

Heterogeneous Supercomputer Advances Viral Research and Animal Health

Atos supercomputer with Intel® processors helps The Pirbright Institute safeguard livestock and humans from the rising threat of viral diseases

When a deadly virus emerges, scientists must respond rapidly to characterize the virus, track its spread, and stop it from devastating livestock and possibly infecting humans. As a global leader in this work, The Pirbright Institute in the UK needs flexible high-performance computing (HPC) resources that can handle a wide variety of workloads. Pirbright deployed a Bull supercomputer from Atos powered by Intel® Xeon® processors. With a unified environment running its diverse applications, Pirbright enhances scientific productivity and helps policymakers respond effectively when a viral outbreak threatens.

Challenge

- Rising demand for computational resources threatened to outstrip the capacity provided by Pirbright's legacy mix of servers, clusters, and workstations.
- Pirbright sought a versatile system that could accelerate progress in areas such as genome assembly of complex viruses and hosts, epidemiology studies to monitor virus migration, and the development of innovative analytics tools.

Solution

- Taking advantage of Intel® technologies and Atos expertise, Pirbright implemented a Bull HPC system with heterogeneous nodes based on the Intel® Xeon® processor E7 and E5 families.

Results

- Scientists get rapid results, gaining insights that can help reduce the impact of viral disease, ensure food security, and improve the quality of life for animals and humans. Whole virus genomes are sequenced within 24 hours, enabling pinpoint tracking of a virus' path and aiding the development of diagnostic tools, vaccines, and treatments.
- By running its diverse workloads on a unified environment, Pirbright lowers management overhead and avoids the time and expense of moving vast amounts of data among multiple systems.

Preserving Life, the Economy, and the Food Supply

Agriculture is a vital element of the UK's economy, putting food on the table and paychecks in the pockets of nearly one-half million people.

Viral diseases can be deadly to livestock as well as economically devastating to the nation. A single, persistent virus—foot and mouth disease (FMD)—has an annual global impact of up to USD 21 billion in animal production losses and vaccination costs. Even where FMD is well controlled, outbreaks of the disease still cause losses of more than USD 1.5 billion each year¹. If a virus such as avian influenza crosses to human populations, the impact is even broader.

Now, virus activity is on the rise, due to factors such as population movement,

global trade, and climate change. "We're seeing an unprecedented movement of viruses globally," says Dr Bryan Charleston, director and CEO of The Pirbright Institute. "Viruses are moving in a much more dynamic system than we've seen for many years. New strains are emerging, and previously localized strains are threatening new geographic regions."

Located in Surrey, England, Pirbright is the UK's flagship research center focused on giving the UK capabilities to predict, detect, understand, and respond to economically important viral diseases of livestock. This work aids in tracking the global spread of deadly viruses and developing diagnostic tools, vaccines, treatments, and policies that can control, contain, and eliminate them. Pirbright's breakthroughs also suggest

breeding strategies that can increase disease resistance. Strategic funding is from the UK's Biotechnology and Biological Sciences Research Council (BBSRC).

Pirbright houses research programs, diagnostic labs, bio-containment labs, and computational scientists in the same facility. "This leads to rapid movement of information from the field, where biological samples are gathered, to our research laboratories," says Charleston. "It also allows our research outputs to move quickly to people who are developing and implementing strategies for the control of diseases in the field. We are constantly providing advice to governments around the world as well as to pharmaceutical and diagnostic companies."

Real-Time Insights

What makes one virus particularly virulent? What enables some animals to resist a viral infection? How is a virus changing at a genomic, cellular, molecular, and population level as it spreads?

To answer these and other questions, Pirbright scientists integrate genomics, proteomics, transcriptomics, and other "omics" data, along with epidemiological, phenotypical, and environmental information, and advanced methods in bioinformatics, statistics, and mathematical modeling. These studies are carried out not just on viruses themselves, but also on vectors (mosquitos and midges) and

hosts (livestock). By combining these methods, molecular epidemiologists can trace the movement of a virus and determine how it is evolving.

Many workloads are compute and/or data intensive. Genomes must be sequenced multiple times, often using different tools and protocols. Many viruses of interest are RNA viruses rather than DNA viruses. These RNA viruses mutate so rapidly that understanding their properties often requires analyzing a swarm of closely related genomes, raising the complexity of the genomics challenges.

Fast turnaround is essential. "With high-throughput sequencing and enough computational power, we can actually plot transmission routes and see the differences at a forensic level and understand how the virus is changing as it spreads," says Charleston. "We can trace an outbreak in real time and pinpoint where it has come from. It's this detailed understanding in real time that allows the outbreak to be brought under control."

¹Data in this paragraph are from T.J.D. Knight-Jones and J. Rushton, The Economic Impacts of Foot and Mouth Disease, Preventive Veterinary Medicine, Nov. 1, 2013. Accessible at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3989032/>.

Powerful, Flexible System for Diverse Workloads

As the demand for computational resources has risen, Pirbright began to outgrow the mix of workstations, servers, and small clusters it had implemented over the previous decade. Storage was a particular problem. Different groups had local storage resources, so collaborative projects and mixed workloads often required transferring large data volumes between aging systems.

Pirbright's leaders saw an opportunity to improve scientific productivity by modernizing their legacy infrastructure. They worked with Atos to design a unified environment that could increase performance and throughput for the institute's workloads.

“Purchasing a supercomputer is not like going to a supermarket and picking something off a shelf, especially when you have so many different user applications to support,” says Dr. Paolo Ribeca, head of Pirbright's Center for Integrative Biology and Bioinformatics. “We did a lot of design and discussion with Atos to define the system we wanted. Now we have a centralized resource with heterogeneous nodes so we can work on all the classes of problems we have to face. With the Intel® processors, we have excellent performance. And for all practical purposes, we have virtually unlimited storage.”

Dr. Paolo Ribeca,
Head Center for Integrative Biology and Bioinformatics
The Pirbright Institute

Performance and Memory with Intel® Technologies

The new system combines primary nodes built around the Intel® Xeon® processor E5 family, with memory-intensive tasks such as genome assembly running on Intel® Xeon® processors E7-4809 configured with 1.5 TB of memory. The environment also includes a half-petabyte of disk storage.

The choice of Intel technologies was straightforward. “Nowadays, there is not much that is suitable for our work that is not Intel,” says Ribeca. “With the Intel Xeon processor E5 family in the primary nodes, we have lots of computational power. With the Intel Xeon processor E7 family, we have large-memory nodes that we use for genome assembly and heavy-duty sequence analysis, where typical algorithms require large amounts of memory. So virtually all our users can perform any type of work in

a friendly environment without having to worry about problems of bandwidth, data transfer, storage, and so forth.”

In addition to accelerating Pirbright's scientific applications, the new platform speeds the development of innovative analytical and computational methods—an area in which Ribeca and the Center are international leaders—that are being customised for the study of viruses.

“We're doing non-standard method development for looking at new viruses that have not been characterized previously,” says Ribeca. “We're studying the genomics of viral swarms and the evolutionary landscapes of viruses with a complex population structure—we need to model the interaction between those viruses, the host and, possibly, one or more vectors.

That requires ingenuity, and the correct combination of algorithms and statistics into novel methods. Thanks to the Atos platform and Intel technologies, we can develop such methods much more quickly, and test them without having to move terabytes of data through the Internet.”

The Atos system with Intel technologies is also more reliable, manageable, and energy-efficient compared to the previous legacy environment, Charleston observes. He and Ribeca also plan to explore new Intel technologies that are highly relevant to their workloads, including Intel® 3D XPoint™ technology, a revolutionary advance in nonvolatile memory and storage, Intel® Xeon Phi™ processors, and future-generation Intel processors.

Faster Time-to-Results

By accelerating time-to-results and improving data sharing, Pirbright's powerful new platform strengthens the Institute's commitment to be the world's leading research center for preventing and controlling viral diseases of livestock. The platform's return on investment will be measured not only in cost savings achieved through a unified modern environment, but in the lives and well-being of animals and ultimately humans.

“In an emergency situation, we need to sequence and analyse larger data volumes within the disease containment envelope. The Atos* platform and Intel® processors give us the speed to carry out that analysis quickly and without all the problems of bandwidth and moving huge volumes of data around. This allows us to respond faster to the disease.”

Dr Bryan Charleston,
Director and CEO
The Pirbright Institute



Photo courtesy of HDR Architecture, Inc., © 2014 James Brittain

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps. Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software, or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Copyright © 2017 Intel Corporation. All rights reserved. Intel, the Intel logo, Intel Xeon, Xeon Phi, and Xeon Phi are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others. Find the solution that's right for your organization. Contact your Intel representative or visit www.intel.com/solutions To learn more about Atos solutions for high performance computing, please visit https://atos.net/en/ and https://bull.com/

All trademarks are the property of their respective owners. Atos, the Atos logo, Atos Copia, Atos Consulting, Atos Workgrid, Bull, Canopy, equisWorkline, Unity, Workline and Zero Email are registered trademarks of the Atos group. Atos reserves the right to modify this document at any time without notice. Some offerings or parts of offerings described in this document may not be available in your country. This document does not represent a contractual commitment. © Atos 2017. All rights reserved. May 2017 © 2017 Atos