

THE BEAST THAT ROARS FOR WORLD-CLASS SCIENTIFIC RESEARCH

A NEXSAN CASE STUDY OF WELLCOME TRUST

The Wellcome Trust Centre for Human Genetics has opened up new pathways toward greater understanding on human genome. The complexity of its multi-disciplinary research has presented a multitude of challenges in data storage and availability. The underlying concern is always resilience to failure. Nexsan offered the industry-leading storage solution to the Centre and brought to it the peace of mind and a world of possibilities.

THE BEAST UNLEASHED

SATABeast from Nexsan is an extreme density, energy efficient storage array designed to meet the challenges of the most demanding storage environments. Featuring both Fibre Channel and iSCSI connectivity, SATABeast provides unmatched flexibility along with enterprise-class features, performance and proven reliability.

- Employs Nexsan's revolutionary AutoMAID® [Automatic Massive Array of Idle Disks] energy-saving technology, reducing power consumption and operational costs.
- Supports standard high performance RAID implementations as well as back up, VTL [Virtual Tape Library] and archiving applications.
- Utilizes dual RAID engines on each SATABeast controller to provide extremely high throughput across 42 spindles.
- Offers high availability features including continuous background integrity scans to ensure drive health and data accessibility.
- Each components is fully redundant and hot swappable with each controller featuring full cache coherency to maintain data integrity ≠ even in the event of a power loss.



CASESTUDY 2

BACKGROUND

Scientific research is now opening up new pathways towards a greater understanding of the genetic basis of many common ailments that have plagued and perplexed humanity for centuries.

Nowhere are these endeavours more in evidence than at the Wellcome Trust Centre for Human Genetics, located at the University of Oxford. Established in 1994, the Centre undertakes vital research into the genetic basis of common diseases, including asthma, diabetes, malaria and cardiovascular disease. A not-for-profit organization, it places all of its findings in the public domain – a tremendous asset to global biomedical research.

The centre houses multi-disciplinary research teams in human genetics, functional genomics, bioinformatics, statistical genetics and structural biology, all of which presents a multitude of challenges when it comes to data storage and availability. With so much research underway, data volumes have expanded at a staggering rate. What has fuelled this frenzy is the explosion in bioinformatics and statistical genetics work involved in genomic research, says Dr. Tim Bardsley, the Centre's IT manager, "particularly since the Sanger Institute mapped the entire human genome."

the Sanger Institute, under founder director Sir. John Sulston, spearheaded the UK contribution to the human Genome Project. It sequenced almost one-third of the human genome, greatly enhancing our ability to study the diseases that afflict people and animals, and was instrumental in ensuring that sequence data were made freely available to researchers worldwide for the benefit of all. As well as the human genome, the Sanger Institute has sequenced the genomes of numerous disease-causing microbes, including those that cause tuberculosis, malaria, leprosy and diphtheria (See panel 'The Sanger Institute: A New Focus').

THE CHALLENGE

"As a result, scientific, research-focused organisations like ours are faced with extremely complex storage management tasks," states Bardsley. "Our expanding research programmes and increasing data are forcing an exponential growth in our data storage requirements, a phenomenon seen right across the scientific research industry."

All this is a far cry from how things were when he joined the centre eight years ago. "Then bioinformatics had a relatively small role, with far fewer stats being generated. Even six years ago, our SAN [Storage Area Network] was the size of a modern hard disc. To meet our expanding requirements, this was first upgraded to eight terabytes and then, through further recent investment, to fifty terabytes."

//

The underlying concern for us always with any equipment is resilience to failure... Though we have experienced failures with other manufacturers, Nexsan's technology has never failed."

DR. TIM BARDSLEY

IT MANAGER, WELLCOME TRUST CENTRE FOR HUMAN GENETICS

A NEW FOCUS

A key aspect of the Sanger Institute's work is to identify sequence variation in human populations, and how specific variants contribute to health and disease. One major initiative, the Cancer Genome Project, is systematically searching all human genes for genetic variations implicated in cancer – a quest that has already unearthed new cancer genes and potential new therapies.

Under current director Allan Bradley, focus has shifted to the use of genome sequence data and the development of high-throughput methods to answer important biological questions and the role of genes in health and disease. The breadth of the institute's work has expanded, and in the past two years nearly 20 new senior project leaders have been recruited. New laboratories are being built to house these new studies.

The Wellcome Trust Sanger Institute is named in honour of Fred Sanger, a double Nobel Laureate who devised the method for DNA sequencing used in the Human Genome Project.

With a total of around 500 network-connected permanent staff to cater for, including administrators and scientists (primarily located within the main building itself, but many also collaborating with associates worldwide), IT has to provide the bandwidth – and data security – that helps to sustain the centre's operations. That means ensuring the vast amounts of data generated, and then relayed to bio-information for detailed analysis, can travel at speed, unhampered, through the network.

"The bottom line is that, when genetic material is being analyzed, large output files are created," states Bardsley, "each of about fifty megabytes in size – and all of these have to be kept online in order to facilitate research." As a result of this activity, the centre's data yield has soared from 20GB a day a couple of years ago to the current daily level of 200-300 GB. With a 120-node Linux cluster and 25 core servers, IT's task is to ensure user data processing needs are always being met."

THE SOLUTION

It is little wonder, then, that the centre was an early adopter of storage network solutions – and one key area of investment has been in Nexsan technology. Last year, the centre bought two 21TB SATABeasts, Nexsan's powerful, high-capacity storage system, and has recently followed that up with the purchase of two more (See panel on page 2 'The Beast Unleashed').

Bradley points specifically to the SATABeast's advanced mechanical design, which not only provides excellent cooling, reduced drive vibration and high levels of energy efficiency, but also makes the SATABeast extremely reliable. For whilst failure of the centre's IT systems would have a highly detrimental effect on all users – and especially its researchers – protection of the data is paramount. "We have research programmes that have been running now for three years or more. Raw data is our life's blood and that is why it is safeguarded so well. There is really no way to put a price on how valuable that is," says Bardsley.

"The underlying concern for us always with any equipment is resilience to failure. Nexsan offered the industry-leading storage density to handle our data and it is superior in a number of ways. Though we have experienced failures with other manufacturers, Nexsan's technology has never failed."

He was also impressed by the level of attention and involvement that Nexsan exhibited after the centre's latest SATABeast purchases from reseller, S-Store. "They brought in one of Nexsan's VPs at one point and we were able to quiz him about the technology in some detail. It's unusual for someone of that standing to become involved in that way and we were able to establish the direction Nexsan is taking, which was very valuable to us, in terms of planning our own future needs."

“ Scientific, research-focused organisations like ours are faced with extremely complex storage management tasks.

Raw data is our life’s blood and that is why it is safeguarded so well. There is really no way to put a price on how valuable that is. ”

DR. TIM BARDSLEY

IT MANAGER, WELLCOME TRUST CENTRE FOR HUMAN GENETICS

The recent investment was triggered in part by a major collaborative project on which the centre is embarked. “The requirement, capacity wise, was to give our people on-line DAS [Direct Attached Storage], so storage consolidation was a major driver. That has increased our overall data capacity to 42TB – and we are using around 30TB of that already.”

The decoding of the human genome has opened up a world of possibilities that have transformed scientific research in a way that could not have been contemplated a few short years ago. SATABeast now forms part of the insurance policy that underpins the work for the Wellcome Trust Centre for Human Genetics, with further investment in Nexsan technology likely in the not so distant future, as data storage needs continue to expand.

ABOUT NEXSAN

Nexsan® is a leading independent provider of disk-based storage systems purpose-built and priced for the mid-market, offering industry-leading reliability, space and power efficiency. Nexsan storage systems provide scalability, integrity and security for growing volumes of unstructured data and are ideal for virtual storage, data protection, secure online archiving, bulk and cloud storage applications. Overcoming the challenges of traditional storage, Nexsan delivers a different kind of storage experience with easy-to-use, efficient and enterprise-class solutions that reduce the complexity and cost of storage. Nexsan delivers its storage systems through a select global partner ecosystem of solution providers, OEMs and system integrators. Nexsan is based in Thousand Oaks, Calif. For more information, visit the company’s website at www.nexsan.com.

©2011 Nexsan Corporation. All rights reserved.