusetts, has been ranked the most powerful supercomputer in New England. We caught up with ISTC for Big Data Principal Investigator and Lincoln Laboratory fellow Dr. Jeremy Kepner, who heads the LLSC, to learn more about the supercomputer, how it's helping ISTC research, and his work for the ISTC.

What role have you and MIT Lincoln Laboratory been playing in the ISTC for Big Data?

My team has been providing the big data computing resources, data sets, and demo integration for the BigDAWG polystore system, the ISTC's capstone project to radically simplify big data management. Our team works on a lot of big data projects and has many decades of experience integrating large and complex data sets to address challenging scientific problems.

How will the new LLSC supercomputer help with the ISTC work?

With its new Dell EMC petaflop-scale supercomputer, the LLSC has 6 times more processing power and 20 times more bandwidth than with its predecessor. This is good news to the more than 1,000 researchers—including ISTC for Big Data researchers—who depend on it in their work.

Researchers use our interactive supercomputing resources to augment their desktop systems—to process large sets of sensor data, create high-fidelity simulations, develop new algorithms and do other compute-intensive work. With the new system, the ISTC can now scale up our BigDAWG architecture to run on much larger platforms.

Can you elaborate a bit on how your own research interests—specifically D4M with its associative arrays—are contributing to the BigDAWG polystore system?

My primary interests are in highperformance computing, parallel algorithms, and computational software. The BigDAWG polystore system is next-generation federation middleware that supports many different data models and databases. It uses a concept called islands of information to unite many, diverse query processing engines, so users can make complex, cross-database queries using their current tools (for example, SQL) and get answers fast and simply.

Associative arrays provide a mathematical model that may encompass many of the diverse databases in BigDAWG. The ability to describe database representations and queries within a single mathematical model is a strong indication that BigDAWG is on the right track.

What's the latest on your work on BigDAWG?

DM4 is one of two cross-system islands implemented in BigDAWG. The other is Myria, from the University of Washington. Each offers a different interface to an overlapping set of backend database engines.

Myria has adopted a programming model of relational algebra extended with iteration. Among other engines, it includes shims (cross-database translators) to SciDB and Postgres. Myria includes a sophisticated optimizer to efficiently process its query language.

On the other hand, D4M uses a new data model, associative arrays, as an access mechanism for existing data stores. This data model unifies multiple storage abstractions, including spreadsheets, matrices, and graphs. D4M has a query language that includes filtering, subsetting, and linear algebra operations, and it contains shims to Accumulo, SciDB and Postgres.

BigDAWG currently consists of a "scope-cast facility" with island implementations from Myria and D4M, along with degenerate islands for three production databases (Accumulo, SciDB, and Postgres). To provide the union of the capabilities of the federator's underlying storage engines, these degenerate islands have the full functionality of a single storage engine— so users don't lose the capabilities of their databases. We also incorporate access to other experimental database systems.
You can read more about the BigDAWG architecture in our paper, "A Demonstration of the BigDAWG Polystore System."

You've got a new book coming out soon from MIT Press: "The Mathematics of Big Data." What's it about?

Broadly, it's about the role of mathematics in taking full advantage of big data. Specifically, it provides a unifying mathematical framework for representing data through all the steps of a machine learning system—one that can reduce the amount of time and effort currently spent on front-end processing for machine learning (typically 90%).

Big data machine learning systems encompass the entire process of parsing, ingesting, querying, and analyzing data to make predictions—the biggest promise of big data. The variety of front-end processing approaches for enabling machine learning systems include data representation, graph construction, graph traversal, and graph centrality metrics. In many cases, well-designed front-end processing can significantly reduce the complexity of the back-end machine learning algorithm and allow a simpler algorithm to be used. "The Mathematics of Big Data" describes the mathematical basis for implementing these approaches.

Laboratory's Supercomputing System Ranked Most Powerful in New England

MIT Lincoln Laboratory

Division 5: Cyber Security and Information Sciences

December 2, 2016

LLSC Systems Engineer Matthew Hubbell inspects the new supercomputer.

The new computing system, TX-Green, at the Lincoln Laboratory Supercomputing Center (LLSC) has been named the most powerful supercomputer in New England and the 3rd most powerful at a United States university. A team of experts at TOP500.org ranks the world's 500 most powerful supercomputers biannually. The systems are ranked based on a LINPACK Benchmark, which is a measure of a system's floating-point computing power, or how fast a computer solves a dense system of linear equations.

Established in early 2016, the LLSC was developed to enhance computing power and accessibility for more than 1,000 researchers across the Laboratory. The LLSC uses interactive supercomputing to augment the processing power of desktop systems to process larger sets of sensor data, create higher-fidelity simulations, and develop new algorithms. Located in Holyoke, Massachusetts, the new system is the only zero-carbon supercomputer on the TOP500 list. It uses energy from a mixture of hydroelectric, wind, solar, and nuclear sources.

In November, Dell EMC installed a new petaflop-scale system, which consists of 41,472 Intel processor cores and is capable of computing 1015 operations per second. Compared to LLSC's previous technology, the new system provides six times more processing power and 20 times more bandwidth. This technology enables new research in several Laboratory research areas, such as space observation, robotic vehicles, communications, cyber security, machine learning, sensor processing, electronic devices, bioinformatics, and air traffic control.

The LLSC mission is to address supercomputing needs, develop new supercomputing capabilities and technologies, and collaborate with MIT campus supercomputing initiatives. "The LLSC vision is to enable the brilliant scientists and engineers at Lincoln Laboratory to analyze and process enormous amounts of information with complex algorithms," said Dr. Jeremy Kepner, Laboratory Fellow and head of the LLSC. "Our new system is one of the largest on the East Coast and is specifically focused on enabling new research in machine learning, advanced physical devices, and autonomous systems.

Because the new processors are similar to the prototypes developed at the Laboratory more than two decades ago, the new petaflop system is compatible with all existing LLSC software. "We have had many years to prepare our computing system for this kind of processor," said Kepner. "This new system is essentially a plug-and-play solution."

After establishing the Supercomputing Center as one of the top systems in the world, the LLSC team will continue to upgrade and expand supercomputing at the Laboratory. Said Kepner, "Our hope is that this system is the first of many such large-scale systems at Lincoln Laboratory."

The center is located in Holyoke, Massachusetts, and is home to the only zero-carbon emissions supercomputer on the TOP500 list.

Establishment of the Lincoln Laboratory Supercomputing Center

MIT Lincoln Laboratory

Division 5: Cyber Security and Information Sciences

April 15, 2016

The Lincoln Laboratory Supercomputing Center team.

Lincoln Laboratory has been a world leader in interactive supercomputing since the 1950s. In 1955, TX-0, the first fully transistor-based computer, was built to support a wide range of research at the Laboratory and the MIT campus, and became the basis for the second largest computing company in the world, Digital Equipment Corporation. In 2001, the Laboratory developed Parallel Matlab, which enabled thousands of researchers worldwide to use interactive supercomputing for high performance data analysis. In 2008, the Laboratory demonstrated the largest single problem ever run on a computer, using its TX-2500 supercomputer, a part of the system called LLGrid. Recently, the Laboratory acknowledged the importance of the LLGrid world-class computing capability with the establishment of the Lincoln Laboratory Supercomputing Center (LLSC) on 1 April.

LLSC is based in part on the LLGrid infrastructure, but has been developed to enhance computing power and accessibility to over 1,000 researchers across the Institute. "By establishing the LLSC, Lincoln Laboratory will be able to better address supercomputing needs across all Laboratory missions, develop new supercomputing capabilities and technologies, and spawn even closer collaborations with MIT campus supercomputing initiatives," said Dr. Jeremy Kepner, Laboratory Fellow, and head of the Supercomputing Center. "These brilliant engineers, scientists, faculty, and students use our capabilities to conduct research in diverse fields such as space observations, robotic vehicles, communications, cyber security, machine learning, sensor processing, electronic devices, bioinformatics, and air traffic control."

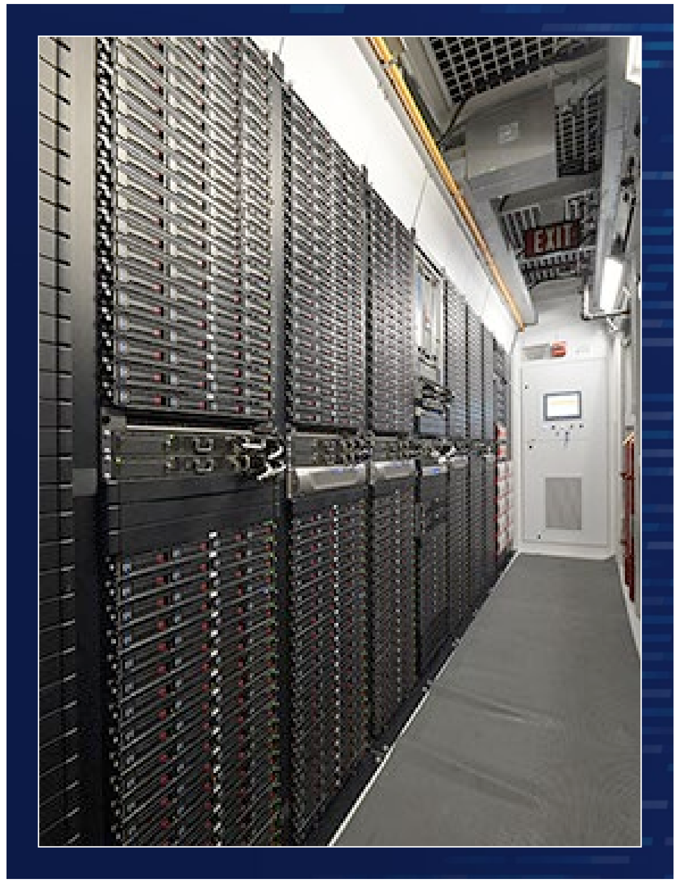
Only 13 years ago, the Laboratory's supercomputing capability, LLGrid, was composed of a single 16-processor system. Dr. Albert Reuther, Manager, Supercomputing Center, said that a "different kind of supercomputing" was clearly needed to meet the needs of Laboratory researchers. Since then, the capability has expanded to 10,000 processors across seven systems. In addition, Reuther said that the center differs from others like it because of the team's "focus on interactive supercomputing for high performance data analysis," and the "extremely 'green' computing center in Holyoke, Massachusetts, which allows our computers to run 93% carbon-free."

"This new level of supercomputing capability will be a key technology for the computational fluid dynamics (CFD) work performed in Group 74," said Dr. Nathan J. Falkiewicz, Structural and Thermal-Fluids Engineering, Group 74. Falkiewicz explained that the new capability will allow his team to take advantage of the parallelism inherent in existing CFD codes to significantly reduce simulation time for computationally taxing problems, as well as enable simulation for certain classes of problems that would otherwise have "prohibitively long" execution times without access to large core-count, high-performance computing clusters.

Reuther said that the LLSC exists today in large part because of the researchers who utilize supercomputing capabilities to produce cutting-edge research results, as well as many other supporters: "LLSC has been blessed to have the support of visionaries in the Director's Office, the Technology Office, and the Steering Committee who have seen the potential of supercomputing to enable all of Lincoln's missions." Reuther also credits the MIT Lincoln Laboratory Beaver Works Center for playing a critical role in the LLSC's collaborations with campus, the group leaders who have supported the team throughout their growth, as well as the Security Services, Information Services, Financial Services, Facilities, Contracting, and Technical Communications departments.

"Creating the Lincoln Laboratory Supercomputing Center has been a goal for the team for many years and it is tremendously rewarding to see it come to fruition," said Kepner. "Laboratory researchers will see continued improvement in the LLSC systems, MIT Campus will benefit from our unique interactive supercomputing technologies, and Laboratory and campus researchers will be able to collaborate more closely on their joint research projects."

For more information about the LLSC, visit the LLSC website or contact the team at LLSC-help@ll.mit.edu.



Interior view of the LLSC Holyoke location.

Dr. Jeremy Kepner Named Vice Chair of the Society for Industrial and Applied Mathematics (SIAM) Data Mining and Analytics Activity Group

MIT Lincoln Laboratory

The Bulletin

January 14, 2014

Dr. Jeremy Kepner, Computing and Analytics, Group 53, has been elected vice chair of the Society for Industrial and Applied Mathematics (SIAM) Activity Group on Data Mining and Analytics (SIAG/DMA) for the 2014-2015 term. SIAM Activity Groups (SIAGs) provide a more focused forum for SIAM members interested in exploring one of the areas of applied mathematics, computational science, or applications.

The purpose of SIAG/DMA in particular is to advance the mathematics of data mining, to highlight the importance and benefits of the application of data mining, and to identify and explore the connections between data mining and other applied sciences. The activity group organizes the yearly SIAM International Conference on Data Mining (SDM), and organizes minisymposia at the SIAM Annual Meeting, among other duties.

The mission of SIAM is to build cooperation between mathematics and the worlds of science and technology through SIAM publications, research, and community. One of four mathematical professional societies in the Joint Policy Board for Mathematics (JPBM), SIAM membershi is composed of students and professionals whose primary interests are in mathematics and computational science and their applications.

www.ll.mit.edu/LLSC