

ANALYZE

OPTIMIZE

SECURE

## From the backup window to the glass atrium

A new definition for the intelligent management of data

A discussion document published by BridgeHead Software | November 2007



# Contents

|  |    |
|--|----|
| Introduction                                 | 3  |
| Managing the data mountain                   | 4  |
| Counting the hidden costs of primary storage | 6  |
| Compliance: carrot or stick                  | 8  |
| Building the Magisterial Residence           | 9  |
| Summary                                      | 11 |
| Sources                                      | 12 |

## Introduction

If your concept of an archive corresponds to visions of a gloomy vault full of forgotten records, think again. The true dictionary definition of the word “archive” is “magisterial residence.” It’s time to elevate the archive from a dark, forgotten place to its true position of authority as a transparent, secure, central repository for valuable data that can be prioritised, searched and accessed, on-demand, by authorised users. While IT professionals continue to invest their time, money and business security in expensive backup technology, many are overlooking the importance of archiving as a fundamental storage management tool that will lower costs, reduce business risk and improve operating performance.

In this document, we will explore the definitions of backup and archiving and unveil the key distinctions. We’ll also reveal that, when deployed in tandem, these two distinct, but complementary disciplines can help organisations achieve competitive advantage through intelligent and cost-effective management of data.

*“Archive isn’t about back-office operational efficiency (although it can contribute to operational efficiency). It is about solving business problems such as compliance, historical auditability, data non-repudiation and intellectual property management and protection. Archive, unlike either backup or HSM, has a full business value case to offer in terms of cost-savings, risk reduction and business process improvement.”*

*John Toigo: Building an Archive: End of a Six Part Series (May 22 2007)<sup>1</sup>*

## Managing the data mountain

In recent years, an increasing volume of data has made backup a real burden to storage managers. However, like the life-jacket under your seat, in a crisis situation, you'll rue the day you ignored the safety announcement.

In 2003, a study by the University of California, Berkeley, calculated that the world's information production was five exabytes with 92% data being stored on magnetic media (primarily hard disks). Based on these figures, Forrester estimates that with expansion of data increasing at a rate of 30% per year, the volume will increase to zettabyte sizes by 2010.<sup>2</sup>

It can take up to 9 hours to perform many corporate backup cycles and while the volume of data continues to grow, there's just no way to extend the after-hours window of opportunity. Once 9 hours extends to 12, you'll be eating into those all-important, productive business hours. Meanwhile, the growing demand for 24x7 business communications renders the term, "out-of-hours" redundant – there's simply no space or resources to allow for down-time.

It is estimated that up to 80% data currently held as primary storage will seldom if ever be accessed again. Yet night-after-night, across the globe, the business community continues to do the right thing by backing up data, "just-in-case" disaster strikes.

The "emergency" factor is one of the most compelling distinctions between backup and archiving. In essence, backup is your disaster-recovery plan. By taking a point-in-time copy of your data and storing it offline on disk, optical or tape media, you can prepare for malicious attack, system/hardware failure or natural disaster.

In order to ensure business continuity, backups are repeated and scheduled at regular intervals to minimise the impact of loss of production data. The problem is, that backup does nothing to manage the volume of data held on primary storage. As you acquire more data, you simply increase the pressure on the back-up window, creating a monster with an insatiable appetite for growth.

By contrast, an archive transfers important but rarely accessed data, captured in its final form, away from primary storage to fixed or removable media. This takes the pressure off the back-up window, reducing the burden of managing high volumes of data on primary storage.

*“Archive is not backup, nor is it hierarchical storage management (HSM), though it has some elements in common with both. Archive is a more granular technique for selecting and moving specific data to long-term storage in accordance with business policy and in response to business requirements.”*

*John Toigo: Building an Archive: End of a Six Part Series<sup>3</sup>*

By nature, archive data should be prioritised in terms of its long-term value and required frequency of access for future scrutiny. For the purposes of compliance, it may be protected from modification and stored for long periods of time. By selecting appropriate storage media, it is possible to specify the seek-speed for access and longevity of retention. Advances in technology mean that archive data can be stored securely and intelligently using sophisticated search features that can quickly and efficiently identify files to make short work of e-discovery initiatives.

## Counting the hidden costs of primary storage

The fear of losing valuable data has prompted IT Managers to rely on universal backup of data as their insurance policy. It is commonly accepted that the declining cost of primary storage media has masked the through-life costs of maintaining high volumes of data.

Until recently, purchasing cheap, additional storage, seemed the easy option. Archiving was centred on point solutions that appeared to increase the complexity of managing data. The sheer effort of classifying and selecting appropriate data for archiving across multiple data types, platforms and media deterred storage managers from attempting to migrate primary data to lower-cost storage options.

Now, with growing concerns about the rising cost of fuel and uncertainty regarding the security of energy supplies needed to power vast data centres, storage managers are beginning to take a different perspective on archiving data. There are grim warnings from industry experts:

*“..by 2008 50 percent of current data centers will have insufficient power and cooling capacity to meet the demands of high-density equipment”*

*Gartner Press Release November 29 2006*

*“...within five years, the cost to power and cool a server cabinet over its three-year projected life could rise from the current \$206,000 to as much as \$2.3 million. That's anywhere from 300 percent to 2,250 percent of the equipment purchase price.”*

*Rising Energy Costs Reduce Processor Performance Gains, CIO Magazine, March 1 2007*

The through-life costs of following a universal backup policy in which most of your data is held on primary storage are observed not only in terms of energy consumption but also in the increased investment required for specialist IT resources to run extended, high-volume backup cycles, locate individual lost files or respond efficiently to e-discovery initiatives.

Another hidden cost of clinging to a universal backup policy could be the painful responsibility of defending an expensive legal action for failing to adhere to tightening international business requirements for compliance and corporate governance. While the UK's new Civil Procedures Rules (CPR) which published guidance for e-disclosure may not generate the high-anxiety of Sarbanes-Oxley, a recent survey of 100 litigators in 22 leading UK-based law firms conducted by KPMG Forensic revealed significant levels of expense and confusion. According to KPMG's survey, nearly half (48 percent) of e-disclosure cases cost £500,000 or more, with over a quarter (26 percent) costing more than £1 million.

*"E-disclosure can be immensely complex, costly and challenging, and litigators have called for renewed energy in agreeing clearer case management guidelines. Many of them also clearly believe that some training for judges could be beneficial.*

*Technology may have created the problem, but technology can also be part of the solution. Lawyers need to be aware of the tools that exist to help them manage the information load, as well as devise review strategies to make the task easier."*

*Paul Tombleson, Head of Forensic Technology at KPMG Forensic4*

## Compliance: carrot or stick?

Depending on whether you see the glass half-empty or half-full, compliance is either an additional burden on the overworked IT department or a valuable opportunity to set your house in order. Whatever your view, compliance is driving more Storage Managers towards intelligent archiving as a tool to automate the e-discovery process with increased control over retention, authentication, security and access control.

Organisations are under increased pressure to hold data securely and efficiently for extended periods of time. A best-in-class archive solution will automate the responsibility of maintaining records by keeping data encrypted and tamper-proof and providing an audit trail of access history, controls and custody. The new BridgeHead Enterprise Data Archiving Platform provides peace of mind with built-in disaster recovery by writing multiple archive copies to multiple locations on different media to deliver new options for expensive, routine backup and replication of archive data.

According to The BridgeHead Software ILM Audit 2007, compliance has jumped into second place, overtaking data growth as a key driver for archiving on both sides of the Atlantic. Since 2005, the significance of compliance has risen from just 15% in 2005, to over 56% in 2006 and this year it is up to xx% [insert combined US/UK figures from Exec Summary 2007]

In spite of these encouraging figures, in the UK 16% IT Managers surveyed in the BridgeHead Software ILM Audit 2007 never archive their data. Mindful of the tenet: “no pain, no gain,” the onerous task of selecting and prioritising data may have deterred Storage Managers from replacing universal back-up with a combined backup and archiving strategy. Those who continue to ignore the drum-beat of compliance may discover, to their cost, that the pain of meeting the requirements a legal e-discovery action, overshadow the effort of classifying data to build an enterprise-level archiving solution.

## Building the Magisterial Residence

Data is one of your most valuable assets. It contains vital information that can map the future prosperity of your business, documenting achievements and highlighting bad investments, so you can make strategic business decisions, based on the hard lessons of experience. As you work towards building a magisterial residence for your archive, think in terms of replacing the dim vault with a central, glass atrium, protected by robust security screening that welcomes authorised visitors but denies access to unwanted guests.

Transparency is the key to an efficient and productive archiving system. In order to work effectively, archive data must be readily available to authorised users. Users and applications should be unaware that data is being accessed from a more secure and lower-cost environment. Selection of appropriate data for archiving is a primary concern. Universal archiving of data is as misguided as universal backup.

*“The devil is always in the details, isn’t it? When you introduce automated archiving, you need to understand what’s going to happen when you flick the switch. Simplistic programs that just begin to madly find and write data off to the archive can be dangerous.”*

*Patrick Dowling, senior vice president of marketing, BridgeHead Software<sup>5</sup>*

Dynamic data, subject to constant review and modification and should remain as primary data, protected by routine backup. Archives should contain fixed-content files, such as email, legal documents, medical images or video footage. An intelligent archive system will allow you to prioritise data and set retention policies at file level.

While setting up an effective, archive system that is transparent to users, will require careful upfront analysis of data and definition of workload, the advent of new, rules based automated archiving engines is helping to simplify the process. A truly enterprise-level archiving solution will avoid the duplication of effort associated with point archiving solutions by managing all activities across all types of data, storage media and physical locations.

Once user-rules are defined, the system will assign data to the appropriate storage media and provide an end-to-end management policy for maintenance, offsite rotation, migration and retirement, including secure shredding of data at end-of-life. Advanced features such as built-in analysis, simulation and reporting, enable Storage Managers to gain a more detailed understanding of their data environment and to test the impact of rules before they are applied live on the system.

*“Archiving at the enterprise level is where the solution to out-of-control data growth will be resolved. Archiving initiatives are currently very fragmented across most organisations, limited to addressing specific data types such as e-mail or driven by departmental groups outside of IT focused on a particular piece of the puzzle, such as financial compliance. By bringing archiving together through a core enterprise system as we’re trying to do, you start to enjoy a real critical mass in terms of the data of the archive. You also create management efficiencies by streamlining and simplifying processes, cutting out duplication and improving control.”*

*Tony Cotterill, CEO and President, BridgeHead Software<sup>6</sup>*

## Summary

In this document, we have reviewed the distinctions between backup and archiving data to establish the discrete roles of these two disciplines in addressing the challenges of out-of-control data growth:

- Backup is ideal for copying dynamic data stored on primary systems
- Archiving involves moving fixed data away from primary storage onto lower-cost media
- Backup copies are stored offline whereas an archive stores data online or near-line so that it is transparent and readily available to users
- Backup involves copying all files held on primary systems, whereas intelligent archiving involves classifying data by setting retention policies at file level

The expanding data mountain will shatter the backup window if organisations continue to retain the bulk of their data on primary storage. Storage Managers must take action to manage the hidden, through-life costs of investing in primary storage and the need to prepare for compliance.

Exciting, new advances in archive technology that deliver sophisticated search, retrieve and built-in disaster recovery capabilities, empower organisations with increased control and intelligence about the content, location and status of their data.

The emerging adoption of archiving does not mean that routine backup of primary data is now redundant. When applied as complementary disciplines, intelligent archiving of fixed data and routine backup of dynamic files can enhance productivity to achieve long-term competitive advantage. As Charles Garry, independent industry analyst concludes:

*“Through the practice of active archiving only the most important, most relevant, most frequently accessed data remains in the application’s database. By doing so, the application’s performance is maintained while keeping infrastructure and personnel costs down. In addition, end users become more productive, customer satisfaction goes up, and data in the application is better protected.”*

*Charles Garry, Independent Industry Analyst<sup>7</sup>*

## Sources

- 1 Building an Archive: End of a Six Part Series John Toigo: Enterprise Systems (May 22 2007)
- 2 Data, Data Everywhere! Boris Evelson, Forrester (July 23 2007)
- 3 Building an Archive: End of a Six Part Series John Toigo: Enterprise Systems (May 22 2007)
- 4 Lawyers call for clarity on 3-disclosure – and training for Judges and masters: Press Release KPMG Forensic (October 10 2007)
- 5 Building an Archive Strategy, Part V: The Devil in the Details, John Toigo: Enterprise Systems (May 15 2007)
- 6 IT Must Take Back Control of Archiving Throughout the Enterprise: BridgeHead Software Press Release, Storage Expo (October 17 2007)
- 7 Active Archiving – A Key Enterprise ILM Practice, Charles Garry, Independent Industry Analyst, Storage Networking Solutions (SNS) Europe (March 1 2007)