WHITE PAPER
December 2005

Information lifecycle management vision

ABSTRACT

Information lifecycle management has been widely discussed without the formalization of a clear and compelling vision. This paper describes the Sun vision for information lifecycle management, defines the keys to attaining it and provides some insight into how current decisions will facilitate or inhibit attainment of the vision.

Sun microsystems

1 Executive summary

Information lifecycle management (ILM) is a sustainable storage strategy that balances the cost of storing and managing information with its business value. Achieving an advanced state of information lifecycle management consistent with this vision includes:

- Clear linkage of business intent to storage management reality at the application level, including "application-aware" storage management
- Highly granular management capabilities that obviate the need for mass customization in storage management
- Automation of resource management tools and processes, providing management, measurement and dynamic modeling of storage infrastructure, with linkage to business services
- Integrated policy and metadata management, linking business rules derived from service level agreements to storage management actions
- Integrated ILM services for security, retention management, data protection and storage infrastructure optimization, working from a common repository of rules and metadata

An advanced state of information lifecycle management encompasses a storage management world where business information objects are managed automatically, based on their business value. The storage infrastructure is not only business aware, but also is conscious of the content of a business object. Business decision makers determine the rules for managing their information, and the complex service level requirements of a particular information element are consistently and automatically met.

Information technology (IT) and business decision makers armed with a clear vision of information lifecycle management can take steps now to allow their organizations to fully benefit from information lifecycle management in the future. Key areas of focus should include:

- **Enhancing IT/business alignment so that applications or business services eventually form the basis for the management of the storage infrastructure** — Taking the steps today to enhance service management and to model and manage infrastructure based on applications is an important major step.
- **Integrating your organization's information management and information use efforts** — Ensuring the integration of data classification efforts for the purposes of both storage management and access (taxonomy) will help make significant progress toward the effective use of information as a company business asset.
- **Implementing the right ILM-related services** — Solutions for continuous data protection, archive/retention management and device virtualization that show a clear statement of direction toward integration will enable many of the important future benefits of advanced information lifecycle management.

Attaining an advanced state of information lifecycle management can result in significant business benefits, the most significant of which include improved organizational adaptability and improved service to customers, suppliers and internal end users. Fully mature information lifecycle management may result in business performance benefits such as increasing revenue opportunities, reducing costs and driving competitive advantage.

Storage management decisions made during the early stages of development of information lifecycle management can have long-lasting effects on the ability of the organization to mature. A focus on business/IT alignment, acquiring technologies with integration capabilities and using a rich approach to metadata management will facilitate

the evolution of information lifecycle management in your company. Organizations with mature information lifecycle management that links information management to information use should attain solid business benefits and improve competitiveness.

Information lifecycle management is a sustainable storage strategy that balances the cost of storing and managing information with its business value. It provides a practical methodology for aligning storage costs with business priorities.

2 Introduction

Paradoxically, achieving an advanced state of information management generally means that management becomes effortless. Much is happening, but little human intervention is involved. An advanced state of information lifecycle management offers storage management that is practical, sustainable, cost-effective and a source of significant business value. It is both ubiquitous and momentous, critical and automatic.

3 ILM vision

We envision an advanced state of information lifecycle management that involves both pervasiveness and importance. Our vision includes:

ILM vision:

- *Advanced business/IT alignment*
- *Granular management*
- *Automated resource management*
- *Integrated policy and metadata management*

- Business and IT (specifically storage management) are aligned via extensive service level agreements (SLAs) that clearly define the expectations of service. Storage is managed based on policies that are directly linked to the service agreements. The integration of service level management and storage management is adaptive and responsive. The SLA/storage policy linkage immediately absorbs changes to business needs.
- The storage management system is highly granular with the ability to cost-effectively deal with many different policies and data classes. The concept of standard service offerings adopted by sophisticated IT organizations today as a means to simplify storage management becomes obsolete. It is no longer necessary to suboptimize the data classification scheme to maintain a “manageable” number of data classes and polices.
- Resources are optimized via transparent resource management solutions that provide automated modeling, measurement, mapping and planning capabilities. Administrative decisions regarding allocation, provisioning and capacity planning for storage are largely completed within the intelligence of the storage system. These systems provide IT and business users with balanced scorecard metrics that are used to make strategic decisions regarding storage. Execution is completely automated.
- Information management is based on extensive metadata and policy integration. Rich metadata is maintained for all business data objects to facilitate effective management. A semantic information abstraction layer is in place to link the management of data objects with their business use.
- Four key storage management services are integrated: security, retention, protection and infrastructure optimization.
  - Security, including role-based access management, authentication and encryption services, is implemented as defined in service agreements. Storage security services, in addition to perimeter security, are sufficient to meet security objectives providing a second (or third) level of safeguards and ensuring controlling access should other security layers fail.
  - Retention management, including support for compliance requirements, manages reference information so that retrieval and discovery requirements are optimally met. Archived information is available for integration with revenue-producing activities, such as customer service applications and decision support. Within retention management, tiered storage infrastructure is used to contain costs. Broad audit, nonrepudiation, and access control services (applied as indicated by service requirements) are available and are managed by policy. Object-based storage management provides content-based access independent of data type.

ILM vision (continued):

- *Integrated ILM services*
  - *Security*
  - *Retention management*
  - *Data protection*
  - *Storage productivity*
- *Optimized infrastructure*

- Protection of dynamic data is generally served by continuous data protection (CDP) technologies that provide management based on recovery time/point objectives. Data objects managed by CDP services are available for use by applications outside of the primary data paths used by business applications.
- The primary storage infrastructure is optimized via a tiered management approach that automates data movement based on policy driven by service level. The primary data paths are insulated from the CDP and retention data paths allowing business application servers to access information with minimal management overhead.

- The storage infrastructure is extremely granular without large steps between location performance characteristics; in a mature environment, it will be difficult to tell the difference between two similar storage tiers. The infrastructure is self-optimizing and self-healing. Primary storage is separate from protection and retention storage and is accessed and managed via separate data paths. Devices are logically separated from control structures and are accessed via extensive virtualization capabilities. The storage infrastructure provides automated balancing, re-pathing and reallocation. It can be modified to meet changes in business needs in a matter of minutes.

ILM vision building blocks:

Two key building blocks must be in place for this vision to occur.

- *Business/IT alignment must be as fully integrated and automated as possible so that business intent is linked to storage management reality.*
- *Information lifecycle management services to manage, place and move data must be integrated and coordinated. An underlying metadata management repository and policy management system must be available to link business.*

4 Key building blocks

Two key building blocks must be in place to support the information lifecycle management vision: business/IT alignment and integrated ILM services.

4.1 Business/IT alignment

Ensuring alignment between IT operations and business goals has traditionally been viewed as the chief information officer's job. However, successful IT/business alignment entails more than executive-level oversight. Linking business intent with storage reality requires a complex chain of information, decisions, tools and capable infrastructure. Companies need IT management solutions that provide an understanding of how technology supports, affects and enables their business-critical services. Additionally, companies need solutions that allow management to assess the impact of new business plans (for example, new applications or business growth) on IT operations.

Our information lifecycle management vision is based on a view of storage management we describe as "service-level-driven policy management" (SLPM). Service level agreements are defined based on business needs, including specifications for access (performance), recovery time, recovery point, retention and specialized protection (e.g., compliance requirements).

ILM vision building blocks (continued):

information objects to appropriate management actions on the basis of service objectives.

As a company's information lifecycle management strategy matures, integration with business needs evolves. Agreements generally begin as basic agreements for critical services, and IT makes the majority of the service decisions regarding protection, security and retention. This is followed by a stage of maturity during which complex and detailed agreements are implemented, and then by a stage when service offerings are standardized (a standard service "catalog") and selected in the agreements. Finally, as the storage management system matures to require fewer specifications, simple agreements return, specifying only key objectives and relying on intelligent storage management to address the bulk of the details.

An example of an early-maturity agreement would be one that defines the data as mission-critical and allows IT to decide to double-mirror, complete an incremental backup every four hours and do full backups daily. At a later stage, detailed specification of data protection technologies would likely be included. At the final stage, it would be necessary to specify only the important metrics (recovery time and recovery point).

Integrated ILM services:

- *Services*
  - *Continuous data protection*
  - *Archive services*
  - *Virtualization services*
- *Integrated external to the production data path*
  - *Extract, transform and load applications from CDP to archive*
  - *Full-volume anti-virus scans*
  - *Other administrative tasks*

4.2 Integrated ILM services

The integration of retention, protection and virtualization for data outside the primary data path is fundamental to the ILM vision. Service integration is based on a set of integrated software appliances that virtualize a range of storage services. The services can be presented as different types of storage-related applications based on customer needs. Three key capabilities are needed (see Figure 1).

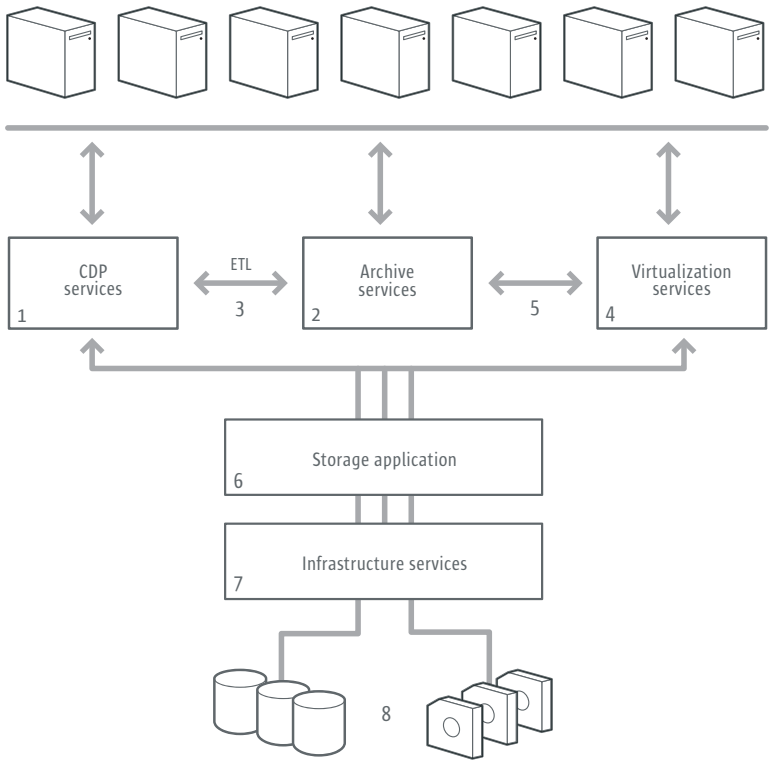

Figure 1: Application servers and storage

First, continuous data protection (CDP) Services [1] (1 in the figure on the previous page) produce and administer point-in-time copies that are managed based on SLA recovery point and recovery time objectives. Image copies of dynamic data are available on a continuous basis, hence the name CDP.

Second, archive services provide a policy-managed object store to serve reference information and compliance needs. Objects are ingested either directly into the object store from primary data [2] or, more interestingly, are extracted from point-in-time copies managed by the CDP system via extract, transform and load (ETL) applications running within the ILM system [3].

Finally, virtualization services for tape (and other devices) will be included and accessible to services within the ILM system (e.g., archive and CDP) [5] or directly from business applications [4].

The outward-facing services of the information lifecycle management system are linked to the physical storage [8] by two layers of management software. The storage application layer [6] (e.g., replication, indexing) and the infrastructure services layer [7] (e.g., device access and management, volume manager, file system, administration, high-

availability, security, messaging) provide common functionality to the CDP, archive and virtualization services.

The combination and integration of these services provides for management of data protection, archive and virtualization outside the primary data path. This capability provides several benefits. Perhaps the strongest benefits are processing of archive data external to business applications, and removal of the synchronization tasks necessary to collect archive data from dynamic storage. Many other opportunities exist for removal of administrative workloads from primary data paths; these include volume-level virus scanning (done on CDP point-in-time copies), version control and software distribution, and creation of synchronized test data.

5 Process and practice

Large-scale IT operations cannot be expected to implement the information lifecycle management vision immediately. Based on the state of ILM evolution in most large businesses and on the development of the required technologies, it is likely several years away. This is partially due to the ability of organizations to change, but is also dependent on the development of several parts of the vision expressed previously.

We believe that an advanced state of information lifecycle management will be most successfully implemented through a set of rational, step-by-step initiatives. Once an organization has reached a state of information lifecycle management maturity consistent with the technology available today, the next steps toward the vision largely involve optimization of storage management processes; broadening of optimized processes across the entire enterprise; linkage of policies that support separate data protection, retention and storage optimization efforts; and adoption of key technologies.

Next steps:

- *Early-maturity decisions and later-stage development: Depending on the tools and approaches selected in the early stages of ILM development, further maturation may be a simple integration project that broadens the coverage of the tools.*

5.1 Early-maturity decisions and later-stage development

Some implementations will result in a need to replace the storage management tool set, which may be very costly in terms of headcount, training and software. We advise using significant care today in the selection of tools and approaches to facilitate high levels of maturity. In each initiative that follows, we will address the characteristics necessary to enable further evolution toward the information lifecycle management vision.

Initiatives:

- *Enhanced business alignment*
  - *Using automated storage resource management for dynamic modeling, application-aware storage management and measurement*
- *Cross-enterprise expansion*
  - *Service-level-driven policy management*
  - *All platforms*
- *Taxonomy development*
  - *Integration of information management and information science*
- *ILM-related services*
  - *Continuous data protection*
  - *Archive/retention management*
  - *Device virtualization*
  - *Begin with nonintegrated approach with future integration capability*

5.2 Initiatives

Several key initiatives will drive an organization toward the vision for information lifecycle management:

5.2.1 Enhanced business alignment and resource management

Organizations that have achieved a moderately mature state of ILM growth have most likely solidified their business/IT integration via a service level management process. To achieve this vision of information lifecycle management, a close linkage is required between infrastructure and the applications served by that infrastructure.

This process involves planning for application requirements with business end users, integrating application requirements into a dynamic model for IT infrastructure, developing a set of service level objectives and measuring and managing to those objectives.

The difficulty in aligning specific elements of storage management in this context is most likely the linkage of an application with its data objects and with the specific elements of infrastructure involved. The effective use of a system resources manager solution that includes modeling and applications linkage is a key success factor in achieving application-aware storage management.

Implementing an initiative aimed at achieving application-aware storage management requires management focus, people, modeling and measurement tools. It is likely to be most successful if viewed as an evolving initiative rather than a "mega-project."

5.2.2 Cross-enterprise expansion

The complete expansion of ILM storage management practices and polices (across disciplines and the enterprise) is a cornerstone of the information lifecycle management vision. To achieve the benefits of service integration, administration that is out of primary data path, business alignment and eventual ILM ubiquity, the ILM storage management toolset must be applicable to all servers, storage platforms and devices.

Service-level-driven policy management (SLPM) relies on encompassing control, management and measurement systems. It simply doesn't work if a storage appliance is outside the control of the management system. Generally, this means that subsystems in the storage hierarchy need to have standard measurement and management interfaces.

Standard policy/metadata systems are also a requirement. Policy in an SLPM environment is forged at the data classification level and is executed by specific service engines. Metadata and policy reside in a single repository or in an intelligent hierarchical system of repositories having a common overarching engine to resolve policy conflicts.

5.2.3 Taxonomy development

Rich data classification, while pivotal to SLPM from a storage management perspective, will also enable optimized use of enterprise information assets. One of the interesting things that is likely happen on the path to this information lifecycle management vision is a blending of information management with information use. Information technology organizations should expect that a librarian/information scientist or two (or several) will be intimately involved in their ILM maturation process.

Several analysts have suggested that this linkage will evolve into an additional layer in the infrastructure/information management stack, calling it a "semantic information abstraction" or "organic information abstraction" layer.[1] What it represents is the emergence of information asset management as a high-value business process, and it signals a changing role for IT: an evolution from data managers to enterprise information stewards.

5.2.4 ILM services and related technology

Evolution to this vision of ILM requires a platform/technology approach to continuous data protection, virtualization and archive/reference information services with solid potential for future integration. Linked CDP, virtualization and archive/reference information services can significantly impact primary data path productivity when integrated. IT organizations that want to take advantage of separating administrative and secondary functions from their primary data paths should seek solutions that have consistent approaches to policy, metadata management and measurement.

6 Benefits

Achieving the Sun information lifecycle management vision — or any truly advanced state of ILM — can accrue substantial benefits to an IT organization in terms of productivity and control. More importantly, sophisticated information lifecycle management will advance the management of enterprise information to help achieve competitive advantage for the company.

The largest benefits of the evolution to advanced information lifecycle management are business-oriented rather than IT-oriented. Advanced ILM helps improve service to business

[1] *Orlov, Laurie M., and Laura Ramos. May 2004. "Organic Information Abstraction." Cambridge, MA: Forrester Research.*

Achieving an advanced state of information lifecycle management should result in significant business benefits:

- *Organizational adaptability*
- *Improved service to business end users*
- *Ability to improve data integration — dynamic and archive data*
- *Improved performance of production data systems*
- *Reduced IT costs*
- *Reduced business risk*

Combined benefits generate business performance results:

- *Increased revenue*
- *Decreased cost*
- *Enhanced competitive advantage*

end users, enable the integration of information for reuse in customer- and supplier-facing applications and improve the adaptability of the organization.

- **Business/IT alignment** — IT closely linked to business processes and applications should result in improved end-user service, can focus storage management on important business requirements and should improve the organization's ability to adapt to business changes.
- **Integrated information management and use** — Information management and information use are more closely aligned via the linkage of data classification to information taxonomy. The business value extracted from information is optimized; information becomes an asset that is managed for competitive advantage.
- **Data reuse/integration** — The reuse/data integration of reference or archive information managed and facilitated by integrated ILM services and metadata management should result in enhanced competitive advantage driving increased loyalty of suppliers and customers.
- **Improved service to business end users** — Out-of-primary-path processing for archive workloads and administrative processing (anti-virus, software management, test data) should improve performance of existing application servers. Implementation of a policy-managed storage hierarchy can maintain service levels at the lowest possible cost.
- **Cost containment** — Broad policy-based storage management can reduce errors, omissions and headcount requirements, driving cost and risk out of storage management. Further, automated storage management is scalable without incremental cost, allowing IT organizations to achieve the benefits of technology advances without incremental costs beyond cost/capacity.

7 Conclusion

Storage management decisions made during the early stages of information lifecycle management development can have long-lasting effects on the ability of the organization to mature. Achieving the benefits of advanced information lifecycle management will likely be facilitated by defining your vision for ILM and overlaying that knowledge on the decisions made today. Driving business/IT alignment, acquiring a service base with integration capabilities and using an information-rich approach to metadata management will drive your organization to an application-aware storage management model that will help you link information management to information use.

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN **Web** sun.com

JT 0012 C 12/05