

Pharmaceutical Company

Client profile

OCF's client is a global pharmaceutical company, with business units focused on every element of the discovery, development, trials, marketing and sales of drugs. Our client in particular is the Chemistry department, responsible for early compound screening and target identification.

Summary

OCF and IBM worked in partnership with the computational Chemistry department to design and deliver a high performance clustered-storage solution. The solution when fully integrated will deliver 200% more data at less than half the CPU overhead of the previous system. This enables more compounds to be screened as well as providing a more responsive service to discovery and medicinal chemists using the compound screening

The Challenge

With block-buster drug pipelines suffering too many late-phase losses, resulting in business costs of hundreds of millions the pressure is on make the process more efficient. One way to achieve this is to utilise the enormous compound libraries of the corporation more effectively; to help weed-out compounds that will fail at trial and to identify more potential targets.

This process is highly compute and data-intensive, requiring the cross-referencing of millions of compounds against a constantly increasing flow of targets. Delivering against this exponential growth curve in demand is critical to feeding the drug discovery pipeline and therefore the future success of the corporation.

OCF and our partners, IBM were tasked with optimising the delivery of data to the existing OCF/IBM supplied computational cluster. Our target was to reduce the significant overhead generated by the classical NFS infrastructure, increase the delivery rate (aggregate and node) and provide a scalable solution.

Solution

Although the client's corporate standard was HP and EMC for storage technology, the client needed a solution that could integrate seamlessly with their existing IBM high-performance compute clusters. OCF could offer the unique blend of integration skills and

IBM technology to deliver a high performance, resilient, manageable and scalable solution.

Our solution was to integrate high-throughput IBM's xSeries Servers, DS4000 storage and innovative GPFS file-system. The solution delivered more than 4TB of high-throughput storage, optimised for the delivery of data into a parallel clustered compute environment,

GPFS is a unique, parallel file-system designed to operate in high-performance environments; it provides a unified, shared file-system to multiple, parallel storage resources and distributed clients. It has the unique ability to enable DS4000 storage sub-systems out-perform high-end proprietary systems.

Benefits

The biggest benefit delivered to our client is the acceleration of the compound screening cluster. We are able to deliver data 200% faster and with less impact on the performance of the cluster using GPFS's unique and innovative technology. This will impact on the scientific productivity of the computational chemistry group, saving potential losses in late-stage failures and increasing the likelihood of target identification.

Conclusion

The client has plans to take advantage of the improvement in the delivery of compound data by upgrade the computational cluster in 2006 using IBM BladeCenter technology.

As a result of this project, an additional site within the same corporation is in the process of procuring a similar system to help speed the screening of compounds in another department.