Backup and Recovery: The Open Source Option Comes of Age

May 2008

Adapted from *Worldwide Data Protection and Recovery Software 2007–2011 Forecast: The Focus on Compressing Recovery Requirements* by Laura DuBois, IDC #206199

Sponsored by Zmanda

Businesses are generating more digital data than at any time in history. With limited budgets and the same or reduced staff headcount, organizations are looking to automate and simplify their complex data backup and recovery infrastructure and processes. As businesses continue to add storage capacity at record rates, they are also reevaluating data protection solutions to identify additional opportunities to increase data security, operational efficiency, and cost savings. This paper examines the need for data protection in today's business environment and looks at the role of Zmanda as an open source solutions provider in the midsize and enterprise market spaces.

The Ever-Increasing Demands for Backup and Recovery

There has been an explosive growth in the amount of business data. According to IDC's *Digital Universe* study, the amount of data stored has grown fiftyfold in the past three years. Business is turning more and more data into useful information, and that trend, combined with increased regulatory and legal accountability, has compounded the problem of how to store, manage, archive, and ensure the safety and security of all this data. Moreover, the window for backing up data has shrunk; for example, an immediate audit request for email records puts incredible demands on IT staff and backup systems.

Regulatory compliance has become a critical issue for most businesses, and in many instances, it transforms data into a strategic corporate asset. Electronic medical records (EMRs), for example, are a strong growth area in healthcare, driving the need for more and better storage and data security. The Health Insurance Portability and Accountability Act (HIPAA) mandates that healthcare providers and insurers comply with security and confidentiality regulations for electronic health information. In addition, public companies are concentrating much of their Sarbanes-Oxley compliance spending on business audits that include risk-gap analysis of IT processes in data protection. Meanwhile, some 25 states now have laws in place governing computer and network security.

Data Protection: Every Company's Number 1 Job

Information and its subset, data, are a business' most precious assets. Data protection issues are a concern for all businesses, regardless of size. By its very nature, data is difficult to organize and manage, especially when in digital format. In the enterprise, data is distributed on a variety of storage devices, servers, endpoint machines, and removable storage devices and across far-flung networks.

For small and medium-sized businesses (SMBs), many, if not most, of the previously cited issues apply, with the exception that the SMB may lack the dedicated IT resources of the large enterprise. If the SMB has only one or two IT people, they are often either power users or generalists charged with managing a broad range of storage and hardware devices, software, and networking. Many SMBs simply cannot afford expensive commercial backup solutions, nor do they have the IT resources needed to manage these often complex solutions.

Yet with potential data loss threats from equipment or system failure, a natural disaster, user error, lost or stolen endpoint devices, cybercriminal intrusions, and more, the risk could not be greater, regardless of company size. Some of the most important long-term risks frequently cited in IDC surveys are the following:

■ Inconsistent backup policies for different applications and different locations

■ Lack of established IT processes to ensure successful completion of backups

■ Little or infrequent testing to ensure that backed-up data and applications are recoverable in a timely fashion

Simply put, backup and recovery is not an optional behavior for IT. With limited human and financial resources, companies, regardless of size, must implement data protection policies that best suit their needs while mitigating risks. Based on recent IDC research, IT and business management have the following key data protection requirements:

■ **Distributed data protection.** With the increase in distributed data, either at endpoints (e.g., notebooks, mobile devices, thumb drives) or at "edge" remote office/branch office locations, there is an ever-increasing requirement to protect data and automate the backup process while making it transparent to users and easy for technical staff to manage.

■ **Improving reliability.** More reliable backups are needed to ensure business continuity and fast recovery in case of disaster and mishaps. Increasingly, this means using disk storage or online storage in addition to tape cartridges, since disks and storage clouds such as Amazon S3 (Simple Storage Service) are more reliable and less subject to failure.

■ **Data security.** Security for "data at rest" and "data in transit" is needed on both the client side and the server side using encryption technologies. There is also increased demand for layered security measures, including role-based authorization and authentication to access backup data.

■ **Scalability.** SMBs require disaster recovery and expanded storage capacity; enterprise-level companies have these same requirements, as well as iron-clad data protection and recovery systems. SMBs may use the same systems as enterprises, due to flexibility and ease of use across multiple platforms, but with some unneeded functionality stripped out.

■ **Total cost of ownership (TCO).** Organizations are taking a TCO approach in evaluating data protection solutions and looking at software, hardware, and people costs associated with operating these solutions.

The Unique Requirements of SMBs

Midsize businesses, although smaller in scale than larger enterprises, have similar data protection complexity and management challenges. In many ways, tech-savvy SMBs are leading the data protection revolution by demanding data protection solutions that address the following unique requirements of SMBs:

- **Easy management.** The backup and recovery solution must be easy to administer and provide tools that make monitoring the backup process a minimal task. A well-designed Web-based interface can greatly increase usability. In other words, the interface should be turnkey and intuitively simple to operate.

- **Flexible and scalable.** The backup and recovery solution must grow with the organization to gracefully manage heterogeneous operating systems, databases and applications, and media options such as online storage without sacrificing reliability and performance.

- **Open formats.** Locking up data in proprietary formats comes with huge costs. Any data backed up using a proprietary media format can be restored using only the original backup application, effectively imposing a vendor-specific lock-in. In addition, proprietary data formats reduce the value of archives by block access by third-party regulatory or ediscovery applications. Open data formats support restores using native tools and utilities such as tar, rsync, cpio, and dump on Linux and Unix systems and ntbackup on Windows. Furthermore, open data formats allow easy migration of data from one platform to another by simply restoring it to a new target platform.

- **Lower cost of ownership.** Low cost applies not only to the initial purchase price but also to the costs of installation, configuration, integration, and ongoing administration. Simplicity also lowers labor and operational costs that must be considered in the context of TCO.

The Case for Better Backup and Recovery

Enterprise IT understands the need to deploy the best backup and recovery resources possible. And while enterprise systems and needs often change over time, existing solutions ought to be reviewed periodically and newer alternatives evaluated. This is particularly important in light of the following trends:

- **Workforce mobility.** More and more IT professionals are conducting administrative tasks remotely, including backup and recovery.

- **Centralized management.** The need to control the configuration, execution, monitoring, and reporting for distributed backup processes — including operating systems, databases, and applications — from a single touchpoint has led to the growing popularity of the "management console" concept.

- **Growth of open source use.** Open systems and adherence to standards are being embraced by many IT organizations, and SMBs are following suit. Openness and standards have paved the way for solutions that can read and manage data in de facto operating systems and common applications of the day, thus enabling data portability and future readiness. Perhaps most important, businesses are increasingly resisting being locked into proprietary data formats.

- **Migration to Web-based solutions.** Following the trend in software-as-a-service applications, many businesses are considering storage as a service (SaaS) to complement or augment the use of onsite backup.

For SMBs, whoever makes the acquisition recommendations and is responsible for keeping the system running smoothly will likely be more a generalist than a storage specialist. IDC believes that, given the challenges and priorities associated with improving data availability, application recovery times, and business continuance, as well as the mission-critical responsibilities to respond to regulatory audits and discovery requests, there is a distinct and ongoing need for data protection, disaster recovery, and archiving. The primary drivers are ease of use, faster backups, and more reliable media.

This makes open source data protection solutions particularly desirable for midsize companies, which generally have less investment in tape and tape automation and are more likely to have fewer concerns about supporting legacy technology. However, solutions need to be offered as a turnkey approach, at the right capacity and price points, requiring minimum levels of training, installation, and administration time.

In addition, SMBs are more frequently turning to offsite storage as an integral part of their backup application. Online storage supports both onsite media and offsite storage, providing backup for target data, applications, and archival data to ensure compliance and disaster recovery requirements are met. SaaS delivers a highly scalable, reliable, and low-cost storage resource that integrates easily and transparently with the company's existing databases and applications, even for hot online backup of the data. A primary concern is ensuring data security and integrity during transmission and while in storage. Encrypting data in motion and data at rest effectively addresses these concerns.

Considering Zmanda

Zmanda, headquartered in Sunnyvale, California, offers commercial open source backup and recovery solutions for Linux, Unix, Windows, and Mac systems as well as popular databases and applications. In addition to championing Amanda, its popular open source data backup and recovery software, Zmanda has a strong commitment to the open source community, contributing development and testing resources to open source data protection projects that share its vision and providing hardware, collaboration tools, and financial assistance to key open source developers.

Zmanda has grown rapidly, with more than 500,000 current active installations of its backup and recovery software worldwide. Its subscriber base increased 900% from 2006 to 2007. The company has customers in 23 countries, and its subscriber list includes up-and-coming Web 2.0 companies, Fortune 500 enterprises, and SMBs. Zmanda is an integral part of the LAMP (Linux, Apache, MySQL and Python/Perl/PHP) open source ecosystem, and the leading open source vendors, including Sun Microsystems, MySQL, and Red Hat, resell and comarket Zmanda solutions.

Zmanda offers two product lines: Amanda Enterprise and Zmanda Recovery Manager (ZRM) for MySQL. Amanda Enterprise is a network solution that enables the backup and recovery of business-critical data associated with servers, workstations, desktops, and databases and business-critical large-scale applications, such as Microsoft Exchange, across the entire network. ZRM for MySQL is the first advanced backup and recovery solution specifically designed for the MySQL database and provides continuous data protection (CDP) capabilities including hot backup and fast point-in-time recovery.

Amanda Enterprise

Amanda Enterprise allows users to set up a single master backup server to back up multiple Linux, Unix, Windows, and Mac clients to a large selection of tape, disk, online storage (e.g., Amazon S3), and optical devices including tape libraries, auto changers, optical jukeboxes, RAID arrays, NAS devices, and many others. Amanda Enterprise supports both client-side and server-side compression and security options and is certified for use by the U.S. Department of Homeland Security; it is the only open source data protection product to have achieved such a distinction.

Amanda Enterprise has an intelligent scheduler designed to automatically determine backup levels based on load balancing, enabling users to optimize their storage resource requirements and utilization. This is a key differentiator from other commercial backup software that requires users to specify which backup levels are to be performed and on which specific days to perform these backup levels. For organizations with a large number of backup clients to configure, this can be a tedious and complicated task, not to mention the purchase of a lot of additional storage capacity that remains underutilized except for peak conditions.

The biggest advantage of Amanda Enterprise, according to Zmanda, is that the solution does not use any proprietary data formats or tools when backing up data. Amanda Enterprise uses tar, dump, or Schily tar as backup tools, which are readily available, industry-standard tools. Other commercial backup products use their own proprietary backup algorithms, tools, and data layouts, which lock users into the product, since it provides the only way to recover data.

Other noteworthy Amanda Enterprise features include the following:

- **Easy-to-use Web-based management console.** This console simplifies daily activities for backup administrators who can remotely access Amanda Enterprise from any Internet-enabled PDA or phone to perform backup and recovery operations.

- **Scalable and heterogeneous.** A single backup server can back up and recover hundreds of file systems.

- **Layered security.** An encryption option is available on the client side and the server side, plus role-based access and user authentication to back up data.

- **Low cost of ownership.** Amanda Enterprise, depending on configuration, costs 50% to 80% less than other commercial products. Users pay an annual subscription for each system protected regardless of data volume.

- **Technical support.** Users may choose from three levels of support, complete with knowledge base access, email support, phone support, case management, and software enhancement updates.

Zmanda Recovery Manager for MySQL

Zmanda developed ZRM for MySQL with the explicit goal of providing simple, secure, and consistent backups with minimal impact on applications and users. With its recent 2.2 release, Zmanda is now reportedly the only vendor to provide CDP for MySQL databases. Zmanda has gained significant traction within the MySQL database administrator (DBA) community and was recently recognized as the "MySQL Partner of the Year" at the April 2008 MySQL User Conference. With ZRM for MySQL, DBAs can deploy MySQL databases in mission-critical, high-transaction environments with confidence because they know their data is protected by Zmanda. True to its open source mission, ZRM doesn't use any proprietary data formats. Users can always recover their MySQL data even without ZRM being installed.

With ZRM for MySQL, DBAs can do the following:

- **Perform flexible backups.** DBAs can easily schedule full, incremental, raw, or logical backups and use snapshot technologies to perform hot backup or activate CDP to enable instantaneous point-in-time recovery.

- **Back up Live MySQL databases without impacting applications and users.** Background backup is accomplished by taking advantage of Linux LVM, Microsoft VSS, Symantec VxFS, Solaris ZFS, and NetApp SnapManager technologies.

- **Perform one-click point-in-time recovery.** DBAs can easily identify the recovery point and activate database recovery right from the Visual Log Analyzer to recover the whole MySQL server, database, table, or transaction.

- **Centralize global backup management.** DBAs can automate backups across local, remote, or clustered MySQL servers via an easy-to-use Web console and back up MySQL servers with MyISAM, InnoDB, NDB, and other storage engines on Linux, Solaris, Windows, and Mac.

- **Maximize data security.** DBAs can encrypt data at rest and data in transit using industry-standard encryption technologies. Role-based access control provides added layers of security.

- **Leverage power reporting.** DBAs can stay informed via automated alerts and reports and easily fulfill audit or compliance requirements using highly configurable reports.

- **Easily deploy and manage backup and recovery operations.** ZRM is quick and easy to install. Similar to Amanda Enterprise users, ZRM users may choose from three levels of support that best fit their requirements.

Since no two environments are the same, additional services available from Zmanda include:

- **QuickStart.** Implementation is done by a Zmanda professional services team, which works closely with Amanda core developers.

- **Implementation Services.** Zmanda professional services engineers design a custom backup and recovery solution and assist in implementation.

- **Solution Services.** Turnkey solutions include backup migration design, industry-specific data protection, and, when necessary, bare metal recovery.

- **Continuity and Compliance Consulting.** Additional solutions include disaster recovery readiness, business continuity compliance, and global storage management.

Challenges

Zmanda finds itself in a large but rapidly consolidating and increasingly competitive marketplace. The company has an established reputation in the midmarket and is strongly positioned in the open source market segment, as well as the entrenched commercial solutions segment. This enables Zmanda to be a single source for solutions. With Sun, MySQL, and Red Hat actively reselling and comarketing its products, Zmanda is well-positioned to outpace industry growth through its strong multichannel go-to-market strategy. With the stated mission to "make it easy and affordable for people to back up and recover data in an increasingly complex and heterogeneous IT environment," Zmanda is addressing a real market need for simple and cost-effective midlevel backup and recovery solutions.

Zmanda has demonstrated impressive success since its inception, especially given the fact that the company directed its initial product offering at SMBs rather than taking the more conventional approach of beginning at the enterprise level. The company has succeeded in the traditionally underserved midmarket; now it must demonstrate prowess in the highly competitive enterprise market segment. One challenge will be to maintain a balance of support for both midmarket and enterprise solutions.

Further, Zmanda must continue to lead with innovation and customer satisfaction. This means aggressively marketing the Amanda Enterprise solution as more feature-rich than its predecessor. As it expands and enhances Amanda Enterprise's competitive positioning, Zmanda would be well-advised to give enterprise IT the same attention to detail in customer support and professional service offerings that it gave to SMBs. Zmanda will be well-served to continue its innovative edge in the areas of SaaS and MySQL backup and recovery.

Zmanda has garnered considerable credibility by receiving the Homeland Security endorsement, and the company should continue to enrich and expand its security feature set, given that computer systems security is the second most critical concern for business.

Conclusion

Midsize firms face many of the same pressures as large firms when it comes to storing and managing electronic data. Routine business and regulatory compliance requirements are increasing the quantity of data managed and raising its value to the organization. Regardless of company size, data protection priorities include the following:

■ Enabling centralized protection and recovery of distributed data

■ Improving recovery time objectives (RTOs) and recovery point objectives (RPOs) to minimize data loss while optimizing recovery in the event of logical or physical errors

■ Providing secure information retention and retrieval from backup or redundant data sources in the event of audits or ediscovery, or for general business use

Midsize firms often face greater challenges than larger firms, especially with respect to their limited resources and lack of IT expertise. There is a demonstrated need for data protection solutions tailored for midsize firms that provide enterprise-level capabilities, coupled with ease of implementation and management. Zmanda recognized this need early on. Its product and service offerings for both SMB and enterprise markets reflect the company's thoughtful approach, based on an open source foundation that widens their usefulness and appeal.

Open source storage solutions have been gaining momentum over the past five years. The trend is away from tape backup and toward fixed disks for backup and archiving, in part because fixed disks provide better support for RTOs and RPOs. These trends accentuate the need for a thoughtful backup and recovery software solution. Zmanda can expect its two complementary backup and recovery solutions to grow in concert with the need for more storage — regardless of the media — so long as it maintains a competitive edge by continuously innovating and refining its offerings.

IDC expects the midtier segment of the data protection market to continue to grow, and Zmanda is well-positioned in this market. Entering the enterprise market poses a greater challenge, but if Zmanda can address the challenges described in this paper, the company has a significant opportunity for success.