BRINGING AI INTO YOUR AGENCY HPC ENVIRONMENT

Introduction

"Automation of federal government employee tasks could save between 96.7 million and 1.2 billion hours annually, with potential savings between $3.3 billion and $41.1 billion, respectively."

— AI-augmented government: Using cognitive technologies to redesign public sector work. Deloitte University Press

Is Your Agency Ready for the Convergence of AI and HPC?

Today's advancements in high performance computing (HPC) present new opportunities to tackle exceedingly complex workflows. HPC is well established as a key field for advanced computing. Many academic researchers, government agencies and private sector businesses have deployed HPC solutions, leading to breakthrough discoveries in fields like astrophysics, meteorology, seismology, engineering, genomics, medicine, and nuclear simulation.

The rapid pace of technology advancement means government agencies must consistently evaluate their HPC-enabled approaches. Artificial intelligence (AI), one of the most disruptive technologies of our time, creates new opportunities to transform how the public sector gets work done.

Agencies can harness the power of HPC and AI to solve unique problems, increase government responsiveness, improve employee productivity, while also saving time and money. Those embracing AI will maintain their position on the forefront of science; those who ignore its potential run the risk of falling behind their peers.

In this eGuide, we offer practical considerations for public sector HPC managers to incorporate AI into their agency's HPC environment and scale those capabilities to accommodate emerging workloads and increasing end-user demand.

Resources

Artificial Intelligence for the American People
This fact sheet provides details on how the U.S. Government is funding AI research and computing infrastructure, machine learning, and autonomous systems to improve every aspect of citizens' lives.

Artificial Intelligence and the Future of Government
Read this article to find out how AI will transform the way government employees will deliver citizen services in the future.

How Artificial Intelligence Could Transform Government
Smart technology can transform the public sector by automating tasks, but agencies must make difficult decisions on where and how to introduce new technologies.

How Governments Can Help Advance Artificial Intelligence
Read this article to discover how the world's technology ecosystem and government entities will need to work together to realize the full promise of AI.

Introduction

"HPC is a force multiplier for the U.S. economy, amplifying the impact of American labor and ingenuity."

— Patricia Damkroger, vice president, Intel Data Center Group

Evolving HPC and AI Workloads

Government institutions push the limits of HPC solutions for advanced problem solving. Deep learning capabilities and advanced neural networks can address challenges unimagined only a couple of decades ago. Commonly used HPC workloads like analytics are augmented – or supplanted – by AI to generate more meaningful insights from data.

Unlike past approaches, AI empowers HPC systems beyond simplistic rule-based instructions. Instead, AI evaluates data using an instruction set of 'theories' and algorithms. By learning from these theories, AI can better predict and understand the context using inference to fill in data gaps. AI models complement more traditional HPC solutions to reveal insights faster, and more comprehensively, than data-processing and analytics-based applications can by themselves.

As HPC-driven capabilities like simulation, modeling, big data analytics, and AI converge, hybrid workloads are becoming more and more common. One step in a hybrid workload might involve modeling and simulation, while the next step requires deep learning capability. Traditional HPC use cases like modeling and simulation are more compute-intensive, while newer workloads like AI are more data-flow intensive. The varied nature of these workflows can represent a big challenge for non-optimized HPC systems. However, a balanced multi-node system excelling at both unleashes new possibilities for increasingly automated discovery and human benefit.

Deploying and Augmenting Scalable Solutions for HPC

If your agency plans updates to its current HPC infrastructure to accommodate advanced AI workloads, consider these three steps:

1. Think holistically about your stakeholders and HPC needs.
2. Think holistically about your HPC solution.
3. Validate your HPC technology first, then scale it up.

Resources

The Case for Investing in High-performance Computing
If America fails to invest in HPC, our nation's economic power, ability to innovate and national security will be at risk.

The New Age of Innovation: Government's Role in Artificial Intelligence
Listen in on a dynamic conversation regarding the role of government and its implications for AI growth in national public safety, privacy and civil rights.

Is Government Ready for AI?
Everyday uses of AI that can talk, listen and see are coming. Is government ready?

STEP 1: THINK HOLISTICALLY ABOUT YOUR STAKEHOLDERS AND HPC NEEDS

"Right now, in the United States there's an incredible discussion taking place about AI, automation, the future of work, and privacy – but all of that is taking place in the private sector, academia, and think tanks. Government is woefully behind. Elected officials need to look at the facts and work to find common ground. Whether you are a conservative or a progressive, this future is coming."

— Rep. John K. Delany (MD-6), U.S. Congress

HPC systems are more than the sum of individual parts. To evolve and grow optimal HPC infrastructure for your agency's needs, hardware, software, and human skill sets must converge to develop a holistic HPC and AI implementation strategy. HPC managers seek to provide their users with the HPC infrastructure, support, and tools to make projects successful.

While private sector enterprises regularly adapt their HPC systems for defined workloads, government agencies face a very different challenge. In many government agencies, HPC systems support a diverse group of individuals, each with unique research needs. To maintain quality of service (QoS) for stakeholders, an HPC system supporting a diverse array of workloads requires scalability, performance, adaptability, and future-proofing. HPC managers must always seek to understand the needs of their users so the HPC investment is used as much as possible by stakeholders.

AI can solve problems by "learning" from supervised and unsupervised examples rather than depending on a formalized set of equations or rules. For this reason, the role of a combined HPC and AI solution offers new capabilities and opportunities. By mapping your organization's HPC and AI needs against use cases and stakeholder demands, you can determine if your current HPC implementation meets its requirements, or what other upgrades or modifications must be prioritized to enable more complex, demanding future scenarios including AI.

Takeaways and Outcomes

- **User needs drive underlying HPC technology.**
- **Evaluate needs now, be prepared for the future.**

Resources

AI on HPC Infrastructure in Three Steps
This interactive document outlines the steps, considerations, and insights for any organization planning on using its current HPC infrastructure to accommodate AI workloads.

AI and HPC Workload Convergence
Watch this video to discover how AI and HPC workload convergence running on an Intel® - based HPC infrastructure can accelerate your innovation.

Planning for the Convergence of AI and HPC
This insideHPC special report looks at the convergence of the HPC and the data-driven AI communities as they are arguably running similar data and compute intensive workloads.

STEP 2: THINK HOLISTICALLY ABOUT YOUR HPC SOLUTION

Software Selection

End-user experience is paramount. As your agency's needs grow, software capability needs to grow with it. Various scientific endeavors like modeling, simulation, visualization, and AI benefit from applications designed and optimized for those tasks. If you are unsure where to begin the process of identifying the ideal software solution, a good starting point is researching which software solutions best serve other experts driving similar workloads successfully.

Depending on the intended HPC usage scenarios, software resources from Intel, independent software vendors (ISV), and the open source community can offer a jumpstart toward AI-capable systems. By first considering the software solutions needed, it is easier to plan for a physical infrastructure that provides the ideal underpinnings for applications.

Resources

Advanced Analytics and AI Strategy
Advanced analytics and AI can shape the government's future. Watch this video to learn about two key imperatives that will help the public sector to achieve success.

Framework Optimization Training
Whether you're starting out or already an expert, the Intel® AI Academy provides essential learning materials, community, tools, and technology to boost your AI development.

Intel Software Libraries
Simplify and speed your development with software optimized to maximize hardware performance.

STEP 2: THINK HOLISTICALLY ABOUT YOUR HPC SOLUTION

Software Application and Development Environment

If ready-made applications are inadequate for your agency-specific usage scenarios, you will need the right development skills to modify or create them. For scientific computing, the HPC community offers a large number of high-performance libraries for HPC systems of all sizes.

However, optimizing some open source options will require developers with skills and experience in parallel computing or mathematical algorithms. Fortran*, C/C++*, Java*, and Python* code many HPC and AI applications. Intel's HPC interoperable framework, which supports all these languages, helps accelerate the development process. It also supports a common development model for improved code portability, plus software tools to accelerate HPC application modernization initiatives.

	HPC	Big Data (Hadoop/Spark)	AI
Primary Languages	C/C++/Fortran	Java/Python/Scala/R	Python/C++/R
Scaling	Strong	Weak	Both
Cluster Stack Management	Custom	Cloudera Manager/Apache/ Ambari/MapR Control System	NA
Schedulers	Slurm, PBS	YARN, Mesos	NA
File Systems	High-performance Distributed POSIX	Distributed Across Node Local Storage – HDFS	Posix or HDFS
OS	RedHat/CentOS SuSE	RedHat/CentOS SuSE	Ubuntu

STEP 2: THINK HOLISTICALLY ABOUT YOUR HPC SOLUTION

Deep learning frameworks have a central role in maximizing the potential of AI on HPC systems. Popular frameworks like TensorFlow* and Caffe*, optimized for performance on Intel® architecture-based platforms, speed both deep learning training and inference workflows. Tapping the Intel® Math Kernel Library (Intel® MKL) and the Intel® Math Kernel Library for Deep Neural Networks (Intel® MKL-DNN) TensorFlow offers significant performance gains. Similar benefits occur when pairing the Intel® Distribution of Caffe* with Intel® Xeon® processors.

Physical Infrastructure

Before building upon your existing HPC infrastructure, determine first which ingredients perform at an optimal level, and which create bottlenecks that impede advanced workflows. A platform based on Intel® Xeon® Scalable processors and other innovative Intel® technologies for HPC offer the foundation for the most demanding workloads. Elements like processors, memory, storage, and fabric all weigh into the equation.

Wherever possible, preserve your existing infrastructure investments. However, if anticipated use requirements exceed your HPC system's hardware capability, upgrades may be necessary. Budget for the needed components and prioritize their implementation over time as workloads require it. By taking a staged approach, you can help your existing HPC infrastructure evolve and scale over time.

Flexibility of HPC systems for academic environments

The **Texas Advanced Computing Center's (TACC) Stampede2 HPC system** helps thousands of scientists and researchers achieve breakthroughs in AI, modeling, simulations, and much more. To accommodate the diversity of projects, Stampede's flexible and scalable HPC infrastructure supports more than 40 major science and engineering applications at any given time, alongside thousands of applications used by small groups or individual researchers. The current iteration Stampede2[1] ranks at #12 on the June 2017 Top500.org[2] listing of the fastest supercomputers worldwide.

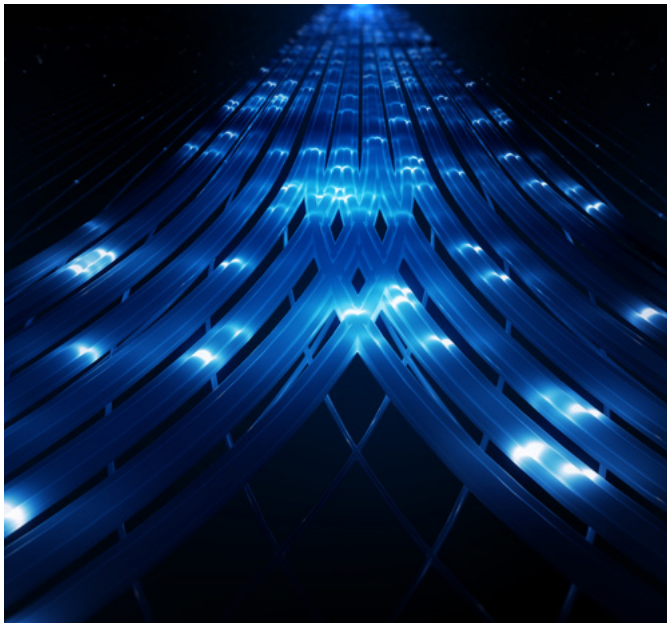
Takeaways and Outcomes

- **Determine if current data and compute infrastructure generate desired insights.**
- **Map technology needs against today's business opportunities.**
- **Consider software and evaluate key AI frameworks early on.**
- **Choose a flexible, scalable HPC platform which can scale to accommodate complex AI workloads.**

Resources

Accelerating AI on Intel® Xeon® Scalable Processors
Intel enhanced the Intel Xeon Scalable platform specifically to run high-performance AI workloads alongside the other cloud and data center workloads agencies already run.

Optimizing HPC Architecture for AI Convergence
Explore key considerations for organizations looking to bring AI into their HPC environment, and steps they can take to ensure the success of their first forays into HPC/AI convergence.



STEP 3: VALIDATE YOUR HPC TECHNOLOGY FIRST, THEN SCALE IT UP

Different agencies have different needs for HPC and AI deployment, so the process is not a one-size-fits-all proposition. Some supercomputers used in government and academic environments are supported by many internal staff members with all the expertise necessary to manage and grow their HPC environment. Organizations lacking dedicated human resources and expertise should consider working with Intel, an original equipment manufacturer (OEM), or a consultant who can help them accelerate system deployment.

With your team in place, validate your test system before completing full-scale implementation. The testing process helps determine many factors like the adequacy of performance levels for challenging workloads, whether bottlenecks exist, and the value of data insights obtained. If the test system proves unable to meet your organization's needs, regroup to determine alternate, more effective solutions.

Takeaways and Outcomes

- **Identify your deployment team.**
- **Perform testing and validation to confirm system adequacy before rollout.**
- **Identify experts for ongoing administration, maintenance, and system growth.**

InsideHPC: How AI is Helping Scientists with the Large Hadron Collider

Watch this video to learn how AI is being used in the design of experiments for the Large Hadron Collider, the world's largest and most powerful particle accelerator.

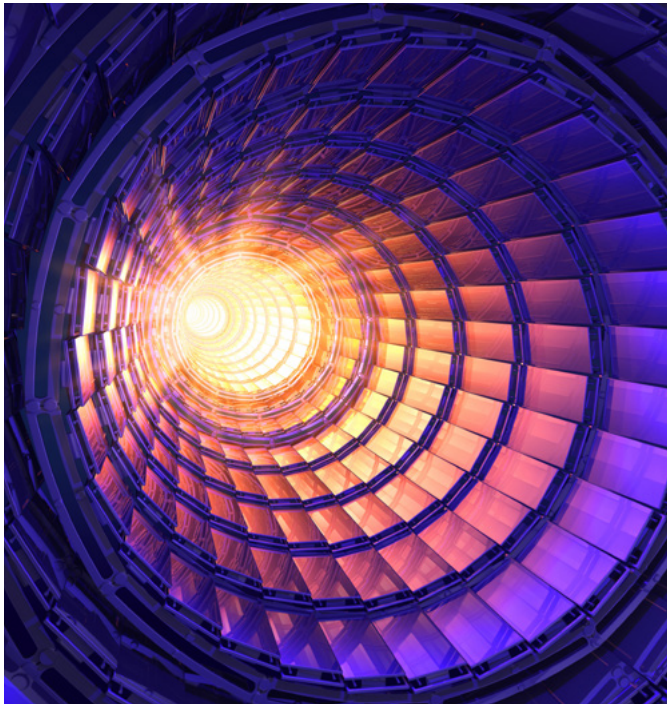
HPC Implementation Enables Extreme Weather Modeling[3]

Read this article to find out how one organization dedicated to weather monitoring has applied a careful evaluation process to define its HPC infrastructure, and software needs to solve the major challenge of predicting hazardous weather conditions. Their task – quite literally – can be a matter of life or death. Driven by the need to predict the path of destructive hurricanes well in advance of landfall, researchers optimized their HPC systems for the important task. Through rapid analysis of available meteorological data, researchers can alert local authorities well in advance should weather conditions indicate the need for evacuation.

Resources

AI: How the Infrastructure You Have Can Give You What You Need
Reduce AI risk and lower the barriers to entry with existing infrastructure.

Start Small, Scale Up: AI's 'Not-So-Secret' to Success
The pace and scope of AI breakthroughs is increasing, but enterprises that are just starting out are by no means lagging.



Summary

Three practical steps can help you determine ideal infrastructure requirements for a future-proofed HPC system:

1. **Think holistically about your stakeholders and HPC needs.**
2. **Think holistically about your HPC solutions.**
3. **Validate your HPC technology first, then scale it up.**

Learn More

To find out more about Intel's HPC technologies and how they can advance your agency's business and scale to meet the needs of increasingly complex workloads like AI, talk to your preferred system provider or learn more at intel.com/hpc.

Other Resources

AI-HPC is Happening Now
This white paper addresses the convergence of HPC and data driven AI communities, the existing ecosystem of affordable technologies, and the current software ecosystem that everyone can use.

Intel AI at SC18
The intersection of AI and HPC is transforming industries from healthcare to transportation while solving some of the toughest socio-economic and scientific challenges in our world. This blog provides a look at how Intel's hardware and enabling software portfolio of unprecedented choice and scale is powering this next wave of discovery.

Intel® technologies enable advanced AI workloads today

Intel AI can empower government agencies and their partners to do more with less. Satellite images and overhead video combined with deep learning can unlock new possibilities in defense, disaster response and mapping. National Laboratories and research institutes can make breakthrough discoveries, leveraging datacenters and high-performance compute that are optimized for AI and have the sophisticated memory to handle unprecedented amounts of data.

The ever-expanding demands of today's large and complex HPC workloads require powerful underlying technologies that deliver:

- Improved performance
- System resilience
- Energy efficiency
- Usage flexibility
- Code portability

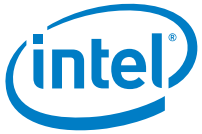
Potential bottlenecks such as processor speed, memory, and storage must each be addressed in a holistic HPC system to derive the greatest benefit from each element. For vast multi-rack HPC systems, the latency of communication among nodes can also impact overall system speed. The interconnecting fabric can have a profound impact on overall system throughput too.

Intel's leadership in platform innovation, supported by a broad software ecosystem, is paving the way for future AI capabilities. Intel's HPC technologies include many innovations to help organizations get the most from their most demanding workloads:

- Intel® Xeon® Scalable processors
- Intel® Omni-Path Architecture
- Intel® FPGAs
- Intel® Optane™ technology
- Intel® 3D NAND SSDs
- Intel® Advanced Vector Extensions 512 (Intel® AVX-512)

Contact Us

Intel Public Sector Solutions
Intel.com/publicsector
publicsector@intel.com



[1] **https://www.nextplatform.com/2017/07/25/texas-advanced-supercomputing-center-taps-latest-hpc-tech/**

[2] **https://www.top500.org/list/2017/06/**

[3] **https://www.hpcwire.com/2017/08/28/nersc-scales-deep-learning15-pflops/**