# Extreme-Scale Computing for Precision Medicine

## Sustained innovation in extreme-scale computing will help usher in a transformative era of personalized biomedicine

## Executive Summary

Precision medicine promises a transformative era of personalized health, with customized treatments and diagnostics tailored to each individual's unique biology. Powered by data- and simulation-driven insights into biomedical and biochemical processes, precision medicine has the potential to replace today's trial-and-error medicine with personalized, predictive treatments for cancer and other diseases. In doing so, precision medicine promises to transform healthcare and improve lives and economies around the world.

But these promises will only be fulfilled through dramatic improvements in computational performance and capacity, along with advances in software, tools, and algorithms. Extreme-scale computers—machines that perform one billion billion calculations per second and are over 100 times more powerful than today's fastest systems—will be needed to analyze vast stores of clinical and genomic data and develop predictive treatments based on advanced 3D multi-scale simulations with uncertainty quantification. Precision medicine will also require scaling systems down, so clinicians can incorporate research breakthroughs into everyday practice.

As an innovation catalyst for the global technology industry, Intel is leading the way to address the computational challenges of precision medicine. Intel is:

- **Creating** a framework for designing next-generation systems with higher performance, throughput, and scalability
- **Holistically advancing** the system elements needed for exascale and beyond
- **Collaborating** with life science and technology leaders to facilitate the development and commercialization of tools and algorithms for precision medicine

Because precision medicine and exascale computing are complex and strategically important, they require sustained investments and coordinated R&D between the public and private sectors. In addition to improving the quality of healthcare, these investments will provide capabilities that can more accurately predict climate change, preserve biodiversity, safeguard national security, drive economic growth, and improve quality of life. The research firm IDC, in a project funded by the US Department of Energy (DOE), found that each dollar of high-performance computing (HPC) investment generates a return on investment of USD 514, and

**William Magro**
Intel Fellow and Chief Technologist,
HPC Software
Intel Corporation

**Ketan Paranjape**
General Manager, Life Sciences
Intel Corporation

**Mark K. Seager**
Intel Fellow and HPC Chief Technical
Strategist
Intel Corporation

## Table of Contents

that the average time from HPC investment to innovation is 1.9 years.[1] By contrast, the current average lag between typical medical investment and innovation in practice is estimated at 25 years.[2]

Government leaders and policymakers have the opportunity to maximize healthcare, drive vibrant economies, and promote national leadership by:

- **Investing in extreme-scale computing, biomedical research**, and other areas to advance precision medicine and move it into clinical practice.
- **Developing patient-centered precision medicine initiatives**, ensuring that patients are empowered stakeholders who benefit from their genomic and health data.
- **Bringing government labs, regulatory agencies, and resources** together with private industries of technology, pharma, healthcare providers, health instruments, research, and academia to address emerging issues.
- **Collaborating across agencies** and with technology innovators to ensure that next-generation platforms meet the needs of relevant stakeholders.
- **Establishing an open, interoperable environment** that fosters innovation, collaboration, and economic growth, and enables innovations to be broadly applied at all scales.
- **Creating policy frameworks** that address issues such as ethical standards, patient privacy, and consent.
- **Expanding workforce skills** to develop and utilize emerging capabilities.

## A Digital Revolution in Medicine

Digital technology has brought powerful capabilities to medicine and healthcare. Today's pharmaceutical researchers use sophisticated 3D molecular dynamics models to design targeted drug treatments. Health professionals review a comprehensive, up-to-date record of each patient's health and medical history as they formulate treatment plans. Clinical researchers mine de-identified records to develop evidence-based treatments. At leading-edge cancer centers, clinicians sequence the DNA of patients and their tumors, and scour vast databases to identify treatments that target a given mutation while leaving healthy cells unharmed.

These are significant advances, but they're evolutionary in nature. Now, as digital technologies advance and our understanding of the life sciences increases, we can see the shape of a revolutionary era of predictive biology, precision medicine, and personalized healthcare.

### Predictive Biology, Personalized Medicine: Approaching an Inflection Point

Precision medicine starts with genome sequencing, a data-intensive process in which scientists use massive computational resources to assemble, analyze, compare, and explore genomes and achieve and share novel insights.

These steps are demanding but manageable with current and near-term technologies. Given the political will and continued advances in computing performance and algorithms, it's not unreasonable to expect that genomic sequencing can become integrated into clinical workflows within a few years. By 2020, cancer patients and their care teams may well be able to diagnose the disease and conduct secondary analysis

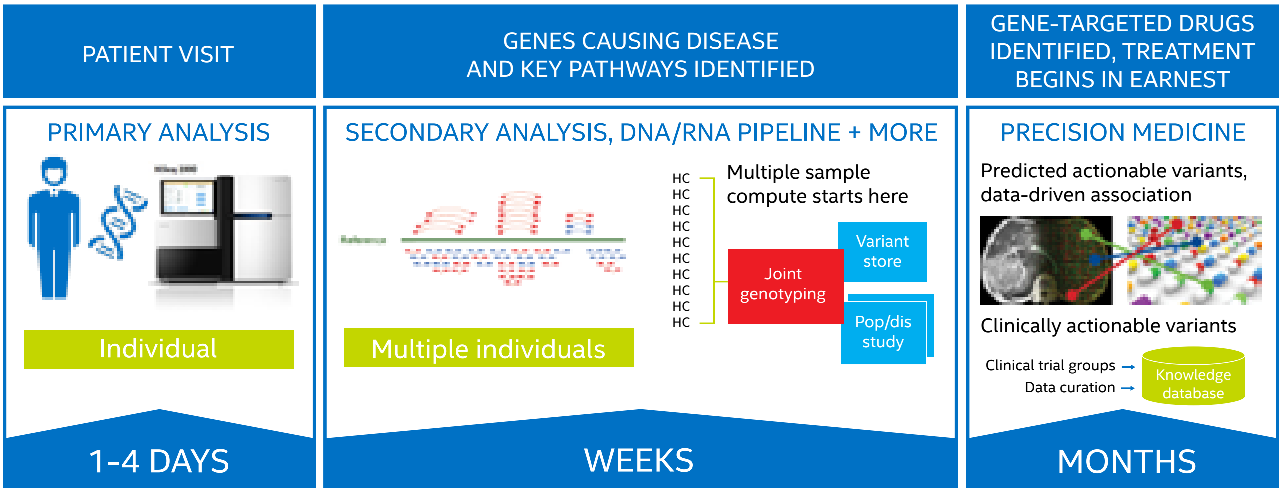

**Figure 1.** Current state of precision medicine

to identify causal genes and resulting pathways to disease expression within a single day—and initiate treatment with targeted therapeutics within another few days (Figure 1).

Precision medicine involves much more than genome sequencing, however. The biological world is becoming more digital and measurable, with data being generated by devices ranging from personal health monitors to state-of-the-art cryo-electron microscopes (cryo-EM). This and other data provides a foundation for new insights into biological constructs and processes. With rigorous consent and privacy protections in place, life science and medical researchers will seek to combine and analyze data sources such as:

- Data from the Internet of Things, generated by billions of lab instruments, clinical monitoring systems, implantable diagnostics devices, and other sensor-based and embedded technologies
- DNA data from cancer tumors at multiple points throughout the treatment cycle
- Longitudinal electronic health records (EHRs) that include each patient's genomic profile, treatments, and outcomes, with both structured and unstructured data from primary and specialty care providers, labs, and other sources
- Data from individuals' personal health devices and personal health records
- Clinical research data
- Pharmaceutical data

The challenge is to capture the full value these rich data stores—to bring the data together, scrutinize it with increasingly sophisticated analytic methodologies, apply cognitive computing capabilities, and combine the answers with new biological insights. The resulting knowledge can provide medicine with the types of modeling and simulation capabilities that have already transformed fields ranging from oil and gas exploration to airplane design. At this point, precision medicine becomes revolutionary in its impact.

## Amplifying the Science of Medicine

Computer simulation is increasingly recognized as a third method of scientific discovery, supplementing the traditional aspects of theory and experimentation. Modeling, simulation, and the scientific study of uncertainty (uncertainty quantification) have helped transform many scientific and technical fields where complex and subtle physics and chemistry are accurately modeled with sophisticated mathematics and highly efficient solution techniques. Powered by high-performance computing, practitioners and researchers perform virtual experiments that are too costly or impractical—or simply impossible—in the physical world. For example:

- Weather forecasters routinely simulate a hurricane's potential paths, providing advance warnings that save lives.
- Automobile and aircraft manufacturers simulate engine systems and body designs, reducing environmental impacts and creating more fuel-efficient, aerodynamic vehicles that save lives through better crash-worthiness.
- Civil engineers model the properties of bridge dynamics and aging, enabling innovative and distinctive bridge designs while creating a finished product they are confident will bear its designated loads over the bridge's lifetime.

In each case, modeling and simulation are used to optimize costs and outcomes, reduce time-to-results,

## A Watershed Moment for Biology

Extreme-scale computing is going to be a watershed moment for biology. We'd never ask an engineer to build a bridge or design an airplane without modeling how it's going to perform in the real world. But doctors do the equivalent every day.

The dream of extreme-scale computing is to model each patient's health at a given moment in time and predict how it will be affected by a given intervention. That's a national-scale challenge. It's Big Data to analyze the huge volumes of data from longitudinal EHRs, genomics, proteomics, medical imaging, and so forth. It's Big Compute to simulate biological processes, combine that with the data, and predict the health trajectory for each unique patient.

If we succeed, we're going to see a fundamental shift that will transform health. But we're not going to get there without national investments to achieve exascale computing and beyond. Exascale enables computing at the highest scale (for data and for calculations per second). Equally important, it will enable affordable and deployable HPC everywhere at every scale: teraflops at the bedside, gigaflops inside our medical sensors, petaflops at our regional hospitals, and of course exaflops at our research centers.

There are positive incentives for agencies and industry in the US to work together and accelerate cross-transfer of knowledge in modeling and simulation. Converging agency missions yield innovative solutions to the nation's most pressing challenges. Together, the technology and health communities need to all strive toward the shared national imperative of saving lives, improving quality of life while reducing cost.

**Jason Paragas, MD**
Director of Innovation
Lawrence Livermore National Laboratory

and reduce guesswork and rework. They amplify the science of weather forecasting, auto design, and bridge construction. Human ingenuity and creativity remain paramount, but there's less reliance on the art of those disciplines.

Human biology is incredibly complex, but as our understanding of the body's biomedical and molecular systems and subsystems increases, our ability to model their behavior also advances. We don't know how far we'll be able to go in unlocking the biological world, but there's every reason to expect that much progress is possible toward building more accurate and detailed biological models and simulating the workings of the body—from large-scale elements such as the brain or circulatory system down to the molecular workings of a mutated cancer cell and the trillions of microorganisms that live in and on us.

If we arm researchers and clinicians with powerful, knowledge-based simulation tools—along with data-rich models and next-generation methods of data analysis—we empower them to potentially improve diagnosis, treatment, and outcomes. These approaches provide opportunities to:

- Simulate and predict how a disease is likely to progress based on all that is known about the patient's health, including the patient's full genomic profile, proteomic pathways, health history, lifestyle factors, environmental exposures, DNA data about the microbiome, and biological and biochemical data regarding the proteomic and epigenetic factors governing the behavior of relevant genes.
- Compare the results against diverse clinical and population databases, to make more informed predictions about the course of an individual's illness and make more informed choices among treatment options.
- Develop and validate innovative treatments more rapidly, simulating the risks and benefits of different treatments and conducting virtual experiments and virtual clinical trials. Virtual trials can test a new treatment across a greater population, enabling pharmaceutical researchers to validate the efficacy and safety of potential treatment breakthroughs more quickly. This can help reduce the likelihood of adverse drugs reactions and enable innovations to reach patients more quickly and at lower cost.

These capabilities can help healthcare become more personalized, predictive, preventive, and cost-effective. They have transformative potential for human lives, societies, and economies. Consider the following examples, which are internationally significant as well as being among those targeted by the United States' Precision Medicine Initiative. Extreme-scale computing and advanced data analytics can help accelerate progress in addressing each of the following.

### Cancer

Cancer kills 8.2 million people worldwide each year,[3] and data forecasts show that more than 1.6 million Americans were diagnosed with cancer in 2015.[4] Can we create digital models sophisticated enough and personalized enough to accurately predict which treatments will work best for an individual patient? Can we determine which cancer-causing gene mutations are about to occur and prevent that from happening? Extreme-scale computing can help deepen our understanding of cancer biology and make it possible to rapidly evaluate new therapies.

### Cardiovascular Disease

Heart attacks, stroke, and other cardiovascular diseases account for 17.5 million deaths around the world each year.[5] Today's cardiovascular surgeons can call on robotic assistants and data-intensive 3D imaging to enhance success. Tomorrow's medical teams want to simulate the effects of a surgical or medical intervention using a much deeper and more accurate digital model of the patient's heart—not a generic model of the human heart and circulatory system, but a model that reflects an individual patient's full and unique physiology and health picture. Advanced modeling can also provide deeper understanding of how to prevent and manage cardiovascular illness.

### Neurologic Disorders

Up to one billion people around the world suffer from diseases of the brain and nervous system, including Alzheimer's, multiple sclerosis, traumatic brain injury, and hundreds of other conditions.[6] Dementia alone is expected to exact a worldwide economic toll of USD 1 trillion by 2018.[7] Parkinson's afflicts as many as 7 million people globally and currently has no cure.

The causes and cures of neurologic disorders are often poorly understood. Modeling and simulating the structure and functioning of the brain and nervous system can help neuroscientists tease out the complex interplay of genetic, biochemical, electrical, structural, and other factors that are often at work.

### Sepsis

Sepsis occurs when the immune system goes into all-out attack mode in response to a serious infection. It strikes more than 30 million people each year,[8] including one million Americans.[9] The condition is so deadly that a 2014 review found that even when treated in an intensive care unit, one-third of sepsis patients died without leaving the hospital.[10]

Sepsis often kills within hours. It is difficult to identify and its underlying biology is still poorly understood. Researchers envision a time when precision medicine will let clinicians simulate the development of an infection and intervene to prevent sepsis or halt its damage.

### Antimicrobial Resistance

Antibiotic-resistant infections strike millions of people in every region of the world[11] and kill more than 23,000 Americans each year.[12] Pathogens that are resistant to the dominant medications used to treat HIV, tuberculous, malaria, and influenza have also emerged. Advances in life science modeling and simulation will be key to identifying and treating antimicrobial-resistant infections as well as developing new antimicrobials.

## Bringing Precision Medicine to Life: Extreme-Scale Computing

Are these scenarios achievable? We'll never know if we don't try. Nations around the world are engaged in the effort, calling precision medicine the next global space race. The People's Republic China is finalizing plans for what is expected to be a 15-year, RMB 60 billion genome sequencing program focusing on specific cancers.[13] Genomics England, established by the United Kingdom's Department of Health, is sequencing the genomes of 100,000 National Health Service patients who suffer from rare genetic disorders.[14] The United States is launching a USD 1 billion National Cancer Moonshot to fund research aimed at new cancer detection and treatments.[15] The US National Institutes of Health (NIH) is also spearheading a one-million-patient cohort program as a first step in a USD 215 million Precision Medicine Initiative.[16]

Extreme-scale computers—capable of exascale and beyond that zettascale levels of performance—will be needed to enable rapid progress in precision medicine, including the five diseases and disorders mentioned above. Researchers will need these powerful systems to manage and analyze the exponentially-expanding world of biological data, develop sophisticated models that represent the body and its structures, and simulate biomedical processes. They will need powerful systems to visualize their findings, explore complex questions through interdisciplinary collaborations, and translate scientific results into practical insights. This need for extreme-scale computing aligns medicine and the life sciences with fields such as aerodynamics, biofuels, renewable energy resources, innovative materials science, climatological analysis, and others that will benefit from accelerated advances in high-performance computing.

Precision medicine will also benefit from advances in natural language processing, deep learning, and other cognitive computing and analytic techniques that will enable researchers to extract greater insights from their data. Clinicians will need compatible, scaled-down systems and decision-support tools to integrate genomics into daily practice.

> Precision medicine is being called the next global space race. Nations around the world are undertaking large-scale projects to make it a reality.

### Data Drivers

The sheer volume of biomedical data, the diversity of data types, and the number of workflows producing and consuming data are enormous—and growing. Genome sequencers, cryo-electron microscopes, and wearable health devices are just three examples of the way massive data growth is driving the need for exascale computing.

### Genome Sequencers

In early 2014, a single genome sequencer produced 3.6 TB of data every six days.[17] That data generally becomes valuable only when the resulting genes are assembled, analyzed, compared, and studied. With the number of sequencers rising, their performance climbing, and costs falling, researchers say that by 2025, the data requirements for genomics alone will equal or surpass those of the entire field of astronomy or all of YouTube*.[18]

To sequence each new patient's cancer over the course of a year, treatment teams might want to perform one whole genome sequence and then track changes in a tumor's mutations three further times throughout the year. That work is estimated to consume nearly four exabytes of data and require 1.5 billion "processor hours" using today's server platforms. In addition to direct patient care, researchers want to incorporate privacy-protected, de-identified genomic data into the study of populations and subgroups. Data-intensive studies of plant and animal genomics can also contribute to treatment advances along with providing deeper insights into aging, stem cells, neural regeneration, and other subjects related to human health.

### Molecular Imaging and Cryo-Electron Microscopy

Molecular imaging brings together structural biology, computational chemistry, biophysics, and bioinformatics to answer previously unanswerable questions. Computationally intensive cryo-electron microscopy, coupled with high-performance data analytics, is now being developed to elucidate accurate 3D chemical composition, protein structures, and motion at the molecular/atomic level.

Proteins are at the limit of imaging technologies today. With extreme-scale computing, it is projected that cryo-EM will become the dominant method of determining protein structures and protein expression chemical pathways, ushering in a new era of deep insight into the human proteome. Harvard University and the Dana Farber Cancer Institute, working collaboratively as an Intel® Parallel Computing Center, are using cryo-EM to study the structure of proteins relevant to cancer immunology and AIDS research.

### Wearable Technologies

Smart wearable technologies may offer breakthroughs in the diagnosis and treatment of neurodegenerative diseases. In an ongoing collaboration between Intel and the Michael J. Fox Foundation for Parkinson's Research,[19] Parkinson's patients use innovative wearable devices to sense and measure some 50 metrics relating to the velocity, orientation, and gravitational forces of their movements as they go through their normal lives. The research team uses data-intensive algorithms developed by Intel to assess each patient's activity level, tremor, nighttime motor symptoms, and gait every few seconds. The insights gained from analyzing the vast quantities of data being generated and correlating it with other information may lead to therapeutic breakthroughs for Parkinson's disease.

## Challenges for Extreme-Scale Computing

While Intel's commitment to Moore's Law has brought steady increases in computational performance and capacity, extreme-scale computing will require disruptive innovation in platform design, deep collaboration between technology innovators and leading HPC users, and thoughtful reworking of applications and tools.

The US Department of Energy's Advanced Scientific Computing Research program has identified ten challenges that must be solved to make exascale platforms a reality[20]:

- More energy-efficient technologies for circuit, power, and cooling
- Interconnect technology advances to move larger data volumes faster and with less energy
- Higher-capacity and higher-bandwidth memory technologies
- Scalable system software that enhances power efficiency and resilience
- Programming environments that express greater levels of parallelism, resilience, and data locality
- Data management software that can handle the rising volume, velocity, and diversity of data
- Solution algorithms that perform well on extreme-scale systems
- Algorithms to facilitate exascale discovery, design and decision-making
- Techniques to ensure resilience and correctness in systems with tens of thousands of system components
- Tools and environments to increase the productivity of those using extreme-scale computers

## Meeting the Challenges: Intel's Role

Intel is uniquely positioned to lead the drive to extreme-scale computing. The company is today's HPC leader and is aggressively focused on delivering the energy-efficient system performance, scalability, resilience, and other attributes needed for exascale computing and beyond. Intel brings to this challenge holistic HPC systems knowledge, an extensive technology and product portfolio, deep design resources, industry-leading expertise in semiconductor engineering and manufacturing, and a broad network of partnerships and collaborations. With a longstanding commitment to digital solutions for healthcare and life sciences computing, Intel is also collaborating to advance progress toward precision medicine.

Nearly 90 percent of the systems on the November 2015 TOP500* ranking of the world's fastest supercomputers incorporate Intel technologies.[21] In an important step toward exascale computing, the DOE selected Intel in April 2015 to deliver the world's most powerful supercomputer. The Aurora system, which Intel is building collaboratively with Cray, provides a flexible technology blueprint for future large-scale computers, offering advanced integration that will help drive price/performance, power efficiency, and performance to new levels. Aurora will reside at the DOE's Argonne Leadership Computing Facility (ALCF) and will be used by multiple DOE labs to accelerate innovation in extreme-scale computing, the life sciences, and other fields.

Reflecting the complexity and interconnectedness of extreme-scale computing challenges, Intel is taking a systems-level architectural approach to designing platform technologies. As a key step forward, Intel has created Intel® Scalable System Framework (Intel® SSF), a next-generation blueprint for developing high-performance, balanced, efficient, and reliable HPC systems at diverse scales. Intel is also:

- Holistically developing and integrating the system elements needed for exascale and beyond.
- Facilitating the development of open-standards-based tools and algorithms for precision medicine and other workloads.
- Collaborating with leaders in the life sciences and other fields to ensure that next-generation systems will address their requirements.
- Investing in leading-edge silicon manufacturing to ensure the tight integration, system capacity, and capabilities for practical, extreme-scale computing.

Intel's open, integrated approach helps promote rapid progress toward extreme-scale computing and democratize the resulting benefits. Intel has demonstrated a longstanding commitment to industry standards and collaboration, including open-source solutions. In addition, Intel leads an open ecosystem of companies and organizations that create their own systems and solutions on top of Intel platforms. Compared to more closed, proprietary environments, Intel's approach can drive a robust technology economy that delivers cost and flexibility advantages for scientific and business computing.

## Flexible Framework for Next-Generation System Design

As a flexible blueprint for extreme-scale computing, Intel Scalable System Framework aligns with future application requirements and technology directions. Moving key system elements closer to the processor, Intel SSF offers outstanding performance, throughput, and memory capacity for the broad range of HPC workloads, including traditional HPC simulations, big data analytics, visualization, and machine learning.

Intel SSF provides a cohesive yet flexible design space with highly configurable elements, along with a reference architecture that supports

### Unlocking Potential

Extreme-scale computing delivers a tremendous opportunity to explore and model the complexity of biology. It's likely to play a major role in identifying and developing new candidate precision treatments, evaluating their potential, and providing the insight necessary for the advances in precision medicine to reach the general population.

Unlocking the potential in the growing sources and volume of information available to probe complexities of diseases such as cancer will require significant advances in extreme-scale computing technologies. The future of precision medicine will be shaped in part by how extreme-scale computing is utilized and by what advances in computing technologies will enable.

**Eric A. Stahlberg, PhD**
HPC Strategist, Leidos
Biomedical Research, Inc.
Frederick National Laboratory for
Cancer Research

> Extreme-scale computers will enable biomedical researchers to crunch more data, answer deeper questions, and run more sophisticated models of biological processes.

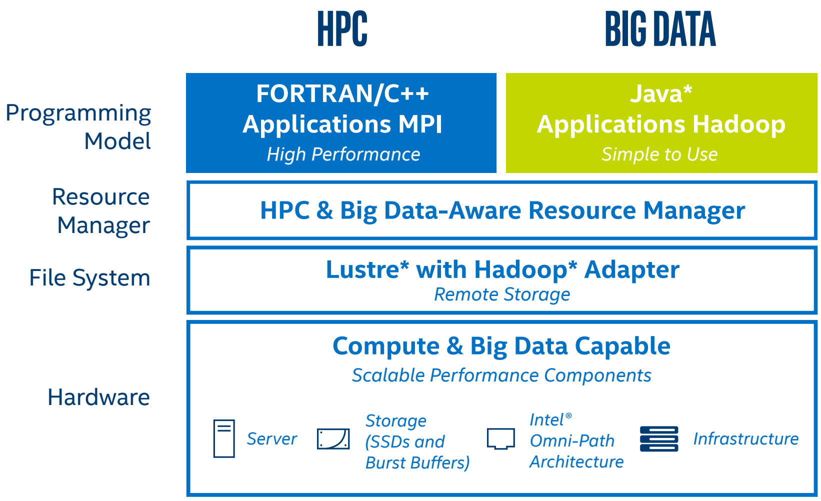

**Figure 2.** Flexible architectural framework for high-performance computing

static and dynamic configurability. Compute elements can be homogenous or heterogeneous, enabling system manufacturers and user organizations to meet diverse workload requirements at multiple scales. This configurability and flexibility promotes wider use of HPC systems in diverse industries, delivering exceptional performance on compatible systems in a range of sizes. Intel SSF also enables application portability and scalability across differently configured systems based on Intel SSF, offering longevity over multiple generations of systems.

These traits allow construction of a converged architecture for high-performance data analytics (HPDA), the field formed by the convergence of traditional HPC and big data workloads. HPC and big data analytics have evolved independently, and the architectures and technologies used for modeling and simulation have differed from those used for data integration and analytics. Increasingly, however, the methodologies of big data analytics are being applied in HPC contexts to gain insight into increasingly vast and complex volumes of scientific data. In the life sciences, next-generation analytics—often powered by real-time data from connected devices and performed with techniques such as deep learning—are providing inputs into sophisticated models that drive new understanding of the natural world.

These new, highly complex HPDA workloads require massive compute resources along with extensive data handling and storage, causing the big data/HPC convergence to occur at level of the system infrastructure as well as middleware, workloads, and workflows. Intel SSF gives converged HPDA workloads the high performance, large memory capacity, and high throughput they require (Figure 2). Intel SSF also:

- Allows data analytics and HPC workloads to run on the same platforms. This infrastructure convergence improves total cost of ownership and provides greater flexibility to manage workflows and increase system utilization.
- Eliminates the need for large-scale data movement between big data and HPC systems. Avoiding this data movement reduces the cost and complexity of data-intensive computing while speeding application performance. It also minimizes data motion by allowing in situ distributed data analysis.
- Enables HPDA workloads to take advantage of the HPC world's high-performance parallel file systems and communication fabrics.
- Simplifies and speeds the development of tools and algorithms through a unified system software stack.

## Breakthrough Technologies

Complementing Intel SSF, Intel is moving away from the traditional view of computers as discrete collections of processors, memory, networking, and storage technologies. Moving forward, we're seeing integrated units holistically designed from a systems perspective into new composite functions that dissolve the historic imbalances between processor performance, memory bandwidth and latency, interconnect bandwidth and latency, and I/O. Servers will have processors with direct connections to high-speed, low-latency interconnect fabrics and Ethernet* networks. New non-volatile memory (NVM) will function as both RAM and extremely low-latency, high-bandwidth storage. This higher degree of convergence and integration helps reduce system costs and overall system power while enhancing system resilience. It enables major advances in modeling and simulation, high-performance data analytics, machine learning, and visualization—the cornerstones of precision medicine and other extreme-scale computing workloads.

Several recent technology advances point the way toward practical, extreme-scale computing platforms. For example, Intel® Optane™ technology ushers in a revolutionary convergence of storage and memory. Intel® Omni-Path Architecture (Intel® OPA) provides an end-to-end fabric with a highly integrated and power-efficient system interconnect technology. Intel's work in silicon photonics offers further opportunities for affordable, high-bandwidth connectivity.

### Memory/Storage Hierarchy: Intel® Optane™ Technology

Intel Optane technology combines the revolutionary 3D XPoint™ NVM media developed by Intel and Micron

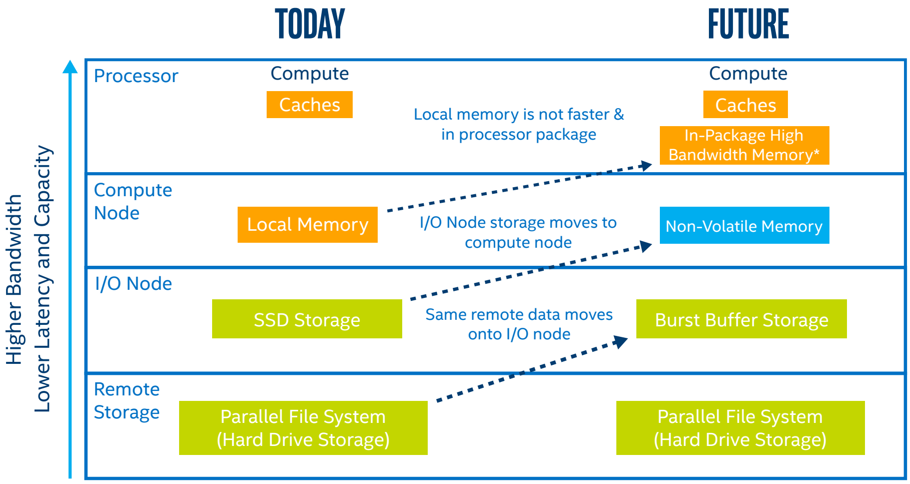

**Figure 3.** Configurable memory/storage hierarchy to support HPC and big data convergence

Technology with Intel's advanced system memory controller, interface hardware, and software. Intel Optane technology (Figure 3) offer dramatic speedups for many applications. Applications limited by data retrieval from large data sets may see orders-of-magnitude improvements in sustained performance when the entire working set of an application fits within the memory.

### Interconnect Fabric: Intel® Omni-Path Architecture

Intel OPA is a next-generation fabric designed to scale to systems with hundreds of thousands of nodes with tightly coupled processor, memory, and storage resources. Demonstrating a high degree of integration/convergence, Intel OPA is a key element of Intel SSF. It will be incorporated into the processor package and PCIe* adapters, integrated CPUs, switches, cables, host software, and management software.

Intel is facilitating industry-wide innovation by utilizing open-source standards such as OpenFabrics Alliance OpenFabrics* Enterprise Distribution (OFED). Intel is making key elements of Intel OPA available as open-source software and establishing the Intel® Fabric Builders community to promote the development of world-class implementations.

## Industry Enablement: System Software, Tools, and Algorithms

While new applications are emerging that are built from the ground up for parallelism, many existing applications, tools, and methodologies must evolve to reap the benefits of innovative architectures and technologies. Intel is facilitating this work through its own innovations and by building on its extensive work in industry enablement. Intel's activities and collaborations include:

- **Application development for analytics and machine learning.** Machine learning is a key enabler as researchers strive to draw more robust inferences from increasingly complex data. Intel has an active research group focused on deep learning and continues to create new tools for data analytics and machine learning.

- **Open-source collaborations.** Intel is driving convergence to a comprehensive, open-source HPC system software stack by joining with major hardware and software vendors and research organizations from around the world under the Linux* Foundation's OpenHPC community. Intel is also an active member and maintainer in the OpenFabrics Alliance, which provides industry-standard high-performance fabric software.

Intel is a leader of OFA's OpenFabrics Interfaces working group, ensuring that innovations in Intel Omni-Path Architecture and other high-performance fabrics are fully exposed to software. Intel also offers advanced open-source tools, platforms, and training, as well as collaborating with ecosystem partners on code optimization.

- **Collaborating for application modernization.** Intel® Parallel Computing Centers (Intel® PCCs) involve collaboration with leading universities, institutions, and labs to provide insight into requirements for next-generation computing and enable key user communities to modernize their codes for emerging architectures.

- **Evolving tools and programming environments.** Intel enables new technologies under existing standards when possible, to bring benefits to a broad range of existing applications. However, to deliver an optimal extreme-scale computing environment, applications, programming models, and resource management tools will in some cases need rethinking, redesign, or creation anew. Intel is actively engaged in these efforts, innovating and/or supporting work to evolve today's widely used Message Passing Interface (MPI) model; enable its performance compilers, libraries, and analysis tools; develop next-generation global and domain-specific resource managers; and identify capabilities needed so extreme-scale operating systems, runtimes, and system services can optimally support future HPC and analytics workloads. Intel's work to strengthen the popular Lustre* file system provides another building block for extreme-scale computing.

## Collaboration for Precision Medicine

Intel works with healthcare and life science leaders to advance meaningful use cases and accelerate development and modernization of applications and tools. This work represents an important step toward enabling life science researchers to take advantage of extreme-scale computing platforms and make the next leap in biomedical discovery.

- **Intel and Oregon Health & Science University (OHSU)** established the Collaborative Cancer Cloud, an open-source-based analytics platform for precision medicine, in 2015. The Cancer Cloud enables medical institutions to securely share insights from their private patient genomic data for potentially lifesaving discoveries. Using today's state-of-the-art technologies, the computing power of the Cancer Cloud speeds genomic sequencing and enhances clinician productivity, giving physicians more time to work with patients and devise effective, personalized treatments.

- **Harvard University and the Dana-Farber Cancer Institute** are applying single-molecule cryo-electron microscopy to the task of visualizing biological molecules in action at atomic resolution. Their work is relevant to the study of the structure of proteins relevant to cancer immunology and AIDS research.

- **Intel engineering teams** have worked with open-source developers since 2011 to produce major performance gains in widely used life sciences tools and applications. These activities help increase the adoption of modern codes in R&D, translate results into clinical insights, and shorten the time from genome sequencing to clinical impact.

- **Intel Corporation in Bangalore worked with researchers from the Indian Institute of Technology and others** to create a whole-genome network of the Arabidopsis thaliana plant (nearly 15,000 genes) from over 168.5 million gene expression values. They used an innovative approach to simplify the construction of genome-scale gene networks and ran their codes on the Tianhe-2 system at China's National University of Defense Technology.[22]

### On the Front Lines of Precision Medicine

Dr. Brian Druker pioneered research that led to the development of Gleevec*, a targeted therapy used to treat leukemia and a number of other cancers. In doing so, he catalyzed the emerging era of precision medicine.

Today, Dr. Druker heads the Oregon Health & Science University (OHSU) Knight Cancer Institute, which completed a USD 1B fundraising record to support the first large-scale program dedicated to early detection of lethal cancers. OHSU and the Knight Cancer Institute have collaborated with Intel since 2013 to advance the development of hardware, software, and workflow solutions for extreme-scale computing.

"The advances we're making in precision medicine allow us to state with some confidence that the end of cancer as we know it is within reach," says Dr. Druker. "But it's going to take breakthroughs on multiple fronts. We need much better screening tools, so we can fight cancer earlier, when it is an easier foe. We need to tie early diagnosis and targeted treatment together, and to distinguish slow and non-lethal cancers from more dangerous malignancies. We need new imaging technologies that let us see molecular abnormalities. And we need dramatic advances in computing to analyze the enormous volumes of genomic, imaging, and clinical data and help clinicians and researchers turn it into effective treatment plans that comprehend the unique characteristics of each patient's individual cancer at that moment in time. Extreme scale computing will help us build the foundation."

## Moving Forward

In addition to technology and scientific innovation, achieving the promises of precision medicine and extreme-scale computing will require creating new health and technology ecosystems and finding solutions to a wide range of policy and other issues. Governments and healthcare organizations have important roles to play, and citizens must be active collaborators in precision medicine planning.

### Crucial Roles for Governments

Governments have vital roles to play as funders, co-designers, eventual users, and policymakers in making precision medicine a reality. To help ensure precision medicine fulfills its promise, governments can:

- **Invest** in extreme-scale computing, computationally oriented life sciences research, and biomedical activities that utilize new and emerging digital capabilities.

- **Collaborate deeply** with technology innovators to co-design next-generation platforms that meet the needs of strategic user communities and use cases. Provide guidance to technology leaders on the types of simulation science and bioinformatics application developments that will have the greatest impact, as well as specific computational, memory, storage, and IO requirements for next-generation applications, workloads, and use cases. Develop strategies on an inter-agency basis while continuing to share results within public institutions.

- **Create open, interoperable, patient-centered environments** that promote rapid innovation and broad dissemination of advances. Promote open standards and flexible, modular designs that can be reconfigured for computing at all scales. Ensure that the benefits of innovations developed through public-private collaboration are shared, as much as possible, between the public and private sectors, to drive vibrant economies and strengthen national leadership positions.

- **Address policy and regulatory issues proactively**. Be prepared to seize the advantage of advances that precision medicine and extreme-scale technology will bring.  Put patient benefits at the center of planning efforts. Work with citizens, clinicians, and technologists to address issues such as ethical standards, data sharing, patient privacy, security, consent, and data interoperability.
- **Provide incentives** for healthcare organizations to incorporate precision medicine into clinical practice. Evaluate the value of whole genomic sequencing within the context of total cost of care for patients, where one test can replace multiple attempts to diagnose and treat disease.
- **Establish** skills programs to ensure the workforce can support and advance the emerging capabilities. Workforce development should target basic and applied scientists, computational scientists focused on healthcare and the life scientists, computer scientists, engineers, mathematicians, and HPC software/application developers.

## Capturing Immediate Value

Healthcare leaders should begin now to integrate genomic data and other new sources of data into electronic health records, clinical workflows, and decision-support tools. Whole genome sequencing and other elements of precision medicine may generate immediate savings by avoiding or reducing the hazards of today's trial-and-error approaches, which include unnecessary tests and drugs, incorrect diagnoses, and inappropriate treatments. Each of these improvements has the potential to reduce costs, improve system efficiency, and reduce human suffering.

Health IT leaders should modernize their technology infrastructure so they're prepared to make the most of clinical and biomedical data and handle rising compute and data requirements. Their voices are critical in setting industry standards and advocating for patient and clinician viewpoints in emerging paradigms.

## Benefits to Quality of Life, Economic Growth, and More

Health is a basic human right and a major influence on quality of life and social well-being. By advancing precision medicine, nations and states can position themselves to:

- **Enhance citizen satisfaction** and deliver affordable, high-quality healthcare to a fast-growing and rapidly aging global population, at lower cost and with better and more predictable outcomes.
- **Improve the care and treatment** of people with chronic and lifestyle diseases and perhaps slow the rising rates of such diseases.
- **Deal** with emerging health challenges—from antibiotic-resistant viruses to bioterrorism to previously rare diseases that will be more common (and likely more deadly) on a warmer planet.
- **Allocate resources** based on a deeper understanding of what approaches are most effective and cost-efficient.

Both health and extreme-scale computing are important elements of economic growth, national security, and competitiveness. By building healthier workforces, nations can drive growth and free funds to spend on innovation and health optimization rather than disease treatment. In addition to furthering precision medicine, investments in extreme-scale computing can enable advances in materials, manufacturing, climate modeling, biodiversity genomics, national defense, and other strategic issues. HPC investments generate benefits that ripple throughout a nation's economy. Intel is driving innovation to advance extreme-scale computing and precision medicine. Through R&D partnerships with Intel and deployment of Intel technologies, public and private-sector organizations can capitalize on Intel's efforts. Join us in making extreme-scale computing and precision medicine a practical reality.

Intel's integrated approach accelerates progress toward extreme-scale computing and helps democratize the resulting benefits. Intel® Scalable System Framework delivers high performance, large memory capacity, and high throughput for converged HPDA workloads.

## Learn More

Learn more about Intel solutions for high performance computing and big data in health and life sciences. Talk to your Intel representative, or visit us on the web:

- www.intel.com/hpc
- www.intel.com/healthcare/bigdata
- www.intel.com/healthcare/optimizecode

**Follow us on Twitter:**

- @IntelHPC, @IntelHealth, @portlandketan

The authors gratefully acknowledge the following for their contributions to this paper: Bill Bartolone, Alice Borrelli, James Brinker, Julie Donnelly, and Dennis Goo of Intel, Brian Druker of the OHSU Knight Cancer Institute, Jason Paragas of LLNL, and Eric Stahlberg of the Frederick National Laboratory for Cancer Research.

[1] HPC ROI: Invest a Dollar to Make $500 Plus Reports IDC, November 2015. http://www.hpcwire.com/2015/11/18/hpc-roi-invest-a-dollar-to-make-500-plus-reports-idc/. To access the full report and economic models, see IDC Economic Models Linking HPC and ROI, https://hpcuserforum.com/roi.html

[2] Comment by Harrison Moses, CEO of Alerion, in a Q&A session at the Informational Technology Innovation Foundation Senate Competitiveness Caucus Panel session, From 10th to 1st: Improving US Innovation and Global Competitiveness, February 10, 2016. https://itif.org/events/2016/02/10/10th-1st-improving-us-innovation-and-global-competitiveness

[3] World Health Organization. Noncommunicable Diseases Fact Sheet, January 2015. http://www.who.int/mediacentre/factsheets/fs355/en/

[4] American Cancer Society, Cancer Facts and Figures 2015. http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf

[5] World Health Organization. Noncommunicable Diseases Fact Sheet, January 2015. http://www.who.int/mediacentre/factsheets/fs355/en/

[6] World Health Organization, Nerurological Disorders: Public Health Challenges, http://www.who.int/mental_health/neurology/neurodiso/en/

[7] Martin Prince, et. al., World Alzheimer Report 2015: The Global Impact of Dementia, An Analysis of Prevalence, Incidence, Costs and Trends. Alzheimer's Disease International, London 2014. , downloadable from http://www.alz.co.uk/research/world-report-2015

[8] World Sepsis Day, Sepsis Facts, http://www.world-sepsis-day.org/CONTENTPIC/2015_WSD_FactSheet_long_English.pdf

[9] National Institute of General Medical Sciences, Sepsis Fact Sheet.

[10] World Sepsis Day, Sepsis Facts, http://www.world-sepsis-day.org/CONTENTPIC/2015_WSD_FactSheet_long_English.pdf

[11] World Health Organization, WHO's First Global Report on Antibiotic Resistance Reveals Serious, Worldwide Threat in Public Health, April 2014. http://www.who.int/mediacentre/news/releases/2014/amr-report/en/

[12] Centers for Disease Control, Antibiotic/Antimicrobial Resistance, November 2015. http://www.cdc.gov/drugresistance/

[13] David Cyranoski, China Embraces Precision Medicine on a Massive Scale, Nature, 06 January 2016, http://www.nature.com/news/china-embraces-precision-medicine-on-a-massive-scale-1.19108

[14] Genomics England, The 100,000 Genomes Project, http://www.genomicsengland.co.uk/the-100000-genomes-project/

[15] White House, Office of the Press Secretary, Investing in the National Cancer Moonshot, February 1, 2016. https://www.whitehouse.gov/the-press-office/2016/02/01/fact-sheet-investing-national-cancer-moonshot

[16] NIH, Precision Medicine Initiative Cohort Program. https://www.nih.gov/precision-medicine-initiative-cohort-program

[17] Next Gen Seek: How Does a Single HiSeq X Compare with HiSeq 2500? January 14, 2014. http://nextgenseek.com/2014/01/how-does-a-single-hiseq-x-compares-with-hiseq-2500/

[18] Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. (2015) Big Data: Astronomical or Genomical? PLoS Biol 13(7): e1002195. doi:10.1371/journal.pbio.1002195. http://www.plosbiology.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pbio.1002195&representation=PDF

[19] Using Wearable Technology to Advance Parkinson's Research, Intel and the Michael J Fox Foundation, 2015. http://www.intel.com/content/www/us/en/big-data/wearable-technology-parkinsons-research.html

[20] US Department of Energy Advanced Scientific Computing Advisory Committee Report, Top Ten Exascale Research Challenges, DOE ASCAC Subcommittee Report, February 10, 2014. http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/20140210/Top10reportFEB14.pdf

[21] Top500 List, November 2015. http://www.top500.org/lists/2015/11/

[22] Sanchit Misra Vasimmunin, MD, et al, Parallel Bayesian Network Structure Learning for Genome-Scale Gene Networks. IEEE Press, SC '14, Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 461-472. http://dl.acm.org/citation.cfm?id=2683644

Printed in USA 0416/KP/WCI/PDF ♻ Please Recycle 334145-001US