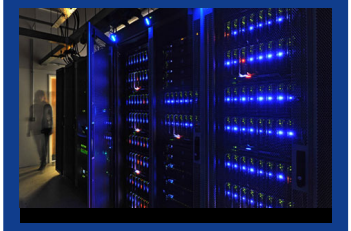
CASE STUDY:
WELLCOME TRUST CENTRE FOR HUMAN GENETICS



The Wellcome Trust Centre for Human Genetics (WTCHG) is a research institute of the Nuffield Department of Medicine at the University of Oxford. The Centre is an international leader in genomics, statistical genetics and structural biology; and collaborates with research teams from across the World on a number of large-scale studies. It houses the second largest next-generation sequencing facility in England. The Centre's research budget from competitively-won grants is close to £20m annually, and it publishes around 300 primary papers each year.

**http://www.well.ox.ac.uk/**

OCF

High performance data processing, management and storage

# Human Genetic Research Boosted By New HPC Cluster

## Business Impact Summary

### Challenge:

Maintain the pace of research and satisfy the demand for HPC with the growing complexity of genetic research.

### Solution:

A new high performance server cluster, built using IBM, Fujitsu, Mellanox and DDN technology.

### Result:

Researchers can put through around 5 times more work since the integration of the new cluster, whilst maintaining the same energy footprint.

## Changes in Science Driving Need for More Memory

In the early days of genetics research, scientists focused on rare diseases with single identifiable causes, because they only needed a small group, to work out what was happening and deliver quick wins.

Today, genetics research now looks at significantly more complex human diseases that require researchers to compare a vast amount of genetic information of individual people, all at the same time.

Genetics projects at WTCHG include national and international studies on various cancers, type-2 diabetes, malaria and analysis of bacterial genomes to trace the spread of infection. It also led the statistical analysis of the genetic make-up of the British population.

There are many genetic features to discover, some of which can lead to diabetes or obesity or heart disease. For example, when studying type-2 diabetes, one of WTCHG's teams found 80 genetic links and predicted it would be possible to identify about 500 more if only they could source the patient data to support their work. The diseases that are now being researched are incredibly complex.

The Centre has already been sequencing over 500 genomes per year, whilst storing around 20,000 genomes - just over half a petabyte – on high-speed disk for analysis.

## Challenge

WTCHG required an increase in HPC compute power to support its genetic researchers and improve the efficiency of analyses.

Dr Robert Esnouf, Head of the Research Computing Core at WTCHG says: "We need to maintain the pace of research right now, and we need more HPC - big file systems, big memory and big clusters. We have learned from past experience that we need to tailor our compute hardware to give us an edge in 'all-against-all' analyses of hundreds of genomes; lining up multiple genomes against each other and using sophisticated statistics to compare them and spot differences that might explain the genetic origin of diseases or susceptibility to diseases."

## Solution

The Centre's new supercomputer, working alongside the previous production cluster, is using Fujitsu, Mellanox and DDN technology and provides a 2.6x performance increase over its predecessor built in 2011.

"OCF worked in partnership with us to design a high-specification cluster and storage system that successfully met our needs," says Dr Esnouf. "OCF then delivered and integrated the new system on time and to budget. It also provided training to the WTCHG team. The OCF-Fujitsu-Mellanox-DDN package was the clear winner from almost all perspectives. It is based on IBM Spectrum Scale (formerly known as GPFS), to which we were already committed, and had the winning price and performance combination, together with low power consumption, fast I/O and simplicity of installation. We even managed to afford a pair of additional cluster nodes with 2TB real memory each for really complex jobs – also through OCF."

"Research is driving our adoption of HPC. Compared to this time last year, our researchers can put through around five times more work and are doing so on a machine with the same energy footprint."

Dr Robert Esnouf, Head of Research Computing, WTCHG

## Benefits

The Centre's Research Computing Core now manages more than 4000 compute cores and 5PB storage, making it one of the largest departmental computing facilities in a UK university.

By understanding the characteristics of key genetics applications and optimising how they map onto the new cluster's architecture, the Centre has also been able to improve dramatically the efficiency of some analyses, from months to weeks. The new cluster has also proved itself to be perfectly suited to supporting research by the Centre's Division of Structural Biology (STRUBI) and it has already produced some of the world's highest-resolution electron microscopy reconstructions – revealing structural details vital to understanding processes such as infection and immunity.

Dr Esnouf concludes: "Research is driving our adoption of HPC. Compared to this time last year, our researchers – and we have about 100 active users – can put through around five times more work and are doing so on a machine with the same energy footprint. With the support of OCF and its hardware partners, like DDN and Fujitsu, we're now fully armed to meet today's challenges."

About

OCF is a high performance data processing, data management, data storage and data analytics provider. We aim to successfully meet the significant "big data" challenges of UK organisations – public and private, SMBs up to complex organisations. Website: www.ocf.co.uk