



## Titan Accelerates Shopzilla Shopping Engine, Delivering Customers Timely, Accurate Results

**“At Shopzilla, we use the best of breed in every vertical. We use BlueArc for its speed. None of what we do is rocket science – we just figured out how to do it well, and constantly improve that process.”**

**“We move terabytes of data through the BlueArc every single day. Users want the best experience they can possibly have, and the way to get the best experience is to have the most available data to the customer.”**

**“The BlueArc is the conduit for the flow of information.”**

**- Burzin Engineer,  
Sr. Director of Infrastructure Services,  
Shopzilla.**

### Summary

Whether searching for a new plasma television, tennis shoes or dining sets, Shopzilla's customers expect rapid, accurate results. Spanning nearly 70,000 stores and 40 million individual products, Shopzilla's massive database, one of the largest on the Internet, is refreshed multiple times daily, ensuring that users receive correct, near real-time pricing. High performance storage from BlueArc's Titan storage system enables the company to deliver updates more quickly, delivering a high quality customer experience.

### The Customer

Shopzilla is one of the world's leading Web-based comparison shopping services. The company's mission is to enable shoppers to quickly and easily find compare and buy anything, sold by virtually anyone, anywhere. Each month, Shopzilla connects tens of millions of consumers with tens of thousands of stores serving tens of millions of products. Shopzilla also operates the BizRate consumer feedback network that collects millions of consumer reviews of stores and products each year. In 2005, Shopzilla was acquired by The E. W. Scripps Company and is now part of its interactive media division.

### The Challenge

Shopzilla's goal is to provide customers with search results of the largest range of products and merchants in the most-organized manner with lightning-fast speed. The company scans data from tens of thousands of stores into a single unified structure and indexes it with a proprietary search engine. The resulting data is then published on the company's consumer-targeted Web site. In 2004, seeing dramatic growth in user accesses, merchant reach and products catalog, Shopzilla looked ahead to what would be a busy holiday shopping season and recognized their existing filer infrastructure would not be up to the task, especially as product pricing and availability were constantly changing. The company needed to find a storage solution that offered the highest throughput available, with capacity to scale, as multiple terabytes would pass through the system each day.

### The Solution

Looking for an enterprise-class network storage solution with uncompromising speed, Shopzilla selected BlueArc's Titan storage system. Initially deploying ten terabytes of high performance storage, Shopzilla has expanded their BlueArc deployment to nearly eighty terabytes over the last three years, as merchants, product listings and user traffic continue to skyrocket. Titan, acting as the conduit for the site's information flow, enables Shopzilla to deliver the freshest data to market in the comparative shopping marketplace, rotating its product database, the company's lifeblood, several times a day, putting the company well ahead of competitors, who can take several days to update content.

## The Results

Burzin Engineer, senior director for infrastructure services at Shopzilla, says that nothing about his technology setup is truly rocket science. Yet, harnessing current, accurate, data for more than 40 million products from 70,000 merchants, some of who change their information as frequently as 24 times per day, and assembling that for a user load nearing 20 million per month, requires an infrastructure that acts like a rocket – in both speed and the ability to reach aggressive targets.

In early 2004, Burzin was looking for a high performance NAS system to replace the company's existing storage filers, which had grown sluggish with age. The company, faced with the possibility of being I/O bound at a critical time in advance of the upcoming holiday shopping crunch, needed a storage solution that scaled in performance and capacity as well as supporting thousands of simultaneous connections. Following some online comparison shopping of his own, Burzin acquired one of BlueArc's first-generation systems, and testing it against their existing equipment, was very impressed with its speed, relative to the competition. "The BlueArc was just crushing it in performance," Burzin said. "We were looking for an enterprise-level NAS solution, and BlueArc was among the top competitors in sheer speed."

Following this initial testing, Shopzilla purchased a Titan storage system equipped with highest performance Fibre Channel drives, and deployed the system to enable faster publication of the site's database, acting as the conduit for the critical flow of information that bridged merchant data to customer searches. By deploying BlueArc, the company was able to accelerate their time to publish, giving them a leg up on the competition.

"The market differentiator was that our data was as fresh as fresh can get in the market," Burzin said. "That's what we can do with the BlueArc."

Shopzilla's proprietary data allocation and assignment process gathers terabytes of disparate information from its submitted merchant data, compacts it, categorizes it, and indexes the results. Index servers then pull the information, through BlueArc's Titan, using a high performance compute cluster that consists of more than 100 individual nodes. This process, which can take days at other shopping comparison sites, and took upwards of eight hours at Shopzilla prior to implementing BlueArc, can now be completed every four hours – giving Shopzilla customers the very best site experience possible, armed with the knowledge that they were seeing near real-time pricing and availability for products.

While Titan has a behind the scenes role at Shopzilla, customers recognize its benefits – ensuring that they don't click on an offering for a plasma TV at \$2,299 and find out its price is actually \$2,499, or try to buy the latest Tickle-Me-Elmo doll, only to find it sold out at the merchant site.

In addition to delivering highest levels of performance, Titan's ability to serve multiple tiers of storage enabled Shopzilla to utilize the system's high capacity threshold to include additional applications with differing demands and characteristics. By deploying high capacity, but lower performance and lower cost Serial ATA shelves in addition to the fibre channel arrays, Shopzilla also uses Titan to for disk-to-disk database backup.

## The Conclusion

To forge ahead in a highly competitive online market, Shopzilla needed to differentiate itself through a higher quality service, featuring timely data and lightning-fast search results. Titan's combination of linear performance and capacity scaling ensures that the company will continue to offer a best of breed product for consumers in a world where seconds of delay can mean the difference between making a sale and losing a customer's interest.

When it comes to getting the best products at the best prices, customers don't care what the technology underpinnings of a site are – only that it responds quickly and offers consistent data that can be relied upon. Where other vendors were unable to meet this unrelenting demand, BlueArc has enabled Shopzilla to grow dramatically from 40,000 merchants in 2004 to the 70,000 merchants with 40 million products offered today.



**BlueArc Corporation**  
 Corporate Headquarters  
 50 Rio Robles Drive  
 San Jose, CA 95134  
 t 408 576 6600  
 f 408 576 6601  
 www.bluearc.com

**BlueArc UK Ltd.**  
 European Headquarters  
 Queensgate House  
 Cookham Road  
 Bracknell RG12 1RB, United Kingdom  
 t +44 (0) 1344 408 200  
 f +44 (0) 1344 408 202