

**Customer Line-Of-Business:****Biological Life Sciences Research****Project Objective:****Reduce Cost and Complexity of Retention Lifecycle Management****Introduction**

GridBank is deployed by a large government sponsored biological research institute in the UK. The group invests US\$500mm per year in research that covers the whole of the microbial, plant and animal kingdoms. It is one of the few public bodies that support basic research in the veterinary and plant sciences. The institute supports over 6000 scientists in universities and other research institutions to promote and encourage knowledge transfer from, and commercialization of, the research it funds.

**Experimental Data**

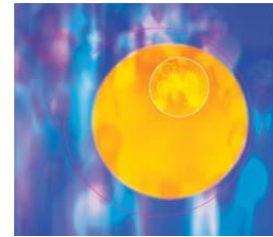
Over the last ten years the success of numerous genome sequencing projects has had a major impact on biology and much effort is now focused on understanding protein function at different levels. As a result, the landscape of global research in human genetics has permanently changed.

The development of mathematical models enabled by data collection from electron microscopes, DNA sequencers and other tools have become vital to more quantitative biological research practices. The data and models facilitate research and testing by computers as a complement to traditional means of observation and experimentation and enabling predictive biology to become a reality. A major challenge for the 21st century is to build on and expand the 'molecular' approach to biological and genetic research.

These advancements are resulting in creation of huge volumes of data that is expensive and time consuming to create however is not accessed or changed frequently. The usage patterns and volume characteristics make it impractical to

store and manage genomic research data on primary disk waiting to be called into action by a project or analytic process.

How the data is preserved however is critical: it must be immutable and stored redundantly due to the time and processing cycles it takes to create.

**Storage Optimization**

Data related to biological research is growing at over 60% annually at the research institute. Several problems related to the management and protection of the data needed to be resolved. Specifically, the institute needed to:

1. Reduce the primary storage footprint;
2. Ensure nightly backups would complete in a predictable timeframe;
3. Classify research data sets by project(s) for fast retrieval;
4. Ensure continuous accessibility to data transparently to researchers;
5. Maintain accessibility to data by key applications;
6. Automate data management and migration based on policies;
7. Maintain failsafe replicas of data at remote site in both object-based storage format and original file formats as well.

**Solution**

GridBank offers a robust platform for managing unstructured data. The solution implemented at the research institute met all of the optimization objectives set forth above.

- Shrinking the primary storage footprint was accomplished by migrating data that was “aging and unchanging” to secondary storage achieving a 50% reduction through single-instancing and compression.
- The system was automated by policies that dictated any dataset that was not accessed in over 21 days was automatically migrated as were any files greater than 2GB in size. These two policies ensured that nightly backups could be run within a 6-hour window.
- Classification policies were defined such that based on the source of the data each object was tagged by the GridBank classification engine with a descriptor identifying the associated project enabling precise recall of all relevant data when a particular project became active.
- Transparent and continuous access to migrated data for end users was guaranteed in several ways. Some of the ingested data left stubs in its original location so users could access the data as if it hadn’t been moved at all. Others have elected to access files through the GridBank tree in their MS Outlook client. For application access, GridBank APIs are being used to integrate some applications directly with GridBank platform so the data can be directly accessed when an application calls for it.
- Finally, GridBank is automatically creating replicas of the migrated data at a remote site. If a node is interrupted, automatic failover takes place ensuring uninterrupted access to data without user or IT intervention.

## About Tarmin

Tarmin Inc. is a leading provider of storage platforms for unstructured data management. Tarmin’s award winning solution organizes data silos spread across a company’s data centers into pools of conventional storage with public or private clouds, optimizing information access while reducing the storage footprint and reducing costs. The policy driven migration engine provides transparent archiving and data access capability. Tarmin delivers infinite scalability, an object-based storage model, and a non-disruptive technology adoption path. While substantially lowering storage and IT CAPEX and OPEX, Tarmin’s data management platform enables unified retention, search and case management, and consistent policy management and security across all data sources. Tarmin creates an easily manageable approach that improves data and storage management, satisfies any organization’s regulatory, governance, and compliance requirements, and delivers a “greener” storage alternative. Tarmin has received numerous industry accolades including: Finalist for Storage Magazine’s Best Enterprise Storage Product of 2009 (Jan, 2010), one of the “Nine Storage Companies to Watch” by Network World and CIO magazines (August 2009), and Storage Magazine UK’s “One to Watch” Award (June 2009). Tarmin is headquartered in Silicon Valley, California. For more information, please visit [www.tarmin.com](http://www.tarmin.com).

## Tarmin, Inc.

465 La Canada Court, Suite 104  
Morgan Hill, CA 94301  
Phone: +1 (866) 4 TARMIN  
Email: [sales@tarmin.com](mailto:sales@tarmin.com) or [info@tarmin.com](mailto:info@tarmin.com)

© 2011 Tarmin™ Inc. All rights reserved. Tarmin, GridBank, NOADS, and the NOADS logo are trademarks or registered trademarks of Tarmin in the United States and other countries. Other names may be trademarks of their respective owners. Tarmin, Inc makes no warranties, expressed or implied, by operation of law or otherwise, relating to this document, the products or the computer software programs described herein. The information contained in this document is subject to change without notice. Tarmin Inc assumes no responsibility for any errors that may appear.

All other trade names, trademarks, registered trademarks and service marks used and mentioned in this document are the rightful property of their respective owners.