

Solution Impact Analysis

EMC Atmos Cloud Storage Helps Vistaprint Cut Storage Infrastructure Costs and Accelerate Production

By Terri McClure

January, 2011

“In the last year, we’ve gained several million new customers; at peak we process more than a million uploads per day. The Atmos cloud architecture gives us unlimited scalability and the flexibility we need to support new business models and fluctuations in demand.”

Company Profile

Industry

Business and Consumer
Marketing and Promotional
Services

~ 2,500 employees
9,600,000 customers per year

Data Center Profile

- Integrated, globally distributed IT and production print facilities
- Microsoft (.net) server environment; EMC storage
- Data centers in Bermuda, Canada, the Netherlands, and Australia
- >1PB of Atmos storage under management

Challenges

Facing tremendous business growth and the need to bring new lines of business to market faster, Vistaprint needed a massively scalable, tiered storage infrastructure that reduced costs, simplified management, and enabled event-driven, policy-based responsiveness to elastic demand cycles.

Solution

Vistaprint deployed a private cloud storage infrastructure anchored by EMC Atmos

Benefits

Vistaprint realized an 83% reduction in storage costs, significant improvements in the performance of storage operations, and nearly limitless scalability to accommodate future growth and changing business demands.

How EMC Atmos Cloud Storage Helps Vistaprint Cut Storage Infrastructure Costs and Accelerate Production

About the Company

Vistaprint is a Venlo, the Netherlands-based company with an innovative, technology-driven business model that has driven growth to service more than 9,000,000 customers annually in 120 countries with a fiscal 2010 revenue stream of \$670M. As an online provider of printed products and services to “micro-businesses” and consumers, Vistaprint’s trajectory of success aligns closely with the widespread adoption of the Internet and its democratizing effects on global business.

Vistaprint gained its advantage in the marketplace by focusing on a rapidly growing but largely underserved micro business market segment (very small businesses with 0-9 employees) and by developing processes and proprietary technology that enabled it to produce short runs of high quality printed materials at a price point that was attractive to that market segment. Vistaprint’s high-volume business model relies heavily on technology to drive down the costs of production and, because all customer interactions take place online (Vistaprint currently processes about 54,000 orders each day), performance and availability are critical IT issues.

With such extensive, online interactions taking place, Vistaprint’s technology team is able to capture and analyze huge amounts of valuable data. The company has developed extensive business intelligence and decision support capabilities and relies heavily on this information when making strategic business and IT decisions. Its development team continually monitors and regularly reports on data that draws correlations between IT and business performance. Modeling exercises typically precede any major changes to the IT infrastructure prompted by new business initiatives.

“We built a strong analytics and business intelligence department that allows us to do analysis of all of the different types of changes that we make. That helps us understand the business impact of IT changes and what that means to the bottom line.”

— Jim Sokoloff,
VP of Technology Operations
Vistaprint

As the business grew from its beginnings in 1995, Vistaprint’s back-end storage architecture evolved through several phases, moving from a NAS environment in 2004 to a high-performance SAN built on modular midrange storage. Business expansion and tremendous growth fueled the deployment of content addressable storage (CAS) and, ultimately, drove Vistaprint’s current storage initiative: deployment of a private cloud architecture based on EMC Atmos.

Leveraging IT to Tackle New Markets

Meeting New Business Needs

Through 2006, Vistaprint's primary business came from offering very small businesses, or micro businesses, customized printed materials such as business cards, letterhead, and postcards. In a typical use case, the customer would utilize Vistaprint's proprietary, online design tools and templates to design "business identity" materials to be printed, most often after having uploaded a logo or small digital photo for use in the design. When the customer ordered printed materials, the uploaded files were moved to remote production facilities, printed, and shipped. Vistaprint charged customers a small fee for each upload and allowed them to maintain a library of these typically small digital images on its storage infrastructure.

A sea change occurred for Vistaprint's IT group in 2007 when the company expanded their presence in the consumer market, offering a range of highly personalized, professionally printed products such as wall and desk calendars, greeting cards and announcements, tote bags, T-shirts, and coffee mugs at affordable prices. The consumer products took off, resulting in an explosion of new upload volume—and a series of new challenges for the storage team.

The most immediate and obvious need was to scale the storage infrastructure to accommodate the flood of new customer uploads. Where micro business products typically entailed customers uploading relatively small image files (such as a company logo) for use in business cards, letterhead, and the like, consumer products tended to be very photo-centric, involving much larger JPEG uploads. Customers ordering the popular calendar products needed to upload 13 or more of these large files.

The consumer business also proved to be highly seasonal as people flocked to Vistaprint's websites to order greeting cards, calendars and holiday gifts. The seasonal fluctuations created highly variable rates of storage demand over the course of a year; to cover anticipated capacity demands, the team would need to provision enough storage headroom to anticipate seasonal spikes of up to 5x or more compared to the year-round average upload volume.

Another critical development intensified the search for a new storage solution: Vistaprint's product and marketing groups had begun modeling adjustments to pricing and revenue mix on the consumer products, seeking a more attractive value proposition for consumers. Micro business customers were willing to pay the \$5.00 to upload their logo and then re-use it on a multitude of printed products. A consumer ordering a wall calendar, however, will upload 13 different photos; charging the fee on each upload increases the price well beyond what consumers are willing to pay. As the necessity of the new pricing models became more apparent, marketing intensified the pressure on IT to come up with a storage solution that could accommodate changes in the pricing model. It was becoming clear that continuing to scale the existing SAN environment would simply be too costly to support the evolution of Vistaprint's business.

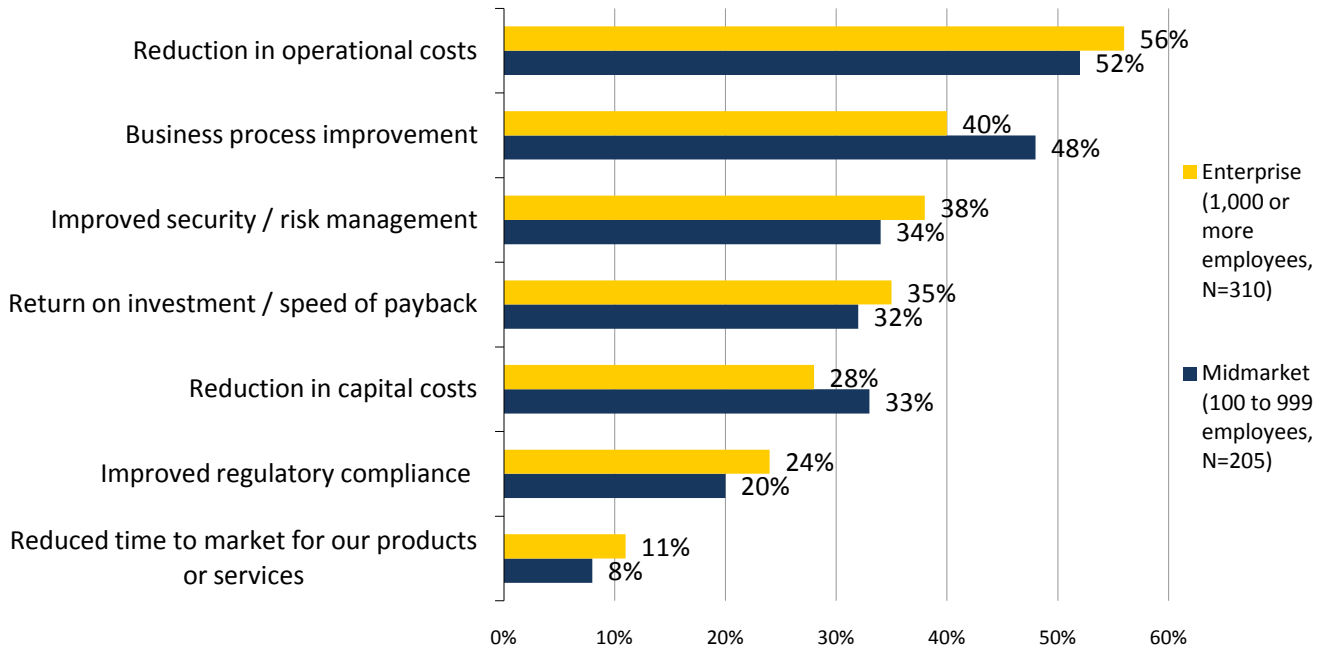
Cost reduction is a key goal for most IT architects today (see Figure 1) and Vistaprint is no different. Its IT architects deployed a new storage strategy in early 2008 to drive costs out of the storage infrastructure. They created a content addressable storage (CAS) environment to serve as a deep archive repository on the back end of the SAN infrastructure. Based on EMC Centera, Vistaprint's CAS solution provided a cost-effective object store for retaining customer uploads over long periods.

"We handle over 50,000 transactions a day, with an average value of nearly \$40 each: that's two million dollars. We can't afford even a slowdown, never mind an outage. Speeding response times has a big impact on bottom line results and converting visits into orders."

—Jim Sokoloff,
VP of Technology Operations
Vistaprint

Figure 1. Reduction in Operational Costs is Key to Justifying IT Spending

Which of the following considerations do you believe will be most important in justifying IT investments to your organization’s business management team over the next 12-18 months? (Percent of respondents, three responses accepted)



Source: Enterprise Strategy Group, 2010

While the CAS strategy went a long way toward solving the company’s immediate upload storage challenges, Vistaprint’s growth showed no signs of slowing. The IT group’s site optimization initiative provided more incentive to develop a long-term next-generation storage strategy, one that helped satisfy the very close correlation between revenue and the speed of Vistaprint’s websites in facilitating customer interaction and transactions—speed means money and high performance drives higher profits.

The team determined that a cloud storage strategy was the best option. IT evaluated both public and private storage cloud options and ran tests on a public, subscription-based model. After projecting Vistaprint’s growth and future capacity requirements and weighing the costs and benefits of the public cloud route against building a private storage cloud, the private option won out. While the subscription-based, public cloud service was attractive on a small scale, the costs became prohibitive at Vistaprint’s current and projected growth levels.

The EMC Atmos Solution: Reducing Costs with a Private Cloud Storage Infrastructure

Facing explosive growth and the need to accommodate evolving business strategies, the IT team developed a new storage strategy that would provide a cost-effective, massively scalable solution that also delivered the performance, reliability, and flexibility required to support Vistaprint’s evolving business requirements. The solution needed to meet Vistaprint’s three core IT initiatives: deliver continuous availability, reduce overall cost of the storage infrastructure, and optimize the user experience. In the spring of 2009, the IT team began deploying a private cloud storage architecture utilizing EMC Atmos. They chose Atmos before EMC had announced it as generally available. The IT team determined that the benefits of a successful Atmos deployment outweighed the potential risks of an early generation offering, having developed a good deal of confidence in working with EMC on the SAN and CAS infrastructure projects. Vistaprint initially deployed Atmos in its two data centers in the Netherlands and Canada, with a total of just over one petabyte of raw capacity.

Table 1 illustrates the impact Atmos has had on Vistaprint’s initiatives.

Table 1. Atmos’ Impact on Vistaprint Core IT Initiatives

Initiative	Results
Availability	<ul style="list-style-type: none"> Achieved 100% uptime for customer uploads Automated, multi-site object replication for disaster recovery “Virtual” object store ensures continuous operations, even during hardware upgrades or migrations
Cost reduction	<ul style="list-style-type: none"> Storage savings estimated at \$1.4 million/year Bandwidth savings estimates at \$300 thousand/year
Site optimization	<ul style="list-style-type: none"> Customers upload directly to the plant nearest them, which speeds the upload process Locating uploads in the manufacturing facility reduces cycle time for work-in-process

Source: Enterprise Strategy Group, 2010.

Storage Costs Reduced by a Factor of Twelve

Vistaprint’s Atmos environment is still young, having been in operation for just over a year. But it has already calculated tremendous costs savings. Where the CAS implementation had cut storage costs by 75%, the new Atmos cloud infrastructure has reduced storage costs by *another* 66%. This means that storage costs in the Atmos cloud are one-twelfth per GB of what they were in 2006. And Sokoloff says that he expects the costs savings will only increase as the cloud infrastructure scales over time.

Harnessing the Power of the REST API for Monitoring and Analytics

The rich data associated with the REST API interface fits neatly into Vistaprint’s advanced site monitoring and analytics capabilities. Running synthetic monitoring transactions, combined with data available on each node and disk, allows Vistaprint IT to know everything that’s happening with the Atmos deployment in real time on displays in the NOC as well as in automated monitoring systems. This not only helps keep the Atmos storage cloud up and running, it helps Vistaprint increase its return on investment by providing information on the number and types of objects accessed, weeks of storage remaining, and replication queue status.

Policy-based Management of Customer Uploads

The API-integrations that can be done with Atmos also allow the storage team to implement a policy-based management structure for customer uploads. The team defined four retention policies which drew associations between the uploaded files and their customer’s position in the buying cycle. The policies ensure that uploaded files from likely buyers are always and quickly available; files unlikely to generate revenue are deleted from the system. This tiered policy structure helps free up valuable capacity and ensures that high value customers always have access to their files. The Atmos environment has also enabled Vistaprint to define policies to automate replication and disk-to-disk backup within and between its data centers.

Unlimited Scalability to Support Continuous Growth

To support the company’s growth trajectory, the storage team needed a storage infrastructure that could scale massively without disrupting operations or creating performance issues. The Atmos cloud environment provides a foundation from which capacity can be scaled almost infinitely. The architecture abstracts back-end storage resources from applications and end-users so that the scaling processes are invisible to them. The initial Atmos implementation for Vistaprint includes just over 1 PB of raw storage; the storage team is currently working with Vistaprint’s business units to determine future capacity requirements and additional Atmos deployments.

High Availability

The storage team configured the Atmos environment so that customer uploads go first by default to the data center—in either the Netherlands or Canada—that is closest to the originating web server. In the event of an outage at the target data center, client-side web logic will force the upload to go to the other. Within about thirty seconds of the ingestion of an upload, the Atmos system automatically replicates the file across production data centers to provide geographic redundancy to reduce risk in downstream operations. According to Sokoloff, Vistaprint “achieved 100% upload availability with the Atmos implementation.”

“Vistaprint achieved 100% upload availability with the Atmos implementation. Atmos has helped us exceed all of our availability targets.”

— Jim Sokoloff,
VP of Technology Operations
Vistaprint

The REST API plays a key role here as well. The object store is fully virtualized, insulating the objects from underlying hardware changes. Atmos can move an object to another disk, another node, another rack, or another site and it automatically abides by the replication policies to ensure it has the right number of copies in the right number of places. When nodes are added for capacity expansion, Atmos can rebalance transparently under the covers with no user intervention.

A Welcome Surprise: Cloud Storage Speeds Production

While Vistaprint’s initial goal in implementing Atmos was cost savings (a goal that, as stated above, was certainly achieved), the solution has also delivered some significant, secondary benefits. By locating the digital asset in the plants for higher availability, Vistaprint also gets much faster reprints. Sokoloff says “it turned out to be very well aligned with the goals of the manufacturing continuous improvement project, which is to reduce work in progress and reduce click to ship time—so essentially the time between when you've given us the payment and the time it's in a bin waiting to be shipped. We've seen some reductions there that weren't a primary goal of the deployment, but that we're pretty pleased to get.”

The Bigger Truth

For many organizations, cloud storage is a new concept and still under evaluation. But for Vistaprint, a company whose business model is based entirely on synchronous, online interactions between customers and the company's IT infrastructure, cloud storage is not only real—it was almost inevitable. Vistaprint's storage strategies and infrastructure evolved out of the necessity to meet demands that came with the company's ongoing success in an Internet-fueled business environment.

As Vistaprint's business advanced along its rapid growth track, the company's IT leadership forged a close partnership with marketing and business groups, helping them craft storage strategies that addressed the company's rapidly evolving business requirements. As a result, Vistaprint's path to the cloud from NAS and file-based storage, through a high performance SAN infrastructure and CAS object-based archiving, and finally to an EMC Atmos-fueled cloud optimized storage (COS) environment played a strategic role in the business. Not only did the cloud storage strategy dramatically reduce costs of managing critical data assets, it became a key enabling technology that helped Vistaprint enter a lucrative new market.

The core challenges for Vistaprint's storage architects emerged from the management of customer uploads, an always important but increasingly critical component of Vistaprint's business model and value proposition. In a continuous effort to provide more value to both small business and consumer customers, Vistaprint continually searched for and tested ways to drive costs out of storing, managing, and serving customer uploads while enhancing the online customer experience by providing applications with high availability, performance, and speed. And as Vistaprint's rapid growth is projected to continue, cost effective scalability has become more important than ever. The Atmos implementation, while only two holiday seasons old, is meeting Vistaprint's requirements in all these categories—and exceeding them in several.

The Atmos environment provides Vistaprint with almost limitless scalability potential, while reducing storage costs more than tenfold over previous architectures and enabling a responsive, flexible, high-performance IT environment. While it might be too soon for Vistaprint's IT group to claim that it has "arrived" at its ultimate storage solution, its deployment of the Atmos private cloud architecture has created a foundation that is poised to meet Vistaprint's storage requirements far into the future.



Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel:508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188. This ESG publication was developed with the assistance and funding of EMC.