Microsoft®

IBM®

Personalizing Healthcare through High Performance Computing at Medical Centers

IBM® - Microsoft® High Performance Computing Solutions for Healthcare and Medical Centers

High Performance Computing for Healthcare and Medical Centers

High performance computing (HPC) continues to power breakthroughs in genomics, biochemistry, and bioinformatics at a rapid pace. More importantly, the integration of these scientific breakthroughs with imaging and data management enables remarkable advances in healthcare, drug discovery, and genomic research. Together, they create exciting personal therapeutic strategies in healthcare, medical centers, and hospitals for longer, healthier lifestyles.

In the next few years, doctors are likely to have access to the entire genome information of individual patients, enabling better informed, customized treatment decisions using drugs genetically tailored for specific genotypes. Rapid advances in biomedical imaging allow clinicians to better visualize and diagnose the onset of disease. Advances in biomarkers and phramacogenomics help doctors track, monitor, and treat diseased conditions with far greater precision than before. These advances would not have been possible without high performance computing solutions that today give doctors the ability to routinely demand and receive a level of detailed information about their patients, unthinkable even two years ago.

HPC powers next generation medical center solutions

Medical researchers now have at their disposal extremely powerful tools for profiling and analyzing molecular, proteomic, and genomic interactions at the atomic level. Tied tightly with the IT industry, these advances closely track advances in CPU processing capabilities and parallel data processing methodologies.

Proteomic biochemical analysis has progressed from single proteins to complex protein mixtures. Gene expression analysis has evolved from microarray based methods to real time PCR quality assay based analysis. The throughput and output of today's microarrays and spectrometers have challenged IT managers with an unprecedented level of computing performance and workflow automation needs and I/O delivery complexity.

IBM® and Microsoft® are at the forefront of the HPC revolution to address and resolve these challenges. Both have demonstrated capabilities of powering innovation and productivity in life sciences and medical research. Microsoft and IBM now combine to offer the power and flexibility of the Windows® HPC Server 2008 software on IBM System x® cluster solutions to help speed time-to-insight and discovery.

NEXT GENERATION DNA SEQUENCING & ANALYSIS

Next generation sequencing helps optimize preventive healthcare and drug therapies while people are still well or at the earliest stages of a disease. Genetic variations, in combination with environmental and lifestyle factors, lead to diseased states involving the interplay of numerous variations and pathways. Investigating these conditions and making statistical genotype-phenotype associations involves sequencing the DNA of diseased populations and comparing them with the Human Genome Project's reference sequence.

These new technologies are bringing healthcare researchers ever closer to personalized patient diagnosis and care. The cost of genomes sequencing has already touched the $10,000 level and the goal of $1,000 is within sight. This change has put high-throughput genome sequencing, once achievable by only a few major sequencing centers, within reach of many smaller research groups and individual research laboratories.

Genome sequencing centers are among the most visible users of high-performance computing. A single next generation sequencing run can generate terabytes of data. Sequencing centers and spectrometry laboratories are now equipped with system farms that collectively create data at many gigabytes per second. Once the data is ingested into the storage environment, it needs to be processed and correlated via a cluster supercomputer using popular sequence analysis tools and software applications.

These workflows have put new demands on file systems that have not previously been seen in bioinformatics as processing capabilities scale from workgroup cluster sizes to much larger facilities that have thousands of

HPC cluster nodes for sequence analysis. Further, consumables such as reagents are a significant factor in the cost per sequencing run, and this makes it even more important for the HPC configuration to work robustly and avoid needless repetition.

IBM and Microsoft partner with leading players in the next generation sequencing ecosystem including Illumina and Applied Biosystems, the leading providers of sequencing equipment. A typical configuration involves pulling high-throughput data in the order of 20-30 terabytes from next generation sequencers and processing it through applications such as BLAST, HMMER, ClustalW, and FASTA using BladeCenter®, iDataPlex™, and the General Parallel File System™ (GPFS™) servers and file systems running on Windows HPC Server 2008.

WELLCOME TRUST SANGER INSTITUTE

The Wellcome Trust Sanger Institute in the UK is one of the leading genomics centers in the world for analyzing and understanding genomes. Through large-scale analysis and focused research and collaborations, the Sanger Institute's programs underpin biological and medical research worldwide. IBM System x clusters and blades, and Microsoft Windows solutions are an integral part of the workflow in next generation sequencing at Sanger.

DAGGETT RESEARCH GROUP

The Daggett Research Group in the University of Washington Department of Bioengineering is world-renowned for the study of protein stability, function, and folding—some of the fundamental unsolved problems in molecular biology. Given that experimental approaches provide only limited amounts of information, the Daggett Research Group uses computer simulations that require massive amounts of computation. The group augmented its existing high-performance computing (HPC) resources with two new HPC systems based on Windows cluster solutions. This has helped enable the group to make fundamental breakthroughs in how the results of simulations are analyzed.

BIOMEDICAL IMAGING

Image analysis is critical in healthcare. It is fundamental to every stage of drug development from the measurement of cell assays to examination of tissue samples to interpretation of medical imaging. Equally, image analysis is critical for early diagnosis and delivery of effective, personalized treatment at the point of care.

Today, bio-imaging technologies are not only a valuable tool for translational research but have also become an integral part of defining how drugs and medical devices are developed and targeted. Imaging pervades the entire drug discovery and development pipeline starting from compound screening and lead optimization, and reaching all the way to preclinical and discovery-stage animal studies.

In clinical research, bio-imaging has become all but omnipresent, providing an enormous amount of patient-specific information that, if linked to clinical and behavioral parameters, can often aid in a proof-of-concept understanding of investigational drugs.

IBM Research collaborates with leading research institutes to use HPC to push back the boundaries in neuroscience research and medical imaging. Current projects include research initiatives with Ecole Polytechnique, Stockholm Brain Institute, Mount Sinai Medical Center, Rockefeller University, and others, in areas ranging from HRRT PET reconstructions to MRI artifact reduction and fMRI analysis to modeling of neural tissue development, cortical map formation, and emotional systems.

The major HPC applications in medical imaging are high content analysis, storage, and retrieval of image data for both molecular imaging and diagnostic imaging. Image data is often unstructured and voluminous, and memory and storage performance is critical.

IBM System x with Windows HPC Server 2008 offers efficient and high performance solutions for medical imaging. IBM BladeCenter servers

integrated with IBM General Parallel File System (GPFS) and IBM DCS9900 provides excellent data storage and retrieval solutions. The IBM Grid Medical Archive Solution (GMAS) leverages IBM System Storage Multilevel Grid Access Manager software delivering flexible and scalable virtualized storage solution for mid to large healthcare organizations with single or multiple sites.

NATIONWIDE CHILDREN'S HOSPITAL

The Biomedical Imaging Team, part of the Center for Childhood Cancer and the Biopathology Center at the Research Institute at Nationwide Children's Hospital at Columbus, Ohio, is currently working with the Microsoft HPC team and the Ohio Supercomputer Center to utilize high performance computing to catalogue and analyze large data sets.

THY-MORS HOSPITAL, DENMARK

An IBM technology innovation for intelligent electronic patient records using a 3D model of the human body has proved practical and valuable at the Thy-Mors Hospital, Denmark. Using an "avatar" or map of the human body, doctors are now able to obtain an up-to-date overview of each patient's record at a glance by easily navigating an electronic patient file. They can choose between different views, for example enabling inspection of the organs or the circulatory, muscular and nervous systems.

"I can see much more information than just what the patient tells me is bothering him or her that day—information for which I would otherwise have to spend considerable time searching through our records system."
Dr Hardy Christoffersen, Head of Surgical Outpatient Clinic, Thys-Mors Hospital.

MAYO CLINIC

The Medical Imaging Informatics Innovation Center (MI3C), a collaborative venture between IBM and Mayo Clinic, brings together researchers, clinicians, and technology experts on Mayo's Rochester, Minnesota campus, to showcase leadership in research, medicine and technology; apply computationally intensive solutions to real-life diagnostic problems; and enable collaborative innovation to drive better patient information and treatment.

Results are already showing up in new solutions for the kinds of diagnostic problems clinicians see every day but can't begin to resolve on their own.

"This facility allows us to explore projects in medical imaging and radiology that can help provide faster and better information for our physicians, and in turn, improved treatments for our patients."
Bradley Erickson, M.D., Ph.D., Head of Mayo's Radiology Informatics Lab and MI3C co-director.

PHARMACOGENOMICS

Pharmacogenomics is one of the most disruptive new approaches to drug discovery and development. In the past, pharmaceutical companies relied on so called blockbuster drugs that could be marketed to a broad population. But this one-size-fits-all approach is increasingly inefficient—it is estimated that only 30 to 50 per cent patients benefit from some drugs due to genetic variations.

Under the pharmacogenomics paradigm, populations are subdivided into several genotypic groups to study the influence of genetic variation on drug response in patients by correlating gene expression or single nucleotide polymorphisms with a drug's efficacy or toxicity. This helps optimize drug therapy with respect to the patients' genotype to ensure maximum efficacy with minimal adverse effects.

From an IT perspective, data collection in support of pharmacogenomics investigations will certainly be more complicated with more tests, more studies, and more intricate data to track and analyze.

PFIZER

Pfizer works with IBM for a pharmacogenomics data warehousing solution. The solution combines clinical patient data with genomic patient data and allows the user to look at the two together as a basis for targeted treatment, and for selecting patients for clinical trials. By knowing which patient has a certain expressed gene, a drug developer will know which patients can safely test a drug for efficacy. Likewise, the physician treating the patient will apply a diagnostic test before prescribing drugs that will be effective for treatment.

THE SCRIPPS RESEARCH INSTITUTE

The Scripps Research Institute, based in La Jolla, California, and Jupiter, Florida, is a leading center for basic biomedical science. Structural biology researchers at the institute used to process test results manually, inhibiting the pace of their work. To boost productivity, they adopted a custom high-performance computing (HPC) solution based on Windows HPC Server 2008. Now, test results are processed 600% faster, giving researchers the time they need to analyze their work more thoroughly and to identify more productive research directions. The HPC applications for the solution were created quickly and easily using Microsoft development tools, and the entire solution life cycle is managed centrally and effectively by a single system developer.

"This is high-performance computing that doesn't require any specialized tools or training. If you know Windows, you know Windows HPC Server."
Josh Kunken, System Developer, The Scripps Research Institute

GENOME RESEARCH INSTITUTE

The University of Cincinnati Genome Research Institute (GRI) is a global hub for biomedical research focusing on new therapies for cancer and metabolic diseases such as obesity and diabetes.

Founded in 2002, GRI brings together a range of academic, nonprofit, and industry partners. The institute uses Microsoft HPC cluster solutions to make it easy for all researchers to use HPC without relying on super users to run jobs for them. Users can now easily submit and monitor jobs from their workstations.

BIOMARKERS

Biomarkers are indicators of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention. Biomarker-based drug development enables better and earlier decision making and genomic biomarkers are beginning to pave the way toward targeted therapeutics. Leading companies such as Novartis, Eli Lilly, Bristol-Myers Squibb, Pfizer, and Bayer have developed new R&D processes based on the use of biomarkers throughout the clinical trials process for cardiology, oncology and neurology.

In the imaging field, biomarkers are increasingly being used non-invasively to assess patients' localized disease progression and response to drug candidates. Biomarkers can help determine whether the drug reaches the target, affects biological activity, and lead to the desired clinical outcome.

However, the science in biomarkers is outpacing widespread acceptance. The path toward acceptance by regulators and the medical community is through discovery and consistent validation of genomic, proteomic, in vitro, and imaging biomarkers.

Since 2005, IBM has hosted The Imaging Biomarker Summits to provide a forum for the exchange of ideas and best practices related to biomarker enabled pharmaceutical R&D, with special emphasis on the impact of imaging technologies. They are attended by representatives from the bio-pharmaceutical

industry, government and academic medical research centers, FDA, medical device manufacturers, and partners.

THE GENOGRAPHIC PROJECT

The Genographic Project, launched in April 2005, is a five-year genetic anthropology study to map historical human migration patterns. It is a joint initiative by National Geographic and IBM, with field research by the Waitt Family Foundation and laboratory research supported by Applied Biosystems.

The project samples DNA from populations around the world and looks for genetic markers—a mutation or modification of the genome as it is passed through generations. Different populations have different genetic markers, and by following them through the generations, scientists are able to identify the different branches of the human tree all the way back to their common African root. Indigenous populations provide geographical and cultural context to the genetic markers in their DNA. These clues can help recreate past migration patterns.

KAROLINSKA INSTITUTET

IBM and Karolinska Institutet have been working together since 2004 to build Sweden's first IT-enabled biobank to advance the understanding of the links between genetics, environment, and disease.

The project integrates research projects across the country by developing data collection standards and providing the IT infrastructure to foster collaboration that could lead to the development of more targeted, timely cures for complex diseases.

IBM high performance cluster solutions with the Windows HPC Server 2008

With the introduction of Windows HPC Server 2008, customers now have access to the latest Microsoft high performance computing platform with IBM System x and BladeCenter servers.

These affordable yet powerful servers offer the enterprise power and control needed to accelerate time-to-insight with a high performance cluster solution that is simple to deploy, operate, and use with existing infrastructure and tools.

Windows HPC Server 2008

Windows HPC Server 2008 is based on the Windows Server® 2008 operating system and is designed to increase productivity, scalability, and manageability. It includes key features, such as high-speed networking, highly efficient and scalable cluster management tools, advanced failover capabilities, a service-oriented architecture (SOA) job scheduler, and support for partners' clustered file systems.

Windows HPC Server 2008 combines the power of a Windows 64-bit Server platform with rich, out-of-the-box functionality to improve the productivity and reduce the complexity, of the HPC environment. It provides a comprehensive set of deployment, administration, and monitoring tools that are easy to deploy, manage, and use with existing infrastructure.

Windows HPC Server 2008 enables broader adoption of HPC by providing a rich and integrated end-user experience scaling from the desktop application to the clusters. A wide range of software vendors, in various verticals, have designed their applications to work seamlessly with Windows HPC Server 2008 so that users can submit and monitor jobs from within familiar applications without having to learn new or complex user interfaces.

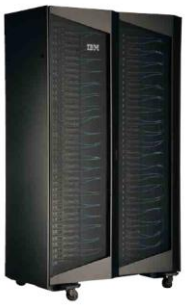
IBM System x® Clusters

IBM System x cluster is a portfolio of robust cluster servers built on IBM X-Architecture® that enable you to jumpstart the production of applications or products faster than the competition. This is realized through research,

experience, innovation, and proven support from IBM.

IBM System x clusters are enhanced by intelligent management tools, leadership offerings in energy-efficiency, consolidation and virtualization, security, and remote maintenance. This gives your data center the much needed optimization and cost reduction boost that is critical for businesses today. Multiple layers of redundancy, memory protection, and high availability tools are integrated into System x design. The IBM System x offering consists of iDataPlex internet-scale computing, BladeCenter servers, Cluster 1350 servers, and Storage servers.

System x iDataPlex®



The System x iDataPlex internet-scale computing server solution from IBM is uniquely positioned to help enterprise clients overcome compute density and energy efficiency constraints. The iDataPlex supports massive scale-out data centers and high performance computing solutions.

The iDataPlex is a half-depth server solution, optimized both mechanically and component-wise for maximum power and cooling efficiency. It is an industry-standards based server platform designed to minimize utilization of data center floor space, power and cooling infrastructure. An easily maintainable solution with individually serviceable servers, front access hard drives/cabling, and common tools across the System x portfolio for management at the node, rack, or data center level, iDataPlex is configurable for customer-specific compute, storage, or I/O needs and delivered pre-configured for rapid deployment.

BladeCenter®

By integrating servers, storage and networking, IBM BladeCenter in healthcare and medical centers sweep aside complexity. Its wide application solution blades, truly efficient chassis, and open design, are packed into an answer to today's straggling racks and overheated server rooms.

Cluster 1350™ Server

IBM Cluster 1350 Server reduces deployment time for Windows clusters from Microsoft, benefits power and cooling through IBM Systems innovation, and offers integrated global hardware support. Healthcare and Medical Center clients can speed up installation of an HPC cluster, simplify its management and support, and reduce mean time to payback.

Storage servers

IBM storage servers come at versatile ranges of entry level disk storage servers, mid-range disk storage systems—the IBM TotalStorage® Enterprise Storage Server® (ESS) and the IBM TotalStorage DS4000® series, the Enterprise disk storage systems—DS6000™ and DS8000® series, Network attached storage or N series products, and IBM's TotalStorage SAN solutions.

The IBM – Microsoft Value Proposition

INNOVATION

Microsoft technologies connect people and systems based on industry standards and help provide a robust platform for adopting future innovations. Microsoft has partnered with industry-leading solution providers to develop tools that help healthcare and medical centers collect, manage, and share the right information with the right people. Microsoft addresses the challenges posed by data formats as standards emerge to deliver information in an accessible way. These solutions are intuitive, building on familiar Microsoft applications to enhance collaboration and help secure your IT environment.

In 1991, Microsoft created Microsoft Research to build a foundation for future technology breakthroughs by focusing on long-term projects independent of day-to-day product development.

Today, Microsoft Research employs more than 700 researchers working in more than 55 research areas. Since 2003, scientists at Microsoft Research have been working on the design of a vaccine for the HIV virus. High Performance Computing Windows cluster solutions help these researchers significantly

increase the range and efficiency of their work—today, researchers run 50 jobs of 200,000 work items each in the time it once took to run one job. The result? Deeper insights and more published research in top scientific journals.

The IBM Computational Biology Center (CBC) promotes research across a wide range of problems including pattern recognition in sequences, complex molecular interactions, biological data analysis, interpretation, and reverse-engineering of complex disease-lifestyle genomic interactions for deeper biological understanding.

CBC tools have been adapted and implemented by a wide range of customers and collaborators at medical research institutions worldwide for biomedical and clinical applications. The algorithms developed at CBC have also been used to analyze the content of patient records enabling information and genomics based personalized medicine and therapies.

ENHANCING PERFORMANCE

Because IBM cluster systems can be deployed to both scale up and scale out, they provide the ability to enable higher levels of performance with minimal disruption to existing environments.

The IBM System x rack-optimized cluster servers have exceeded the most optimistic projections and have become the platform of choice for server consolidation with virtualization, running large scale life sciences workflows.

The IBM BladeCenter enables systems that are even more compact than rack-optimized clusters. In 2004, IBM opened up the specifications to BladeCenter, and this has enabled entrepreneurial innovation to drive the development of blade-based computing.

IBM General Parallel File System (GPFS) can accelerate long-running, parallel calculations such as sequencing analysis, imaging analysis, and biomarker analysis on very large data sets. GPFS achieves higher levels of performance by making it possible to read and write data in parallel and distributed across multiple disks or servers. Because multiple disks or servers contain the data, it also achieves a higher level of fault resilience for clusters and grids across multiple locations.

Microsoft technology makes it possible to deliver solutions that make it easy to help global teams find, aggregate, analyze, and share the right information precisely when they need it. These solutions can seamlessly transfer information from disparate formats to one that is familiar and easy-to-understand. Freed from data-crunching chores, medical researchers can focus their full attention on finding the next breakthrough drug.

REDUCING COST AND COMPLEXITY

Microsoft System Center solutions help IT professionals manage the physical and virtual information technology environments across data centers, desktops, and devices. Using these integrated and automated management solutions, IT organizations can be more productive service providers. With a strong partner ecosystem, a compelling licensing model, and attractive economics, System Center solutions are designed to provide a low total cost of ownership and high return on investment. The functionality and economics of System Center solutions combine to provide the fastest time-to-value while building a more agile IT infrastructure.

IBM data center strategy addresses the design and management of arguably the biggest, most technical and complex challenges. And what can be more complex than high-performance computing? IBM data center services specialize in the design and management of large, complex HPC installations, and IBM Site and Facilities Services can comfortably address the extreme computing needs of life sciences organizations.

Additionally, IBM Computing on Demand (CoD) provides companies with flexible access to security-rich supercomputing environments that can be used just like onsite hardware but without the capital commitment, management, and maintenance costs. The hardware is hosted, maintained, and supported by IBM to deliver cost effective capacity that frees up companies to focus on core activities. IBM customers have discovered that apart from a reduction of 30%-60% in total cost of ownership, they can also gain from CoD through new capabilities to solve previously

intractable problems, faster time-to-results, and the ability to implement innovative new business models.

IMPROVED PRODUCTIVITY

Microsoft Knowledge Driven Life Sciences help information workers collaborate easily and make informed decisions quickly. One of the top issues facing life sciences researchers and professionals is transforming the flood of information into usable medical and therapeutic knowledge.

Most organizations store information in a wide range of isolated systems. As a result, researchers can't always find the information they need when they need it. The latest generation of collaboration products and tools from Microsoft help life sciences organizations tackle this problem.

The Microsoft Office family of interoperable technologies and applications help streamline processes, enhances productivity, and facilitates collaboration.

Microsoft Office InfoPath enables researchers to create interactive forms that make it easy to capture and share data, thus eliminating the need for researchers to reenter the same information in different applications.

Microsoft SharePoint Portal Server provides life sciences organizations with the tools to quickly develop portals that integrate data from various systems into a single solution making it dramatically easier to access and share critical information.

IBM HPC cluster solutions work in a software environment that is designed to take the full advantage of IBM hardware features without compromising reliability and security. The programming and administration environment is based on familiar programming languages, libraries, job management tools, and parallel file systems.

Medical center application developers greatly benefit from these innovative software components without facing a steep learning curve. High Performance Computing cluster software available on IBM clusters includes middleware, the General Parallel File System (GPFS), and a job scheduler designed to maximize resource utilization and job throughput.

IBM clusters also support workload management solutions available from partners. This combination of middleware enables the optimization and scaling of the IBM cluster resources (processors and storage) for several concurrent computations and simulations typical in medical center and life sciences workflows.

FINANCING

IBM Global Financing helps clients reduce cost and optimize their purchase strategies by providing a 'wrapper' of funding solutions, asset disposition strategies, and refurbished equipment. From initial consultation to upgrades to disposition, tailoring simple leases and loans, to highly customizable structures, IBM Global Financing helps organizations and industries with cost effective HPC solutions. By providing holistic financing for the capital intensive data center projects, IBM alleviates capital constraints and offers the opportunity to align up-front costs to anticipated project benefits.

Microsoft Financing provides an affordable and predictable way to acquire the IT solution that companies need to help cut costs and increase efficiencies. Qualified customers are offered immediate access to funds to purchase and deploy the IT solution they need. There is no down payment required, and terms range from two to five years—making it easy to match your payments to realized IT benefits immediately.

Microsoft Financing serves businesses of all sizes and finances a complete range of technology needs, including software, services, partner products, and hardware. Further, Microsoft solution partners can help customers define their technology-related business needs and objectives, set up a payment structure that can work with your budget and help maintain cash flow.

IBM Corporation
Integrated Marketing Communications
Systems and Technology Group
Route 100, Somers, NY 10589

Produced in the United States