MIT Lincoln Laboratory Takes the Mystery Out of Supercomputing

By Dr. Jeremy Kepner

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.

Our long-term familiarity with Big Data allows us to apply the principles of interactive supercomputing to data collection and analysis, designing sensors, and developing algorithms.

In 2008, the 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.

Autonomous systems generate enormous amounts of Big Data. We need to be able to quickly and efficiently process that data and develop algorithms that allow autonomous systems to make intelligent choices.

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, MIT LL 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 labs 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 field-testing.

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 them 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 MIT LL – 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 do not have many traditional computational scientists on staff that are masters of detailed parallel performance techniques. Therefore, part of our job as supercomputing scientists and engineers is to design easy-to-use interfaces that are highly scalable and compatible based on our knowledge of the evolution of processor design.

We have been able to do this because the laws of physics gives 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 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 producing 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.

<sidebar>

Creating a Super Green Supercomputer Center

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.

References
[1] MIT President shares vision for the future of innovation, Sep 30, 2016, https://www.ll.mit.edu/news/reif-innovation-distinguished-lecture.html