

CASE STUDY

e-Therapeutics Leverages GridGain Solution for Complex and Memory-Intensive Network Pharmacology

Modeling and Computations are 20 Times Faster with GridGain In-Memory Data Fabric



e-Therapeutics plc
The Network Pharmacology Company

CHALLENGE

Drug discovery and development company e-Therapeutics was running millions of computations repeatedly in parallel, screening various drug compounds against computer models of diseases. The company needed a solution to enhance the performance and reliability of its system and accelerate the time these cycles took.

SOLUTION

After looking at several open source and commercial solutions, the company selected the GridGain In-Memory Data Fabric to address its large-scale computational challenges.

BENEFITS

- e-Therapeutics enjoys increased reliability of its data infrastructure, a relatively straightforward programming model and higher performance.
- e-Therapeutics modeling and computations are now 20 times faster with GridGain and take days rather than weeks to complete.
- e-Therapeutics now has redundancy for failover so if a node or even a few nodes were to go down, testing and analyses can still go on as usual.

Founded in 2001, e-Therapeutics is a drug discovery and development company based in the United Kingdom. A pioneer of network pharmacology, the company's mission is to apply network pharmacology to find new treatments for human disease.

ADVANCED NETWORK PHARMACOLOGY PLATFORM

Based on advances in chemical biology and network science, "network pharmacology" is a distinctive new approach to drug discovery. It involves the application of network analysis to determine the set of proteins most critical in any disease, and chemical biology to identify molecules capable of targeting that set of proteins. e-Therapeutics has developed a proprietary network pharmacology platform for analyzing networks of proteins associated with particular diseases. After these analyses—which number in the millions—drug candidates with optimal impact are identified.

THE CHALLENGE: MILLIONS OF COMPUTATIONS RUNNING IN PARALLEL

While designing an extension to the platform, e-Therapeutics had to address the challenge of running millions of in-memory analyses repeatedly, screening various compounds against multiple computer models of diseases with different parameters to see which compounds would be most active. Testing these would have taken weeks on the company's system.

"We were looking into how we could implement the system that did the analyses, so we looked at doing it ourselves, and using both open source and commercial solutions," said Jonny Wray, Head of Discovery Informatics for e-Therapeutics. Then Wray remembered another product he'd used in the past – a grid-based solution from GridGain Systems. "I had used GridGain's In-Memory Data Fabric in the past so I knew we should consider it. Without a data fabric solution, our network pharmacology platform would take weeks to test all the compounds against one disease model, so we evaluated GridGain."

"The scientists have no idea the platform is running on GridGain. Things just work."

— Jonny Wray, Head of Discovery Informatics for e-Therapeutics

THE SOLUTION: GRIDGAIN IN-MEMORY DATA FABRIC

The GridGain In-Memory Data Fabric is an advanced data management system, available as an open source distribution, that offers distributed in-memory caching, distributed computations, and streaming. It addresses performance and provides significant scalability for e-Therapeutics by allowing the network pharmacology platform application to compute on it. The GridGain In-Memory Data Fabric is now a key component of e-Therapeutics' network pharmacology platform. Within this heterogeneous environment, GridGain provides a high performance, low latency In-Memory Data Fabric enabling fast computations for multiple different analytic approaches, not just the original platform extension that triggered the use of GridGain.



According to Wray, with GridGain in place, the network pharmacology modeling and computations are 20 times faster than before.

“While that’s impressive, it’s more than the fact that we’re more than 20 times faster than without the GridGain solution. It’s the fact that we can now do what we couldn’t do before,” said Wray. “Just a couple of these tests, without the data fabric, would take weeks and you would need to run many of them in parallel, so it wasn’t something we necessarily would have considered. With GridGain, they take just days and as such, made the platform extensions feasible.”

BENEFITS: SEAMLESS INTEGRATION, TRANSPARENCY AND EFFECTIVE RESULTS

Wray appreciated the ease of integration for the GridGain solution. “While we had to write some code to integrate the solution, GridGain’s programming model makes it really straightforward and easy. They provided all the infrastructure needed to allow us to write our code for the grid, which was one of the reasons we chose GridGain.”

“GridGain’s programming model makes it really straightforward and easy. They provided all the infrastructure needed to allow us to write our code for the grid, which was one of the reasons we chose GridGain.”

— Jonny Wray, Head of Discovery
Informatics for e-Therapeutics

“[While impressive,] it’s more than the fact that we’re more than 20 times faster than without the GridGain solution. It’s the fact that we can now do what we couldn’t do before.”

— Jonny Wray, Head of Discovery
Informatics for e-Therapeutics

Wray says that the GridGain solution is transparent to the e-Therapeutics scientists – in a good way – until something goes wrong, which it seldom does. “The scientists have no idea the platform is running on GridGain. Once it was deployed, they don’t know any differently – things just work,” said Wray. That is until something goes wrong, which actually leaves them with a situation better than the one they had before GridGain anyway.

“We have had a couple of technical issues, where some of our 20 compute nodes crashed, so we’ve had to adjust and work with the smaller grid, but the analysis still works faster than before we had the GridGain solution,” said Wray. “That actually speaks to the reliability of GridGain. If our prior solution had failed, there was no redundancy and we’d have had to shut down altogether.”

e-Therapeutics is happy with its selection of the GridGain In-Memory Data Fabric for its high-performance computing requirements. “We’re pleased with the results, which give us 20x faster data processing performance,” said Wray. “We get great support from GridGain when we need it, and we are looking into working with them on a bigger solution and possibly an on-demand cloud solution.”

ABOUT E-THERAPEUTICS

e-Therapeutics is a public biotechnology company, listed at the London Stock Exchange, which has developed a proprietary platform in network pharmacology representing an innovative approach to drug discovery based on advances in network science and chemical biology. The Company’s discovery and development activity is focused in cancer and disorders of the nervous system. e-Therapeutics is based at sites in Oxford and Newcastle, UK. For more information about the Company please visit www.etherapeutics.co.uk.

ABOUT GRIDGAIN™

GridGain, the leading provider of the open source In-Memory Data Fabric, offers the most comprehensive in-memory computing solution to equip the real-time enterprise with a new level of computing power. Enabling high-performance transactions, real-time streaming and ultra-fast analytics in a single, highly scalable data access and processing layer, GridGain enables customers to predict and innovate ahead of market changes. Fortune 500 companies, top government agencies and innovative mobile and web companies use GridGain to achieve unprecedented computing performance and business insights. GridGain is headquartered in Foster City, California. To download the GridGain In-Memory Data Fabric, please visit <http://www.gridgain.com/download/>.

