

### Overview

The EBI is a non-profit organization that manages and disseminates crucial genomic information to an ever-increasing and internationally diverse research community.

### Challenge

To give international researchers fast and easy access to its genomic databases, the EBI needed to implement a cost-effective, transparent and scalable solution that utilized existing heterogeneous IT resources.

### Solution

To address the ongoing manageability and performance of computing resources, EBI implemented a grid solution using Platform LSF® and Platform MultiCluster.

### Results

- Platform LSF enables EBI to optimize the power of its existing hardware, boosting enterprise productivity and achieving faster results.
- Platform MultiCluster enables a geographically dispersed team to access its genomic and biomolecular databases, accelerating genomic and biomolecular discovery through improved collaboration.

As a non-profit organization, the European Bioinformatics Institute (EBI) is responsible for the management and dissemination of crucial genomic information to an ever-increasing and internationally diverse research community. Located at the Wellcome Trust Genome Campus in the United Kingdom, it is in close proximity to the major sequencing efforts at the Sanger Centre and HGMP Resource Centre.

A Platform LSF® customer since 1997, the EBI uses Platform LSF to optimize its IT resources, which include a 140-CPU Linux PC farm and clusters of Compaq AlphaServers and SGI Origin servers to run the massive 200-gigabyte database of flat files on protein and genomic sequencing. However, the EBI also needs to share its growing genomic sequencing research databases with scientists and researchers around the world. With Platform, it found a proven commercial solution in Platform MultiCluster, which controls and coordinates jobs across multiple geographic locations to enable high throughput capacity and capability.

### Business Challenge

Recently, the Institute decided to make a number of changes to further increase the manageability and performance of its computing resources. This included adopting Red Hat's open source Linux operating environment and splitting the existing domain into a number of sub-domains, spread across its PC farm of 80 CPUs.

According to Peter Stoehr, Head of Database Operations, at EBI, *"We needed to harness all of our compute resources to balance our intensive research and service demands. The data in our databases is doubling each year, so we required a scalable and robust multi-platform solution that could handle significant changes to our system."*

For example, each sequence query can take anything from two minutes to several hours to complete, working from a 200-gigabyte database of protein and DNA sequences. The enormous scale of a genome is expected to double these file sizes every year and EBI must be able to provide its service to the scientific community 24 hours a day.

*"Research, just like IT, is about sharing and furthering knowledge,"* concludes Stoehr. *"With this project, we are combining the two to create an environment in which the scientific community can accelerate its genomic and biomolecular research. This will inevitably benefit generations to come."*

*"Grid computing answered our requirement for a scalable computing infrastructure to support research and development activities within a demanding and heterogeneous computing environment,"* said **Petteri Jokinen, Head of Systems and Networking, EBI.** *"With Platform's solutions, we can harness all of our compute resources to fulfill intensive research demands. In addition, the data in our flat files is doubling every year, so it is important that we have a scalable and robust solution. To date, Platform's solutions have boosted enterprise productivity and enabled us to achieve faster results."*

## Solution

The EBI evaluated several different systems before finally deciding to invest in Platform LSF and Platform MultiCluster. *"We wanted the reassurance and commitment of a commercial organization, as well as having the high level of service and support that such companies offer,"* says Petteri Jokinen, Head of Systems at EBI. *"Once we had made that decision, Platform's reputation in distributed computing, its strong customer references and understanding of our requirements were huge influencers in our choice of solution."*

Platform LSF provides transparent on-demand access to EBI's computing power and allocates the most available compute resources to each search and ensures that the system provides continuous and transparent availability and resource sharing.

*"Since the implementation of Platform LSF we've found that we can harness the power of our existing IT hardware,"* says Petteri Jokinen. *"As a non-profit organization, it is essential for us to maximize the productivity of our existing IT investments and Platform has enabled us to do this. Now, we can confidently provide compute power to the increasing numbers of researchers requiring access to our vast databases."*

To address the ongoing manageability and performance of computing resources, EBI has implemented a grid solution using Platform MultiCluster. The software controls and co-ordinates jobs across multiple geographic locations to enable high throughput capacity and capability. This solution is helping EBI realize its vision to connect its own infrastructure with other international research organizations, with Platform's software controlling and coordinating computing resources across the entire grid.

The EBI is using a reliable and effective grid computing framework for sharing and accessing remote resources across wide-area networks, allowing the organization to work towards its goal of linking up data between geographically diverse research institutions.

## Customer Site

The European Bioinformatics Institute (EBI) ([www.ebi.co.uk](http://www.ebi.co.uk))

The European Bioinformatics Institute (EBI) was established in 1992 to ensure that the growing body of information from molecular biology and genome research is placed in the public domain, freely accessible to the scientific community in ways that promote scientific progress. The non-profit organization is part of the European Molecular Biology Laboratory (EMBL), an international network of research institutes funded by contributions from fifteen countries and dedicated to research in these fields. By building, maintaining and making available databases and information services relevant to molecular biology, EBI serves researchers in molecular biology, genetics, medicine and agriculture from academia, agricultural, biotechnology, chemical and pharmaceutical industries.



## About Platform Computing

Platform Computing's intelligent, practical enterprise grid software solutions help organizations optimize IT resources to Accelerate Intelligence™. We plan, build, run and manage grids that link IT to core business objectives, and help our customers improve service levels, reduce costs and enhance business performance. With industry-leading partnerships and a strong commitment to standards, we are at the forefront of grid software development, propelling over 1,600 clients toward powerful insights that create real, tangible business value. For more please visit [www.platform.com](http://www.platform.com).



### World Headquarters

Platform Computing Inc.  
3760 14th Avenue  
Markham, Ontario  
L3R 3T7 Canada  
Tel: 905 948 8448  
Fax: 905 948 9975  
Toll-free tel: 877 528 3676  
[info@platform.com](mailto:info@platform.com)

### United States

Boston: 781 685 4966  
Detroit: 248 359 7820  
Reston: 703 251 4850  
Newport Beach: 949 798 5654  
New York: 212 672 1770  
San Jose: 408 392 4900

### Asia-Pacific

Beijing: +86 10 62381125  
Hong Kong: +852 2869 5687  
Tokyo: +81 3 5326 3105

### Europe

Düsseldorf: +49 (0) 2102 610390  
London: +44 (0) 1256 370500  
Paris: +33 (0) 1 41 10 09 20

For more information, visit [www.platform.com/contactus](http://www.platform.com/contactus)

Copyright © 2005 Platform Computing Corporation. ® TM Trademarks of Platform Computing Corporation.  
All other logo and product names are the trademarks of their respective owners, errors and omissions excepted. Printed in Canada. Platform and Platform Computing refer to Platform Computing Inc. and each of its subsidiaries.