dolphin

Smart Adaptable Proven™

DATA ARCHIVING

The first Step toward Managing the Information Lifecycle

Best practices for SAP® ILM to improve performance, compliance and cost

Table of Contents

EXECUTIVE SUMMARY

Where data archiving is about performance, Information Lifecycle Management is about compliance. The distinction isn't a break with the past – it's an evolution.

Data archiving strategies are designed to improve performance and manage costs, moving static data from the online database to offline archives. Information Lifecycle Management (ILM) adds a new level to the performance-cost dynamic: manage data in compliance with outside and corporate retention rules and business requirements; maintain ready access to data and documents; protect the organization from legal and business risk.

SAP calls Information Lifecycle Management the combination of "policies, processes, practices, and technologies used to align the business value of information with the most appropriate and effective IT infrastructure from the time information is conceived through its final disposition." At its core, ILM is the practice of thinking and acting strategically about how data is managed in the organization: where it is stored, how quickly can it be accessed, how it is tracked, and how long it is retained.

The rising adoption of information lifecycle management for the mass of documents and records generated by a business is being driven by a range of business priorities, primarily:

1. **Compliance:** Ensuring regulatory compliance through management of the information lifecycle, from creation to record retention and finally, destruction.

2. **Performance:** Online database growth slows enterprise network and database performance. ILM strategies create a process for moving static 'business complete' records from the database to the archive -- freeing online disk space while maintaining seamless archive access for users.

3. **Cost:** SAP users typically experience database growth at 20%-30% per year. Data archiving strategies, in concert with ILM practices, move data from online storage to lower cost archival systems, slowing the growth of infrastructure maintenance and reducing total cost of ownership for data management.

4. **Preparation:** An ILM strategy enables the organization to plan for changes, including system upgrades, mergers and acquisitions, and manages challenges like legacy system decommissioning and legacy data access.

Dolphin helps organizations run crucial business operations better and smarter in SAP. The company has a history of success delivering higher performance and lower total cost of ownership by helping customers using SAP solutions plan and implement strategic ILM and data archiving programs.

This whitepaper discusses the business case for data archiving as the first step in an SAP ILM strategy; strategies for Data Archiving, Retention Management and the Retention Warehouse strategy for legacy decommissioning in SAP; and Dolphin's best-practices approach to an effective, long-term ILM strategy.

THE BUSINESS CASE FOR ILM

Among the most urgent business interests in ILM today is compliance -- and for good reason. Compliance with laws and regulations from Sarbanes-Oxley and Securities and Exchange Commission (SEC) reporting to Health Insurance Portability and Accountability Act (HIPAA) requirements for health data privacy and portability, the Gramm-Leach-Bliley Act for financial services, and the Good Manufacturing Practice Regulations (GMP) for food and drug manufacturing can impact anything from legal disputes and regulatory approvals to the very survival of the business. Moreover, each business has its own policies for retaining the records needed to transact business with its customers.

But not all information needs to be retained forever, and few businesses would want to do so. From a cost perspective, infinite retention is a drain on resources. From a liability standpoint, infinite retention increases liability risks ranging from potential data privacy breaches to legal exposure. The challenge for IT: retain what must be retained, keep it accessible, and destroy what can and should be destroyed – all while managing overall costs.

An ILM strategy begins with defining the policies that will ensure compliance with business and regulatory records retention requirements and integrates these requirements into short and long-term IT resource planning. This gives business leadership confidence that procedures are in place for information management from the time records are created, through archive retention and end-of-life, in ways that are responsible, efficient and cost effective.

The next issue – one that raises flags across the business – is **system performance**. Business operations create documents and records, and users want every record available at the speed of thought. Unfortunately, the more records available online in the production database, the bigger the database. The larger the database, the more often users are affected by slower response times for access to records and the time it takes to run reports.

Keeping records online requires expensive growth in storage hardware. IT must balance cost with service level agreements in terms of their ability to meet business requirements for data backups and recovery speed.

At Dolphin, we have found that 75 to 80 percent of the records in SAP databases are static – data that needs no further modification. Data archiving strategy defines how the business moves these 'business complete' records and documents off of the production database and onto an accessible offline archive, keeping the database free for the most current and critical data, thereby improving runtime and helping to meet service level agreements for backup and recovery.

Figure 1 illustrates how data moves through an organization in an ILM framework. As it becomes less valuable, it is moved from the Tier 1 production database to the lower cost Tier 3 archive system.

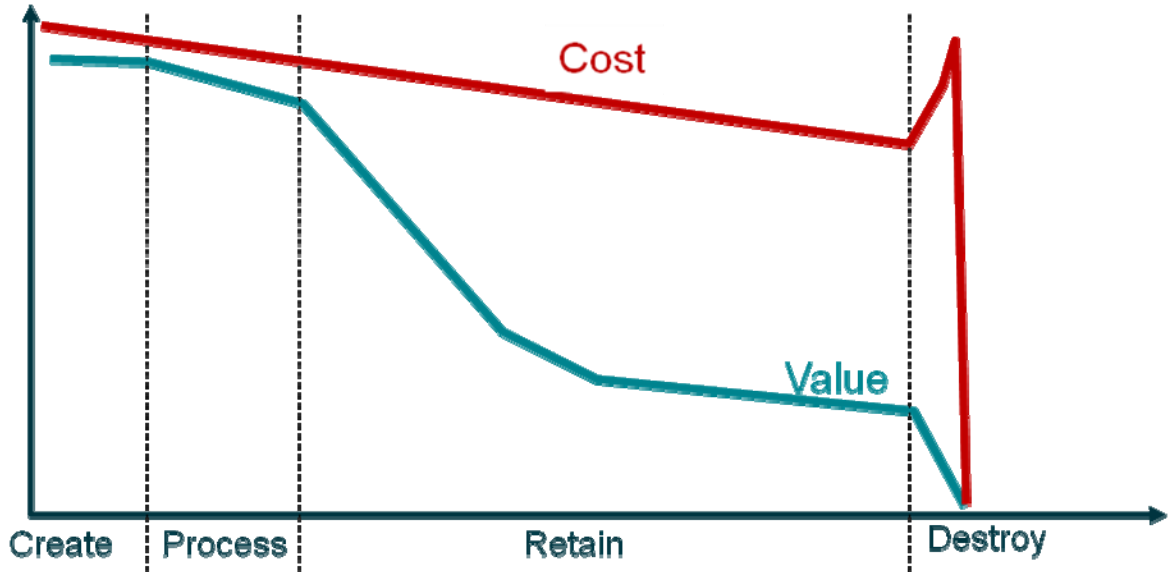

Figure 1

A related business issue driving ILM adoption is **total cost of ownership.** Companies can see a direct correlation between the size of their production database and the cost of storage. The more that data can be moved from expensive "Tier 1" disks to "Tier 3" archive storage, the lower the overall cost of disk space to the organization. Data archiving strategy in an ILM framework helps manage disk capacity growth and lower overall cost of storage.

Moreover, data archiving has a significant impact on **backup performance, database reorgs, system restores, client copies and cost.** When data is taken off of the production database, it is by default

removed from the production backup strategy; as a result, production database backups become faster and less expensive.

Companies are turning to an ILM strategy to prepare for change in their organizations. Upgrades can cause major disruptions for IT organizations and are directly impacted by database size. A smaller production database, through data archiving, simplifies this transition. Another issue is **system consolidation** – moving from multiple SAP systems to a single shared service environment or integrating an acquisition are major drivers of database growth. Smaller production databases ease and speed consolidation.

Organizations are often required to retain legacy data. Rather than running expensive legacy systems, legacy data can be archived into a single point of access so that it can be retained cost effectively and accessed easily, allowing the legacy systems to be decommissioned.

CREATING THE ILM STRATEGY

The three "cornerstones" of ILM in the SAP environment are: Data Management, Retention Management and Retention Warehousing.

SAP announced its ILM from SAP solution in 2008. This is a separate infrastructure component available for users of SAP NetWeaver 7.0 Enhancement Pack 1.0 and above that offers capabilities a customer can implement to enhance how it manages, identifies and categorizes data for each of these cornerstone areas. The new features will be particularly useful for implementing more robust retention management policies and integrating third-party content management systems (CMS) through ArchiveLink.

Data Management, or Archiving, is the first, necessary step of any SAP ILM strategy. Once there is an effective system in place for data archiving, the organization can implement **Retention Management**. The retention management process handles the lifecycle of data, ensuring that archived records are completely removed from the system and destroyed once all business and compliance retention related periods have passed. Retention management also enables business-critical functions such as **"legal holds"** that ensure that data (and only that data) that is directly related to a legal action is found and frozen during a legal proceeding. Further, archiving enables the **Retention Warehouse** – securing continued access to legacy data archives and enabling decommissioning of a legacy system.

An investment in comprehensive up-front analysis and planning accelerates and simplifies implementation. Dolphin's phased approach to effective data management starts with collaboration between the IT team that manages the data and the business people who use it. The goal: a shared definition of business priorities, access needs, and regulatory requirements.

At the end of this up-front analysis, the organization will be able to answer the following questions:

- Who needs what information; and when, where, how fast and for how long from a user standpoint?
- When is a record "business complete"?
- What are the relevant legal and regulatory data retention and data access requirements for the business?

These business policies drive ILM strategy. The resulting blueprint is your map to move forward with your data archiving implementation.

Using SAP's tools and Dolphin's applications, Dolphin analyzes business, technical, and functional needs to create a data archiving blueprint. This analysis considers both immediate requirements and long-term needs for data retention and retrieval to enable IT to manage its storage and infrastructure investment.

As a result, the IT organization can create a smarter, more flexible plan for:

- When information is archived,
- Where information is stored,
- How long information is kept, and
- Ease of access to archived information.

DATA ARCHIVING AND RETRIEVAL IN SAP

A smart data archiving strategy will manage data volumes to meet business needs and compliance requirements. Dolphin's best practice approach to managing data begins with the following principles:

- Start early. Don't wait for a crisis, user complaints or an urgent corporate initiative. Begin defining and implementing a data archiving program early to avoid performance bottlenecks and to manage growth.
- Short online residence. Ideally, an organization should archive records once they are business complete and are no longer being accessed frequently. Records should be moved to archives as soon as it is practical based on business access and regulatory needs to keep the production database lean with in-use, current business data.
- Continuous archiving. Archiving should be a continual process - not a project - to ensure high performance, data protection and that retention requirements are being met. Among other benefits, continuous archiving will facilitate implementation of data retention and end-of-life policies from the archives.

Figure 2 provides an overview of how data archiving works in SAP.

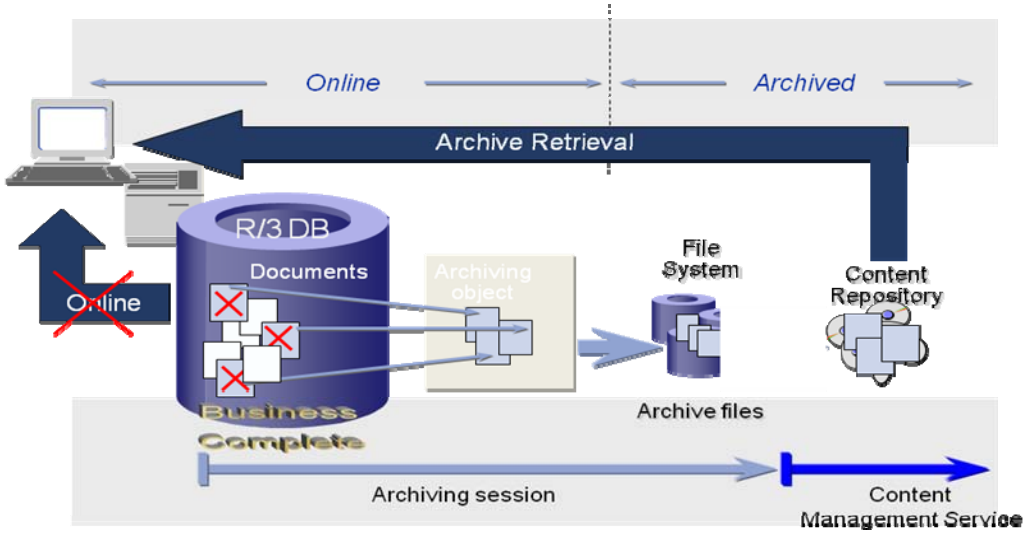

Figure 2

Records created in SAP are stored in the online production database. SAP users can access that transaction through standard transaction codes. Once the transaction is closed, it is considered 'business complete.' When the business complete transaction is no longer accessed frequently, it can be moved to archive using the Archive Development Kit (ADK). In this process, the record is removed from the online production database and saved in the archive, along with all of its fields and retention rules. If the business is using a content management system or third party storage solution, the transaction now is recorded in the content repository, along with the rules in place for data retention for each type of document. The content management system uses these rules to prevent accidental deletion of a record before it reaches its retention period.

DATA RETRIEVAL CHALLENGES

Data archiving projects often never get off the ground because users fear they will lose access to their data. IT's challenge is to assure users that they can retain transparent access to the information. While SAP has a number of archiving function modules, accessing archived data presents unique challenges. For example, once the transaction is in the archive, users cannot access it through the standard transaction codes – these transactions only seek records within the SAP database. Archived data retrieval now requires choosing from SAP function modules.

Figure 3 outlines the standard SAP function modules available for data retrieval. Most of the standard SAP tools either reduce the efficiency of the data archive by adding large index data tables to the online database; or they increase the complexity and time it takes to access data archives.

Single Document Display	• Only available in FI • Reduces archiving efficiency by > 50%
Archive-Enabled Reporting	• Small number of supported reports • Sequential access → long runtime • No archive/online integration
Archive Information System DRB	• Not integrated with standard display • Limited detail display (business view) • Requires index data in database
Print Lists	• Static report output, requires content repository • Supports search and hyperlink features • No dynamic selection, navigation, etc.

Figure 3

The best practice solution for data archive retrieval is PBS archive add ons, which offer transparent, simultaneous look up and access to offline archived data and online data.

PBS archive add ons™	• SAP add-on solution for transparent retrieval • Index information stored outside of the SAP Production database • Archive-enabled display transactions and reports

PBS archive add ons™

PBS archive add ons allow an end user to view and access archived information together with online information. Archive and online data retrieval is seamless – users see the archived data in the same way, and with the same screen layout, as they did when it was still in the database.

Figure 4 highlights the difference that PBS archive add ons make. The left side shows the result of a standard transaction query in the SAP database.

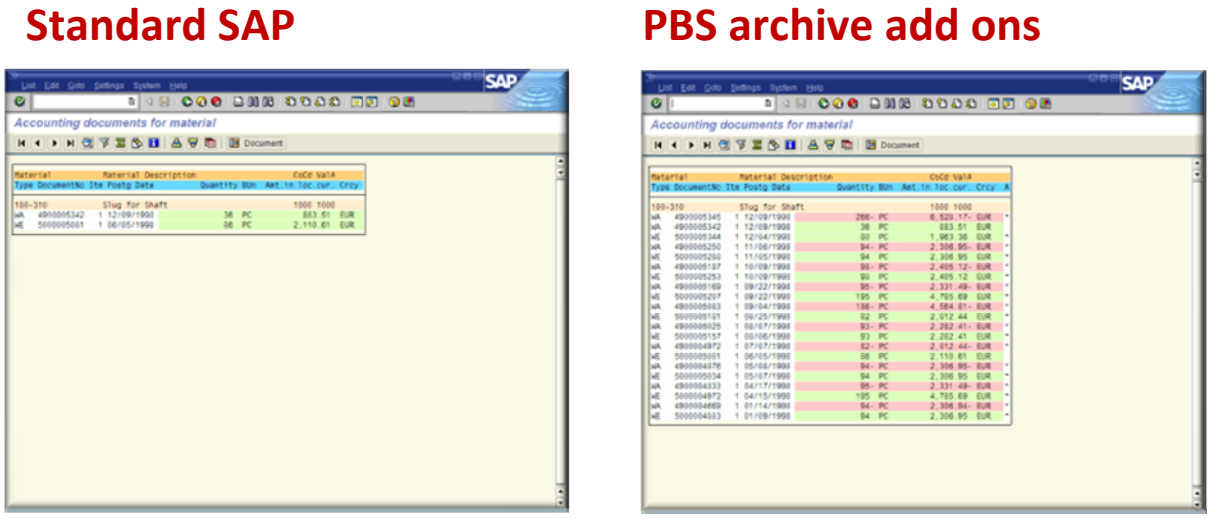

Figure 4

The standard SAP query only searches the production database. If the necessary record is in the archives, it will take approximately 10 more clicks to bring up that record. On the right side, the PBS system brings up the production and offline archived data in the same query, identifying the archived data with an asterisk. But for the user, access to the archived data is entirely transparent – no different than getting an online record.

The PBS archive add ons works by storing archived index information in a compressed format outside of the production database. This improves archive efficiency by keeping the archive's index files outside of the database. The benefits of PBS archive add ons are many:

1. **Aggressive archiving**. Because the PBS archive add ons deliver strong visibility and access into the archived data, IT can archive with confidence. This allows the organization to reduce the time data is kept online, improving database performance, reducing backup times, maintaining user data access and reducing overall storage costs.

2. **Easy implementation**. PBS archive add ons use standard SAP development tools and allow users to access archived data through familiar SAP screens. As a result, Dolphin has found that they reduce custom development efforts by 90% for even the most demanding implementations.
3. **No user training required**. Because PBS archive add ons enhance built-in SAP capabilities and use standard SAP processes to access the archived records, there is no need for user training.
4. **Easy to use**. With PBS archive add ons, users can look up and retrieve relevant transaction records – online and archived – in a single search, a process that takes nearly a dozen searches and clicks through the Archive Information System.

Using ArchiveLink Certified Content Management Systems

In the standard SAP configuration, a user request for an archive record is accessed through the Archive Information Server and then through the Archive Development Kit (ADK) to reach archive files or records in an external ArchiveLink-certified content management system (CMS). PBS archive add ons create their own index files outside of the database. As a result, a user request for a transaction simply searches the PBS index files, then accesses the archived record directly from an ArchiveLink certified CMS.

A key advantage to using an external content management system to manage your SAP archives is data protection. The CMS can enforce the rules established in business and compliance policies as to how data is stored and when it can be purged from the CMS. It also allows the business to share archive information in multiple scenarios. For example, electronic documents and print lists can be linked to SAP transaction records to provide a more complete archive of that transaction.

Archive Sessions Cockpit (ASC)

Archiving is necessary, but it can take up a great deal of staff time. The Archive Sessions Cockpit is designed to free the SAP Archive Administrator from recurring, manual archiving operations.

Available only from Dolphin in North America, the Archive Sessions Cockpit is an SAP add-on that creates a hands-off process that schedules, automates and manages archive runs.

Ongoing archiving processes require numerous actions such as variant creation, jobs release, jobs follow-up and corrective actions. These activities are typically performed by an SAP BASIS resource or Archive Administrator. The Archive Sessions Cockpit performs the complete process in a fully automated manner reducing the need for additional resources.

The ASC frees the SAP Archive Administrator from these recurring, manual archiving operations.

Unlike external job schedulers, ASC was developed specifically for managing complex archiving scenarios with disparate residence/retention settings and diverse archiving schedules for different organizational units and/or document types.

SAP has certified a number of CMS systems for use with ArchiveLink (See Figure 5). These systems offer a number of advantages over other storage management options. For example, Hierarchical Storage Management (HSM) relies on low-cost tape storage, but often is not able to deliver the access speed needed – users often find that requests will "time out" due to the time it takes to load the tape media and locate the archived file.

Another option is accessing the data archive through the local file system. This option is very fast, but requires high administrative and disk costs. It also presents a data protection risk. Typically, any user with access to the local file system can delete an archived file, increasing the chance of accidental or unwanted data loss.

An ArchiveLink CMS is simple to administer and integrate with data storage platforms such as Hard Disk Write Once (HDWO) that deliver the best price performance ratios. Each of the common ArchiveLink CMS options has been ArchiveLink Certified by SAP to offer the same core performance levels and communications links with SAP archives.

PBS ContentLink CMS

One CMS option gaining traction among SAP users is PBS ContentLink. **PBS ContentLink is a powerful high-performance, alternative to traditional content management systems** that enables archiving to meet legal requirements and support an ILM strategy. Unlike traditional content management systems, PBS ContentLink uses SAP as the application layer and hard disk storage systems for long-term data retention to deliver fast, reliable access to stored data, documents, and DART extracts. The result is a content management system that can be implemented faster and administered simply while delivering higher performance at lower costs.

DART Plots the Audit Trail

SAP users can plan ahead for tax audit information requests using DART. The Data Retention Tool (DART) was developed by SAP to enable its users to meet legal and reporting requirements.

DART allows users to extract data from multiple sources – the live database, archives, and combine them into a single report to quickly and efficiently meet the requirements of a tax audit, or legal or regulatory request.

PBS ContentLink is implemented as a "stateless" translation software between the SAP solution and long-term storage devices for archived data – meaning that no information on stored data and documents has to be maintained on the ContentLink content management server. Instead, the PBS ContentLink server maintains only configuration data and, optionally cached data of stored documents for performance improvement. Persistent information resides on the storage hardware layer and within the SAP solution. With stateless implementation, PBS ContentLink takes advantage of existing backup and replication functionality for SAP and storage hardware and requires no additional configuration.

From a hardware standpoint, SAP ILM solutions now can rely on WORM-like storage (write-once, read-many) called HDWO. Available from all major storage hardware vendors, HDWO disks offer high data protection on low-cost disks that feature high read performance, scalable storage volume, and built in replication functionality.

Solution	Hardware
PBS ContentLink	**NetApp Filer**
	IBM DR550
	EMC Centera
	NAS
SAP Content Server	**Disk**
Various CMS; e.g. , IBM CommonStore and FileNet, EMC/Documentum, Open Text, etc.	**Multiple hardware**

Figure 5: Refer to http://www.sap.com/ecosystem/customers/directories/SearchSolution.epx "Certification Category" = "Archiving" for additional ArchiveLink certified products

RETENTION MANAGEMENT

Retention management in SAP requires two key decisions:

- How long do we have to retain archived data and documents?
- How do we control the end-of-life destruction process?

The first is a business decision, based on business need and compliance requirements. The second is a technology decision: How do we align our storage systems, tools and applications to keep and manage retention records and then identify and purge files at their end of life?

Once again, retention management benefits from the up-front planning that goes into the archiving process. In a best practice scenario, data is grouped according to its retention period as part of the archiving process. Groupings can be by record type, country or region. In this way, data and documents with identical retention periods can be stored in the same repository, and the storage and content management system functionality can be used to enforce defined retention periods.

There are some gaps in standard SAP functionality for managing data retention. While SAP Data Archiving and ArchiveLink provide an excellent framework for controlled, unified data storage, there are no out-of-the-box capabilities for applying corporate record retention policies to this data set. Data Archiving allows administrators to define when data can be removed from the database and into archives, but not when and how to dispose of the archived records when their retention period has passed. Nor do the standard functions allow distinct management of data applicable to a legal or compliance hold.

Dolphin developed the **Archive Retention Management Cockpit** as a portal for retention management and end-of-life document and data disposition in SAP. It is a compliance driven, SAP-based application, and supports any ArchiveLink certified content management system.

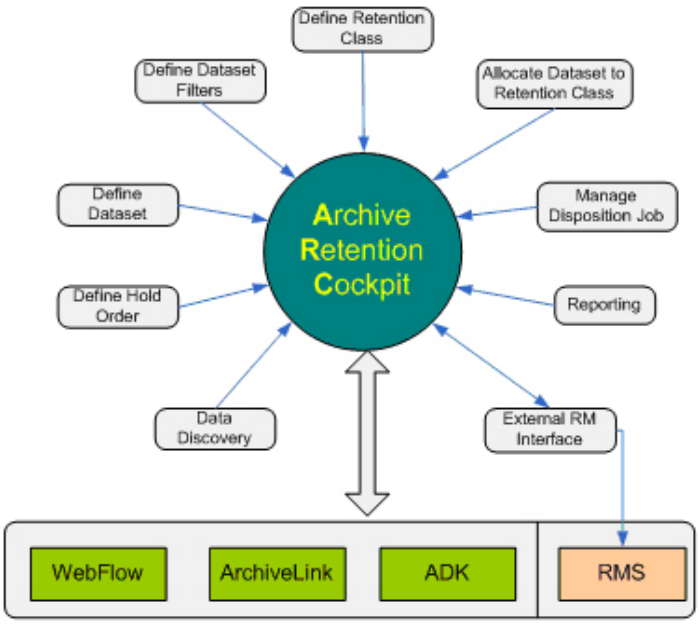

Figure 6: Dolphin Archive Retention Cockpit

The Archive Retention Management Cockpit makes search and discovery simple, allowing users to work within the SAP solution to find records based on a variety of criteria, including open hold orders. Legal hold order management automatically excludes records from destruction during pending litigation or audit. There is also support for structured and unstructured archived data. The cockpit also supports time and event disposition, record disposition workflow, and a record of the activity, and can archive its audit trail for future reference.

RETENTION WAREHOUSE

The issue: How do you consolidate and decommission systems, and still meet data retention and access requirements for legacy data?

Companies implement SAP solutions intending to consolidate or decommission their legacy systems. But often a year or more goes by and the business is still spending resources to maintain the legacy system. Keeping legacy systems operational just to maintain historical information is expensive – they are difficult to maintain, take up valuable space, and require specialized expertise. Moreover, the company runs the risk that the system won't be supported as long as it is required to retain the data. But business and legal compliance requirements make it necessary to ensure efficient access to this information for audit and fiduciary purposes, making management of legacy consolidation in the context of ILM a must.

Approaches to consolidation and legacy decommissioning vary. One approach is to load the data into the new SAP database, but this is complex and time-consuming, and may lead to errors, lost productivity and database performance issues. Others choose to load legacy data onto an independent nearline database and then optionally archive using database-specific nearline archiving tools. However, this database migration does not include business logic – making queries too complicated for most business users. Another alternative is to force the archiving SAP data from the system targeted for decommissioning. These archives can be migrated to the new system and integrated into SAP's Archive Information System for display purposes.

Still another tactic is to bring all the legacy data into the SAP Business Warehouse. The result is an exponential expansion in the size of that database – and the performance issues that go along with it.

System consolidation should lower the cost of IT while improving legacy data access and legal compliance. The key is retaining legacy data in its original application format, keeping the business logic with a single point of access in the target SAP system while allowing decommission of the legacy system. The ease of use and fast access to information fosters user acceptance.

The Retention Warehouse approach for legacy system decommissioning is illustrated in Figure 6. In this scenario, part of SAP's solution for ILM, multiple source systems are decommissioned and archived in accordance with the retention policy into a central Retention Warehouse connected to an ILM-aware content management system. From the Retention Warehouse, users can generate reports on demand using SAP NetWeaver BI solutions.

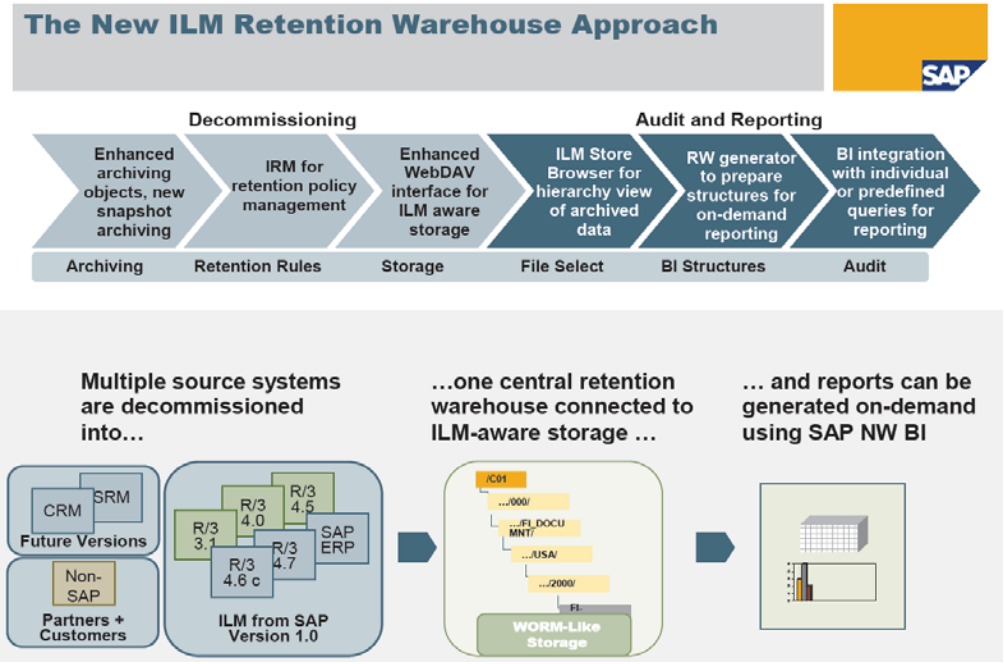

Figure 7: SAP ILM

CONCLUSION

The cost of compliance pales in comparison to the cost of not being in compliance – of not being ready for a legal dispute, audit request or regulatory report; of business data gone missing, or worse, a privacy breach that releases sensitive customer data.

An ILM strategy ensures that data cannot be altered or deleted during its retention period and is available to meet all audit and regulatory requirements. The goals include reducing liability risk by securing data during its retention period and destroying it completely when it should no longer be retained.

It assures high database performance, business continuity, increased productivity and lowering the total cost of ownership in IT through managed storage.

But more than this, the right ILM course gives the business confidence that it is doing everything it can in how it manages its data – for the sake of its customers and itself.

And that is Dolphin's goal. Dolphin's sole focus is on making crucial business operations, like ILM data management and data archiving, run better and smarter for users of SAP solutions.

Dolphin has successfully implemented data archiving across all SAP modules and multiple industries, and our team has extensive archiving experience with all SAP release levels. As the largest provider of PBS software in North America, Dolphin delivers complete services, including SAP data analysis, PBS archive installation, implementation and development of additional retrieval functionality, and support. Our archiving expertise also extends to DART, BI/BW and legacy archiving and decommissioning solutions Dolphin's partnerships with leading storage and archiving add-on tools vendors ensure the best possible solution for your data and document archiving issues.

ABOUT DOLPHIN

Dolphin makes crucial business operations like data management and data archiving, accounts payable, accounts receivable and order management run better and smarter for organizations using SAP solutions. Focused on improving business performance through Information Lifecycle Management and Business Process Management, Dolphin produces the right solution for each customer, faster, through its unmatched experience in SAP technologies, and its proven best practices, tools and SAP add-on applications. Dolphin solutions improve business and IT performance, lower total cost of ownership and deliver high return on investment.

The company was founded in 1995 and has offices in Philadelphia, PA and San Jose, CA. Dolphin solutions are implemented across North America and around the world.