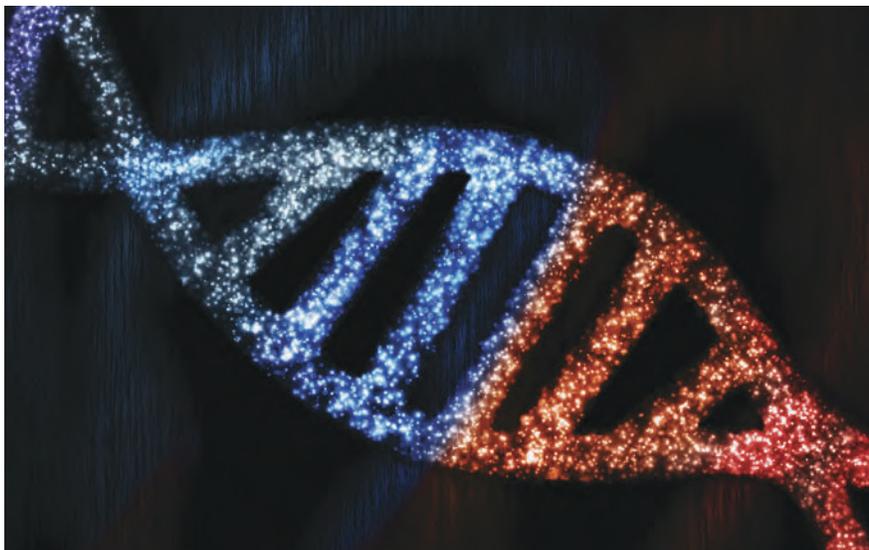# Lenovo™

A comprehensive portfolio of high-performance computing infrastructure designed to accelerate research, reduce IT complexity and maximize asset utilization

Highlights

- Deploy robust, high-performance computing (HPC) systems pre-integrated and optimized for life sciences

- Provide storage and file system expertise for acquisition, analysis and retention of large data sets

- Speed time to results with intelligent scheduling and workflow management

- Accelerate innovation using your choice of open-source or commercial tools

- Lower cost, easily scale and grow your HPC cluster

Life sciences research continues to accelerate at a rapid pace. New approaches to genomic analysis, such as next-generation sequencing, play a vital role in advancing scientific knowledge, facilitating the development of targeted drugs helping to deliver personalized healthcare. For example, new lab equipment, particularly next-generation sequencers, produces multiple terabytes of data per run that must be analyzed and compared to large genomic databases. Increasingly, both research and clinical applications involve rich data sets including MRIs, genome data and ultrasound imaging. Data from medical devices is increasingly being attached to patient records or clinical experiments, driving the need to store and process ever larger data sets with great efficiency.

The shortening of discovery pipelines, greater emphasis on collaboration and the need to conduct more research with fewer scientists means life sciences organizations must be able to accelerate time to results, while being able to manage, process and analyze higher volumes of data with greater efficiency.

To avoid obstacles and speed product innovation, researchers need computing environments that can process tremendous amounts of data rapidly. They also need to collaborate and share large data sets with upstream and downstream partners. And they need an infrastructure that supports automation to simplify data aggregation, assimilation and management.

Top-rated reliability and customer satisfaction

In a recent survey by Technology Business Research Inc. (TBR), customers rated hardware quality, ongoing reliability, performance, virtualization, efficiency and scalability as the most critical attributes to consider when purchasing x86 servers. Results from the same survey showed System x servers achieved the top ranking across all attributes—earning it a No. 1 rating in overall customer satisfaction. The ranking demonstrates the ability of System x servers to meet the evolving demands of today's enterprises.[1]

Lenovo and its network of partners offer a broad portfolio of open, industry-standard infrastructure solutions that can help accelerate product innovation without becoming locked into expensive, inflexible proprietary environments. Lenovo Solution for Life Sciences is comprised of pre-integrated, high-performance servers, storage systems and networking equipment, advanced file systems and integrated workflow management software (see Figure 1). Designed to accommodate large data volumes and compute-intensive applications, this solution can help organizations accelerate time to results at a lower total cost of ownership while making it easier to scale your HPC cluster as business grows.

Speed of analysis is critical in life sciences since it relates directly to the rate of discovery. As an example, information-based models, simulations, virtual molecule screening and other techniques can be used to identify promising new drug candidates quickly and kill off less-promising candidates to avoid incurring development and approval costs in later stages.

System X6 servers equipped with Intel processors provide the robust performance needed to accommodate demanding computational analysis workloads such as molecular dynamics and genomic sequencing. In addition to providing powerful processing, X6 servers offer the large memory capacity needed for in-memory analytics. Large memory machines are vital for applications such as genomic data
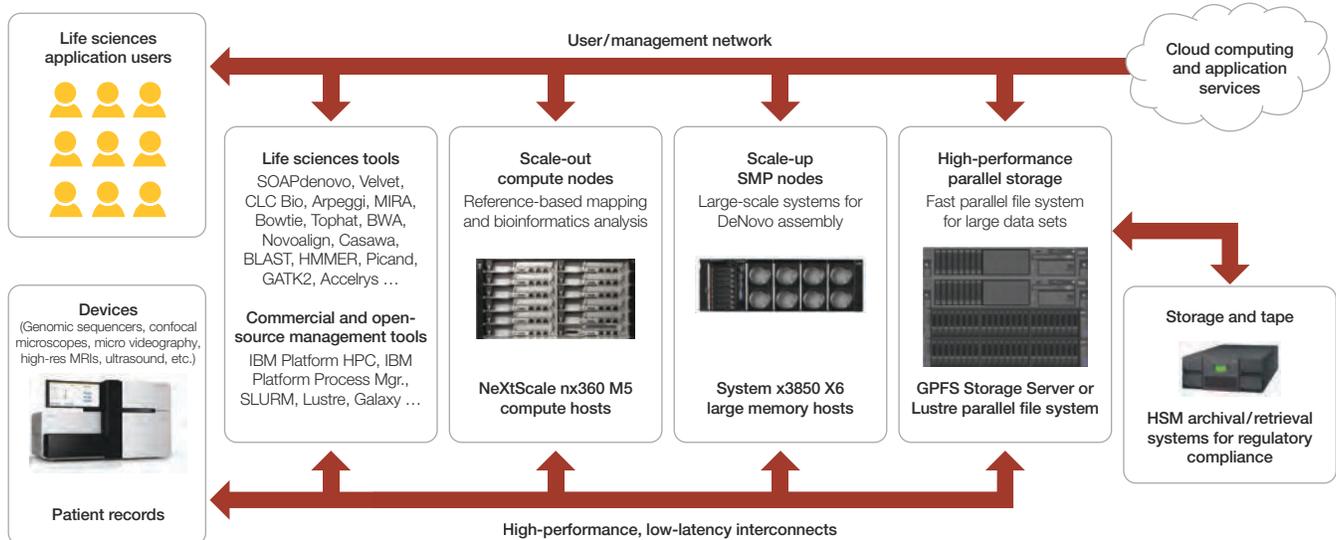
Figure 1. Lenovo Solution for Life Sciences delivers high-performance infrastructure for the acquisition, assembly, analysis and retention of life sciences data.

comparison, sequencing and analysis, and high-speed database searches. Organizations can equip the System x3850 X6 with up to 6 TB of memory, or scale to 12 TB of memory with the System x3950 X6.

The I/O throughput on a genomic system can be extremely demanding due to data volume and the large number of file and directory objects. Lenovo brings unmatched storage design and implementation skills that help tune networks, storage and access patterns to optimize data retrieval for analysis and processing. NeXtScale M5 systems are designed to meet these demands with a simple, flexible and open architecture. Based on a dense chassis design with the ability to pack up to 84 high-performance servers in a 42-inch rack, NeXtScale offers high density and flexibility for large computational analyses. It uses industry-standard components, including I/O cards and top-of-rack networking switches, for flexibility, lower costs and ease of adoption.

Lenovo also offers an extensive software stack to run on top of NeXtScale, providing a choice of powerful, fully supported scheduling, management and optimization tools from IBM Platform Computing or compatibility with a wide-variety of other commercial or open-source tools. The result is a single architecture based on open standards that delivers high performance and high efficiency, allowing organizations to maximize their compute power in a minimum amount of space. Built on industry-standard technology, System x servers can support a full range of life sciences platforms and applications, such as Dassault Systems BIOVIA, CLC bio, Lab 7, Schrödinger and Intel-optimized genomics code.

— **Dr. David Spetzler,**
vice president of research and development, Caris Life Sciences

A greater focus on computational analysis during early-stage research and development is driving up the volumes of data that life sciences organizations must store and manage. These large volumes and the collaborative nature of life sciences research are placing new demands on storage solution performance and data manageability.

Lenovo and Lenovo Business Partners offer a full range of storage technologies designed to accommodate rapidly rising data volumes while maximizing efficiency. The GPFS Storage Server (GSS), based on IBM General Parallel File System (GPFS), System x3650 servers, JBOD storage and networking technologies, offers a high-performance, scalable building-block approach to large-volume storage needs. The GSS allows organizations to start small and build through incremental additions, providing expanded capacity and bandwidth with each additional GSS building-block.

The GSS is built, tested, delivered and supported as an integrated Intelligent Cluster. It comes in a variety of sizes and options including configurations with solid-state disks (SSDs) for added performance. The GSS can be deployed with multiple disk storage tiers, mapping data based on a policy engine to a selected storage tier that incorporates both disk and tape for long-term archival and regulatory requirements. Solution-level support includes components to deliver maximum system availability throughout the life of the system. Research teams can spend less time maintaining systems and more time delivering faster, higher-quality results.

In addition to fast data access, life sciences organizations need the flexibility to store data for longer times on appropriate cost/performance devices, while offering data management tools to migrate and protect that data. Lenovo offers a robust portfolio of storage solutions that can handle the volumes of data generated by complex genomic applications and provide the performance and scalability needed to gain insights from the latest analytics technologies. Organizations can select solutions with a mixture of flash drives and hard disk drives to balance performance and costs.

The increasingly distributed nature of genomics infrastructure requires data management on a much larger and global scale. Not only must data be moved or shared across different sites, its movement or sharing also must be coordinated with computational workload and workflow. Lenovo understands the data management needs of life sciences environments and can provide complete clusters with fully operational process management and scheduling tools.

Lenovo's partnership with IBM allows it to deliver complete solutions to support a wide variety of applications. IBM Platform HPC software offers complete workload management capabilities in a single, low-cost product. Organizations can run a wide range of workloads across a single, shared infrastructure. For firms engaged in collaborative research across centers, Lenovo offers IBM Platform Process Manager (PPM) to help automate complex life-sciences workflows, and provide tools to enable run-time decisions about least-cost approaches to handling large data sets in WAN environments. IBM GPFS Automated File Migration (AFM) can provide a single global name space and efficient access for remote data boosting productivity and avoiding the need for redundant copies of data to help reduce total storage costs.

Lenovo systems are also compatible with a wide variety of open-source codes such as Galaxy workflow, SLURM scheduler and Lustre file system for data acquisition, analysis and retention. Lenovo is also working closely with Intel to help ensure Intel-optimized genomics and other life sciences-related codes run optimally on Lenovo clusters.

As transaction volumes rise, existing compute and storage cluster interconnects may have trouble keeping up. Lenovo low-latency server adapters and switches deliver the performance needed to support clustered databases, parallel processing and high-performance embedded I/O applications, reducing task completion time and lowering cost per operation. Optimized to support Lenovo server and storage solutions, Lenovo and Lenovo partner-provided top-of-rack networking switches and adapters offer 10 Gb and 40 Gb Ethernet connectivity as well as EDR InfiniBand.

The Lenovo Solution for Life Sciences can be deployed as individual building-block components or as a complete Intelligent Cluster solution with a single part number. Intelligent Cluster systems are built, tested, delivered and installed by Lenovo or Lenovo qualified Business Partners, and are supported as a single solution instead of being treated as hundreds of individual components. Lenovo provides single point-of-contact and solution-level support that includes both Lenovo and third-party components to deliver maximum system availability throughout the life of the system.

The integration of new technologies and computing approaches in life sciences is essential to accelerating the rate of discovery and achieving breakthroughs that can improve lives. Lenovo Solution for Life Sciences give organizations powerful, open infrastructure designed to speed research, optimize resource utilization and deliver significant savings in time and investment.

Lenovo is a global personal and enterprise technology company—the largest PC and systems company in the world—serving customers in more than 160 countries. Dedicated to building exceptionally engineered PCs, mobile Internet devices and servers spanning entry through supercomputers, Lenovo has built its business on product innovation, a highly efficient global supply chain and strong strategic execution. The company develops, manufactures and markets reliable, high-quality, secure and easy-to-use technology products and services. Lenovo recently acquired IBM x86 server business. With this acquisition, Lenovo is adding a best-in-class x86 server portfolio along with HPC and life sciences expertise.

To learn more about **Lenovo Solution for Life Sciences**, please contact your Lenovo sales representative or Lenovo Business Partner, or visit:
www.lenovo.com/systems

Lenovo™