

---

# LinkedIn® Scales to 200 Million Users with PCIe Flash Storage from Virident | HGST

## Challenge

After LinkedIn's launch, the number of users grew to 8 million in 2006 then exploded to more than 200 million in 2012. This exceptional increase demanded continuous platform development and the rollout of new features.

## Solution

Beginning in 2010, the Virident team worked in tandem with LinkedIn to help it understand the performance, latency, and TCO benefits of flash; explain the use cases; and go over the deployment details.

## Benefits

The increasing number of users, the company's stock market success, and its growing revenue are all indicators of a successful Web2.0 company. To continue its growth and success LinkedIn depends on the performance and reliability of Virident flash.

**The addition of Virident FlashMAX™ storage devices helped LinkedIn reduce application latency and increase performance, while lowering TCO.**

## Client Overview

LinkedIn is the world's fast-growing professional networking site that allows members to create business contacts, search for jobs and find potential clients. LinkedIn forms a vital link for more than 200 million professionals in more than 200 countries and territories who count on the LinkedIn network and consider it vital to their day-to-day professional life. They use it to stay connected with business contacts, expand their professional networks, develop closer relationships and build personal brand equity. LinkedIn has a diversified business model with revenues from member subscriptions, advertising sales and talent solutions.



The company has been the most successful social-media stock on the market since its initial public offering in May 2011. LinkedIn surpassed expectations by recording its first quarter with more than \$300 million in revenue in February 2013.

### **The Challenge**

After LinkedIn's launch in May 2003, the number of users grew to 8 million in 2006 then exploded to more than 200 million in 2012. This exceptional increase demanded continuous platform development and the rollout of new features. LinkedIn infrastructure started off with RISC UNIX systems, traditional storage-area networks and expensive databases. After experiencing unprecedented growth in just five years, LinkedIn felt it was becoming increasingly difficult to manage and grow its IT infrastructure to meet the needs of the user base.

As with other Web 2.0 industries, the workload of LinkedIn's infrastructure was very unpredictable because even a small increase in the number of users caused a significant increase in stored data, such as blogs, videos and photos. All of this data had to be processed in a very short time. In addition, user experience and application availability were of prime importance. The company's services had to be available 24 x 7 x 365 and these benefits had to be achieved at the lowest possible cost.

It was clear to LinkedIn that its existing data-center infrastructure could not scale to meet its performance needs without an architecture overhaul. The company needed to move to an open-systems architecture, allowing it the flexibility to scale, to deploy applications based on business needs and to deliver a quick response for millions of concurrent users. In short, the company needed to build an agile and flexible infrastructure to manage the large and continually increasing user base.



## The Solution

LinkedIn crunches 120 billion relationships per day and blends large-scale data computation with high volume. Meeting these needs required a storage architecture that could deliver high IOPs, low latency, and the ability to maximize the compute capabilities of the server by pairing it with superior storage. LinkedIn considered moving to the new flash technology and chose Virident FlashMAX for a variety of reasons, including its ability to deliver higher performance through higher server utilization. FlashMAX was also operationally efficient, providing high-performance storage within the server and close to the CPU, with no moving parts. Plus, FlashMAX delivered the IOPS and low latency required to sustain business applications. In addition, LinkedIn sought to use the most reliable flash technology available at that time. The initial deployment comprised Flash SLC (single-level cell) chips that offered significant write performance for critical applications. The most distinguishing feature was Virident's on-card RAID capability, which could support complete chip failure.

## End-to-end solution support

Beginning in 2010, the Virident team worked in tandem with LinkedIn to help it understand the performance, latency, and TCO benefits of flash; explain the use cases; and go over the deployment details. The team provided technical guidance about system bottlenecks within the infrastructure to achieve best performance. The deep technical partnership in early engagement, plus the uncompromising performance, enhanced reliability and data availability of flash-aware RAID, were key points in LinkedIn's selection of Virident as its choice of partner. LinkedIn began the relationship with the purchase of 10 cards to prove out the company's capabilities in their particular environment.

After one year of usage, LinkedIn was comfortable that Virident's flash technology was capable of enterprise-grade deployments. Since LinkedIn's IPO in 2011, the number of users had climbed to more than 130 million. LinkedIn evaluated the newer MLC (multi-level cell) flash technology available and felt that it met their growing needs. MLC offers a good combination of performance, endurance and reliability at a relatively lower cost.

“When we implemented the newer MLC flash technology, our goal was to take advantage of the performance, latency and TCO benefits over our existing infrastructure,” said Sonu Nayyar, senior director of production operations at LinkedIn. “We were immediately impressed with Virident’s customer support as it ramped up our team on the new technology. Once fully implemented, we experienced the high IOPs, low latency, and the ability to maximize the compute capabilities of the server proving we made the right decision.”

### **Steps to widespread deployment**

LinkedIn first deployed FlashMAX in its back-office applications. Once convinced of the products’ performance and results, LinkedIn decided to deploy FlashMAX in one of its business-critical applications running on the Voldemort distributed database system. Now Virident storage was on the front lines, supporting applications that had direct impact on online user experience. For example, FlashMAX enabled features such as “Who’s Viewed My Profile,” which produces high write loads. Other applications at LinkedIn were similarly challenging on the scaling front, such as the one used for “finding similar profiles.” While the set of all user profiles is very large, even a modest subset of all user-profile pairs is quite huge. Other applications that needed to handle hundreds of millions of reads and writes per day were moved to Virident storage; simultaneously, LinkedIn increased the amount of data it stored.

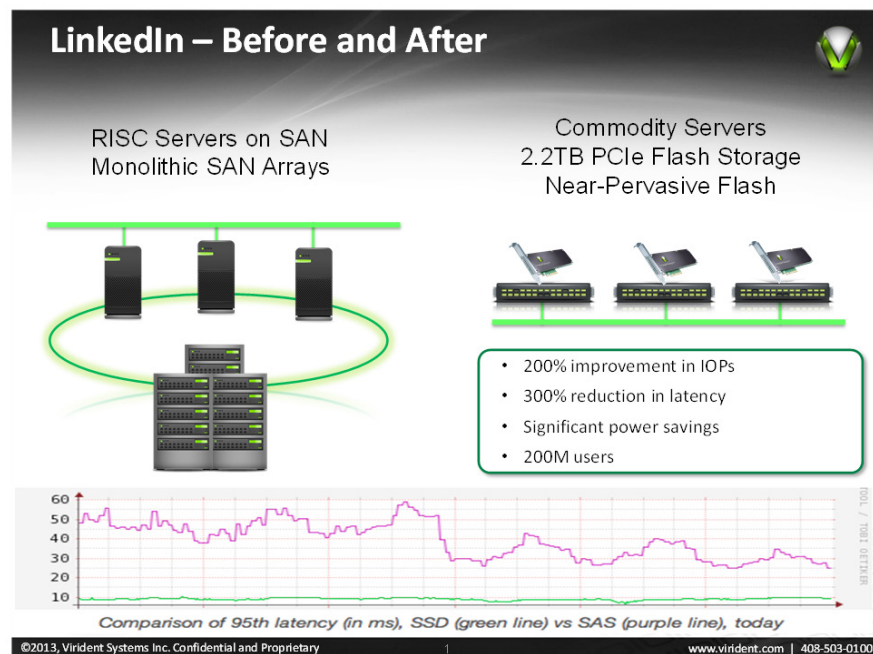
Project Voldemort was a key growth area for the company. Virident flash was used to derive the highest read/writes from the databases and to achieve server consolidation by providing more than one terabyte (TB) of flash per server. Hundreds of Virident units were deployed across all Voldemort servers, a process that took more than six months. Continued deeper technical engagement and knowledge-sharing was significant to the success of this project.

With the success of Voldemort flash-based storage, LinkedIn launched a major initiative for rebuilding infrastructure. Today, all of LinkedIn's data engineering efforts are focused on building services that can work together easily. Flash is being deployed across servers in a "near pervasive" strategy. As part of this initiative, one of the most important things that LinkedIn is building is a new in-house database system originally designed to provide a usability boost for LinkedIn's InMail messaging service. These moves are all part of a mission to create an innovative data environment at LinkedIn, with thousands of Virident flash cards now deployed in pre-production and production environments.

"Flash is transforming the datacenter and Virident is leading the flash platform transformation," said Mike Gustafson, senior vice president and general manager, HGST's SSD, Software and Solutions Business Unit. "LinkedIn has embraced the transformation and is a testament to the fact that the future data center will benefit from a server-side platform. LinkedIn and the Web 2.0 community are only the beginning and as IT becomes more familiar and comfortable with the technology we'll begin to see enterprises mimicking innovators such as LinkedIn and deployed server-side flash platforms."

The increasing number of users, the company's stock market success, and its growing revenue are all indicators of a successful Web2.0 company. LinkedIn's forward momentum is not expected to slow anytime soon. To continue its growth and success LinkedIn depends on the performance and reliability of Virident flash.

## Environment



## About HGST

HGST, a Western Digital company (NASDAQ: WDC), develops innovative, advanced hard disk drives, enterprise-class solid-state drives, external storage solutions and services used to store, preserve and manage the world's most valued data. HGST addresses customers' rapidly changing storage needs by delivering intelligent storage devices that tightly integrate hardware and software to maximize solution performance. Founded by the pioneers of hard drives, HGST provides high-value storage for a broad range of market segments, including Enterprise, Cloud, Datacenter, Mobile Computing, Consumer Electronics and Personal Storage. HGST was established in 2003 and maintains its U.S. headquarters in San Jose, California. For more information, please visit the company's website at <http://www.hgst.com>.



© 2014 HGST, Inc., 3403 Yerba Buena Road, San Jose, CA 95135 USA, Produced in the United States 11/13. All rights reserved.

Virident, Virident FlashMAX, FlashMAX Connect and the Virident logo are trademarks of HGST Inc. All other products and services listed herein are trademarks of their respective companies. HGST trademarks are intended and authorized for use only in countries and jurisdictions in which HGST has obtained the rights to use, market and advertise the brand. Contact HGST for additional information. HGST shall not be liable to third parties for unauthorized use of this document or unauthorized use of its trademarks.

One GB is equal to one billion bytes and one TB equals 1,000 GB (one trillion bytes) when referring to hard drive capacity. Accessible capacity will vary from the stated capacity due to formatting and partitioning of the drive, the computer's operating system, and other factors.

References in this publication to HGST's products, programs, or services do not imply that HGST intends to make these available in all countries in which it operates.

Product specifications provided are sample specifications and do not constitute a warranty. Information is true as of the date of publication and is subject to change. Actual specifications for unique part numbers may vary. Please visit the Support section of our website, [www.hgst.com/support](http://www.hgst.com/support) or [www.virident.com](http://www.virident.com), for additional information on product specifications. Photographs may show design models.