



Tarmin GridBank for Research Data

Revolutionizing Data Management

Overview

Large universities and research establishments generate ever more powerful ways to create, consolidate and crunch huge volumes of data to drive new insights and conclusions. As the volume, velocity and variety of these data repositories grow exponentially, the challenges of storing, maintaining and analyzing them become ever greater. Meanwhile, funding streams for science and research are trending downwards while organizations become more interconnected and reliant on external data sources and commercial and regulatory retention and repository centralization requirements become increasingly stringent. Finally, high performance computing is driving the need for high throughput and IOPS from underlying storage technologies. Tarmin GridBank, the premier Data Defined Storage platform, delivers a data centric platform that consists of flexible automatically tiered storage, which significantly reduces storage costs, while controlling risks and enabling data insights.

Flexible, Low Cost, High Performance Storage

As the volumes of data generated and analyzed trend ever upwards, the need for flexible low cost storage is becoming ever more urgent. Long term data preservation drives the need for sustained cost reduction yet is compounded by the requirement for high performance storage for specific compute tasks. Tarmin GridBank meets these challenges with its data centric comprehensive management capability including Media Independent Data Management, Data Security and Identity Management and Distributed Metadata Repository. GridBank delivers a suite of technologies, including massively scalable hardware agnostic Object Dedupe and compression to reduce data volumes and proprietary lock-in, file system virtualization and centralized management of heterogeneous storage arrays and automated internal data migration between virtualized storage pools to optimize the performance of active projects while moving completed project data to secondary, tertiary or cloud stores. Additionally, institutions aiming to further reduce their storage total cost of ownership can leverage GridBank's chargeback functionality to facilitate the process of operating as a managed service provider for both internal projects and external organizations.

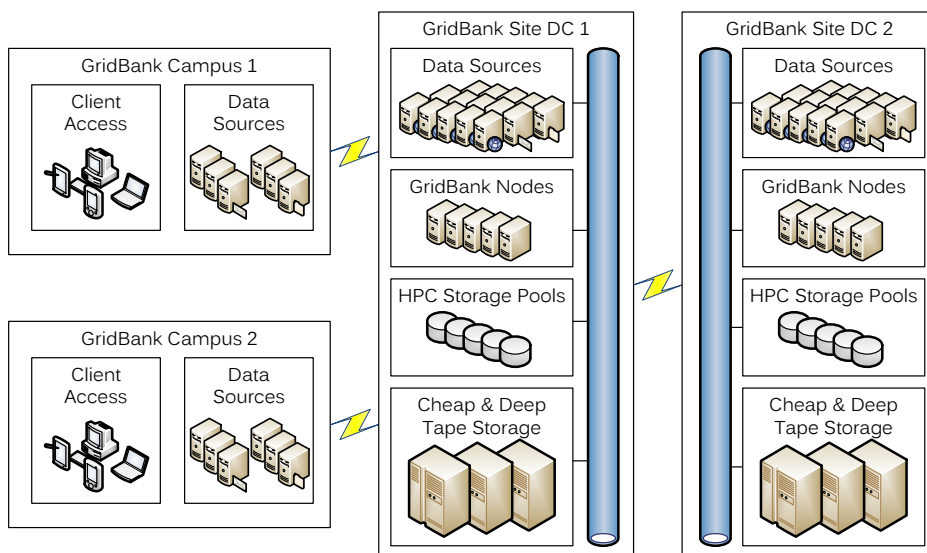
BENEFITS

- Centralize Research Data
- Meet Cost & Performance Targets
- Chargeback Storage to Projects
- Long Term Secure Data Retention
- Bandwidth Optimized Replication
- Resilient Architecture
- Unified On Demand Search
- Integrate Storage & Compute
- Extract Knowledge from Data



*"Tarmin's GridBank software
[addresses] all of the core issues
associated with managing
unstructured data over the long
term."*

Steve Duplessie
Founder, ESG



Typical University Deployment

Product Highlights

- Data Centric Storage
- File System Virtualization
- Distributed Object Dedupe
- AD/LDAP Security
- Information Governance
- GridSync & Smart Mobility
- Content Metadata Filtering
- Unified Enterprise Search
- Big Analytics Integration
- Support for Windows & Linux

Data Retention Compliance

Following high profile cases where research data has been unavailable for public scrutiny, there have been moves towards retention compliance mandates. Research is often publically funded and a project-based activity with budget allocation often controlled by individual researchers. In this environment, storage has typically been selected with a view to addressing the short term requirements for the project duration, without consideration of the long term retention needs and benefits of centralized management. As retention compliance requirements begin to come into force, institutions must offer compelling centralized storage platforms that enable researchers to achieve their objectives of optimally priced storage while delivering compliant long term preservation. Tarmin GridBank delivers a fully featured policy-based information governance framework which facilitates secure long term retention and automated tiered data migrations.

Bandwidth Optimized Replication

Educational institutions and other scientific organizations are regularly reliant on large remote data sources and replicate between remote data centers creating huge demand for bandwidth. While educational institutions have access to high bandwidth connectivity, through providers such as Internet2 or JANET, the massive data volumes involved in particle physics or genomics challenge even these network capacities. Tarmin GridBank provides multi-site replication which optimizes bandwidth consumption with distributed Object Dedupe, locally compressing and deduplicating, significantly reducing transferred data volumes.

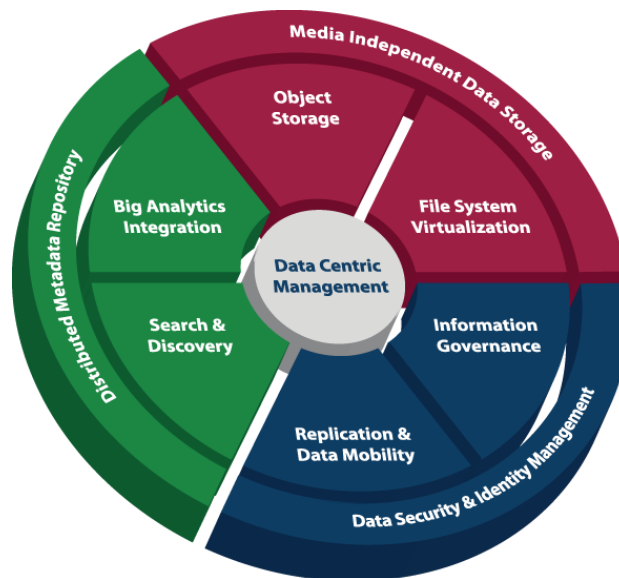
HPC Data Analytics

Universities and research institutions demand high performance computing analysis of vast datasets. This requires active project datasets retained on high throughput, directly addressable data repositories. For these requirements, high performance, reliable access to data is essential. Through enabling hooks for open source analytics frameworks, such as Hadoop, and providing an API for direct access to data, GridBank enables institutions to simplify their analytics workflow, providing a transparent storage abstraction layer between storage for HPC on active projects and storage for longer term retention, automatically migrating data to the appropriate storage tier.



"GridBank 3.0's massively scalable model allows business to consolidate their data globally across silos to reduce cost...and enhance the value of their data."

Ashish Nadkarni
IDC



About Tarmin

Tarmin Inc., the leading provider of Data Defined Storage solutions, unlocks the value of data as a strategic business enabler, delivering a massively scalable, transparent and unified approach for consistent data management, storage, retention, security and search across cloud and traditional storage infrastructure. Tarmin's proprietary GridBank solution empowers organizations to store, control and understand the value of data as a competitive business asset, no matter its size, location or cost by uniting application, information and storage tiers into a single, integrated data centric management architecture.

