

## Association for Information Systems AIS Electronic Library (AISeL)

---

AMCIS 2007 Proceedings

Americas Conference on Information Systems  
(AMCIS)

---

December 2007

# Cost Impacts on Information Lifecycle Management Design

Lars Turczyk

*Technical University Darmstadt*

Martin Behrens

*TU Darmstadt*

Nicolas Liebau

*TU Darmstadt*

Ralf Steinmetz

Follow this and additional works at: <http://aisel.aisnet.org/amcis2007>

---

### Recommended Citation

Turczyk, Lars; Behrens, Martin; Liebau, Nicolas; and Steinmetz, Ralf, "Cost Impacts on Information Lifecycle Management Design" (2007). *AMCIS 2007 Proceedings*. 36.

<http://aisel.aisnet.org/amcis2007/36>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2007 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# Cost Impacts on Information Lifecycle Management Design

**Lars Arne Turczyk**

Technische Universität Darmstadt  
KOM – Multimedia Communications Lab  
Merckstr. 25  
D-64283 Darmstadt  
Germany  
lars.turczyk@siemens.com

**Martin Behrens**

Technische Universität Darmstadt  
KOM – Multimedia Communications Lab  
Merckstr. 25  
D-64283 Darmstadt  
Germany  
martin.behrens@kom.tu-darmstadt.de

**Nicolas Liebau**

Technische Universität Darmstadt  
KOM – Multimedia Communications Lab  
Merckstr. 25  
D-64283 Darmstadt  
Germany  
nicolas.liebau@kom.tu-darmstadt.de

**Ralf Steinmetz**

Technische Universität Darmstadt  
KOM – Multimedia Communications Lab  
Merckstr. 25  
D-64283 Darmstadt  
Germany  
ralf.steinmetz@kom.tu-darmstadt.de

## ABSTRACT

Information Lifecycle Management (ILM) is a strategic concept for storage of information and documents. ILM is based on the idea that in an enterprise information have different values. Information with different values are stored on different storage hierarchies. ILM offers significant potential cost savings by tiering storage and 90% of decision makers consider implementing ILM (Linden 2006). Nonetheless, there are too few experience reports and experimenting and researching in real systems are too expensive. In addition, 66% of IT managers do not have time to put together a basic cost model or a data value model for ILM projects. (Foskett 2006). This paper addresses these issues and contributes to supporting and assisting IT managers in their decision-making process.

We present a cost model for ILM in an enterprise. The model is used for ILM simulations. The simulative approach has two advantages. On the one hand it offers predictions about the dynamic behaviour of an ILM scenario and, on the other hand, it dispenses with real storage hardware. The simulated results lead to design guidelines for ILM environments which help to avoid mismanagement and poor investments in advance. The results raise the awareness of the required number of hierarchies and the choice of storage technologies.

## Keywords

Information Lifecycle Management, Total Cost of Ownership, Design

## INTRODUCTION

Information Lifecycle Management (ILM) is a strategic concept for storage of information and documents (Peterson 2004). ILM is based on the idea that in an enterprise different information have different values. Valuable information is stored on systems with a high quality of service (QoS). The value changes over time and therefore migration of information is required to cheaper storage systems with a lower QoS. Automated migration makes ILM dynamic. Such automation requires storage systems to understand which files are important at what time so that right policies can be applied. In this point ILM nowadays lacks methods and tools.

For tapping cost potentials it is necessary to obtain a broader knowledge about ILM procedures and scenarios. There are too few experience reports on this subject and experimenting and researching in real systems are too expensive. Techtarget found

that 66% of IT-managers do not have time to put together a basic cost model or a data value model for ILM projects. (Foskett 2006). Therefore the aim of this paper is to describe a cost model to generate results and experiences by simulation of ILM scenarios.

The research questions are “What are the relevant cost factors in ILM environments?” and “What are their impacts on ILM design?” To answer the first question we derive a cost model. Thereafter simulations will be used to predict performance of ILM scenarios. The simulative approach for answering the second question has two advantages. First, it offers predictions about the long-term dynamic behaviour of an ILM scenario. Second, it dispenses with real storage hardware which allows comparison of architectural design alternatives.

The model intends to support the design and decision process on ILM implementations. To be truly successful with the design and implementation of an ILM solution, it is important to remember that only about 25% of the challenge is about the selection of products (Patel and Shah 2005). We focus on the other 75% and derive a cost model that leads to a simulation tool. The simulator offers the possibility of checking the cost under dynamic circumstances without implementing real systems. This can help to avoid mismanagement and poor investments in advance.

To create a best possible real cost model we conducted two case studies in the past. In 2005 we analyzed a knowledge database from a consulting department within a large German enterprise. The results showed that, on the one hand, there is demand in industry for ILM and that, on the other hand, Microsoft (MS) Office application formats .doc, .xls and .ppt dominate the file population (Gostner et. al. 2005). This is quite the reverse to the internet where similar studies were conducted (Page et. al. 1999). Therefore in 2006 a second case study focused on the access behaviour of MS files which is different to the access behavior of internet sites (Groepel et. al. 2006).

We now use these results to derive a cost model for dynamic ILM environments. The model will be implemented in a simulator to check its feasibility. The paper concludes with the generation of first results by simulation and their application to ILM design.

The essence of this paper is as follows:

1. We present a cost model for ILM in an enterprise
2. We use the model for ILM simulations
3. We show how the simulation can lead to individual design guidelines for ILM environments before spending money on real systems

## RELATED WORK

ILM is relatively new and has its roots in Hierarchical Storage Management (HSM). Some results from former HSM investigations can be used for ILM. Especially work done on file observation and algorithms is still relevant. Strange examined the long-term access behavior on files in an UNIX system (Strange 1992). His aim was to identify regularities and patterns which can be applied to automated migration strategies for HSM. Further work deals mainly with algorithms which can be used for ILM or other storage strategies. Some examinations focus on the analysis of the access behavior and the development of migration strategies. The file migration protocol listing of a supercomputer was analyzed in a study by Miller and Katz. Migration methods were developed for a corresponding system (Miller and Katz 1993). Miller and Gibson examined the access behaviour in further studies in UNIX environments and designed a "file aging algorithm" as a migration rule (Gibson and Miller 1999).

Today ILM is a strict focus of research. The main results are found in the field of “How” ILM works, i.e. most research was done in the field of procedures and policies. Vendors presented their understanding of ILM (Reiner et. al. 2004). Turczyk et. al. (2006) gave a formal definition usable for ILM abstraction. Beigi et. al. (2005) and Tanaka et. al. (2005) offered proposals for policy description of ILM. Last but not least, Chen (2005) focused on the valuation of files.

Cost models for decision support were not considered in the technical research. Nonetheless ILM is a very real topic on which IT managers have to decide. Of course cost models for IT are found for nearly every situation. For ILM cost models for data center environments (Merill 2003 and Patel and Shah 2005) can be applied if no definite model is available.

The economic view of ILM is driven by analysts. They focus on questions about trends, budgets and decisions. The IT managers' perspective is quite different from the technical point of view. Regulation has a great impact on storage for IT managers (Merrin and Harnetty 2005). Their main focus is on cost impact (Foskett 2006 and Linden 2006). Furthermore Short (2006) recognized that IT Managers develop an ILM-related fear of vendor lock-in. Again Foskett (2006) found a total of 66% of IT Managers having no time to model.

Therefore in this paper we combine the ILM technology with economic IT decision support.

## BASIC PRINCIPLES OF IT ECONOMICS

### Starting point of today's IT departments

Storage volumes are growing at phenomenal rates, yet IT organizations cannot always justify increasing storage budgets. IT organizations expect budget allocations for overall storage services, storage hardware, and software infrastructure to be constant year after year - a percentage of total IT budgets. To make matters worse, cost reductions are still in force and IT continues to be a target for reducing costs while still required to deliver critical services. The argument is IT price erosion. In fact in storage area enterprise, modular and locally attached disk will all continue to experience price erosion. Nevertheless price per megabyte is a poor single metric to use in storage economic decisions. Industry best practice is to take a strategic view of storage architectures and determine the best solutions based on total cost of ownership (TCO) and operating expense (OPEX) minimization.

This strategic view is required for strategic decisions about new storage concepts such as ILM. ILM is a concept for cost-reduction activities that can yield real OPEX savings for the enterprise storage infrastructure.

90% of decision makers consider implementing ILM (Linden 2006). In parallel Paquet from Gartner Inc. predicts ILM to play a large part in the surge in storage management software sales. License revenue for storage management software will reach \$9.4 billion in 2009 with an annual growth rate of 10.9% from 2004 through 2009 (Paquet 2005).

One of the reasons for growth is that businesses are looking to better manage their storage capacity use and are starting to automate some of their management functions. Automated file migration is a key of ILM. It makes ILM dynamic. Such automation requires storage systems to understand which files are important at what time so that right policies can be applied. For design and decision support ILM nowadays lacks methods and tools to map the dynamic behavior resulting from automation.

Our TCO model implemented in a simulator allows the dynamic character of ILM to be reflected. This is the advantage for design and decision support.

## EXPENSE MODEL

### TCO

Best practices indicate that TCO requirements are included in all competitive bid situations. Since many people look at CAPEX (Capital Expenditure) cost as a one-time cost, they are fooled into thinking that the lowest purchase price is also the lowest TCO. This is not the case in most storage and data center environments. Storage purchase decision-makers should be looking at the TCO of competitive solutions, and not just at the lowest price per megabyte presented by vendors in the negotiations.

TCO, on the contrary to return on investment (ROI), is effective for calculating total lifecycle costs for competitive or comparative solutions. This is the case in ILM when planning storage growth or expansion resulting from either vendor options or architecture options.

It is our intention to use TCO for a comparison of architectural approaches.

For data centers dedicated TCO calculation formulas already exist, e.g. Patel and Shah have described the following approach (2005):

$$COST_{total} = COST_{space} + COST_{power\ hardware} + COST_{cooling} + COST_{operation}$$

This formula is suited to the operation of a data center, but is not suited to ILM design consideration with its dynamical character of permanent changing storage demand.

Therefore we summarize the indirect cost space, power, cooling and operation to "administration cost". Furthermore we emphasize the cost for storage capacity. This results in a first formula:

$$COST_{total} = COST_{storage\ capacity} + COST_{administration}$$

We look at the cost over the lifecycle, therefore the time-related aspect is added by consideration of the integral:

$$COST_{total} = c(t) = \int (COST_{storage\ capacity} + COST_{administration})$$

This formula will be specified afterwards. In the following section the ILM-relevant cost factors are considered.

### Expense factors

The cost behaviour of ILM scenarios is generally determined by different influential factors (Short 2006).

#### Data growth

The University of California concluded that in 2003 alone around 5 exabytes ( $10^{18}$  Bytes) of new "stored information" were produced (Lyman and Varian 2003). Even more amounts of data will be produced over the next few years, several analysts report. They all speak of steadily growing capacity demand. The compound annual growth rate (CAGR) varies between 60% and 100%. IDC detected a CAGR of 76% over the years 2000-2004 (Nuygen 2005).

In 2005 META forecasted the annual growth rates of tiered infrastructure as follows: 20% to 25% percent for enterprise (monolithic), 50% to 55% for midrange (modular), and 80% to 85% for capacity-based, yielding aggregate storage growth of about 45%. Like-for-like price/capacity will improve 35% per year. The implication is that the use of lower cost storage alternatives will be a key facet of storage growth over the next few years (Meta 2005)

Although the exact value is not known, the tendency is obvious. In conclusion we assume the aggregate storage growth by a factor  $g \in [0.45; 1.0]$  per year.

#### Price Decline of Hardware

The tendency for storage prices is constantly declining. Over years analysts gave similar prognoses. Between 1998 and 2001 McKinsey determined for the price per gigabyte (GB) a CAGR of -36% (Kraemer 2001). IDC took a look at the prices per GB between 2001 and 2003. In 2003 per-gigabyte external storage prices fell -33%, while in 2002 and 2001, they declined -40% and -43% respectively. So the CAGR between 2001 and 2003 is -36% (Nguyen 2005).

In 2005 Sun Microsystems said prices have been declining annually at approximately 35% (Sun 2005). Prices for enterprise disk were in the \$25 to \$40 per gigabyte range. Midrange disk is at \$12 to \$20 per gigabyte, low-end disk is at \$3 to \$10 per gigabyte. High-end tape automation is at \$1 to \$2 per gigabyte. Knowing that these prices will be obsolete by now, the relation is almost constant. This leads to a relation between enterprise disk (FC, SCSI, FICON, ESCON<sup>1</sup>) and midrange disk (SCSI, FC) is 2:1. The price difference between midrange disk and low-cost disk (S-ATA<sup>2</sup>) is 3:1. The price difference between low-cost disk and automated tape is 5:1

Horizon Information Strategies gave a similar relation in 2003 (Moore 2003). The price difference between enterprise disk and midrange disk is 2.5:1. The price difference between midrange disk and low cost disk is 3:1. The price difference between low cost disk and automated tape is 6:1.

It is important to know the relations when designing ILM from the aspect of cost optimization. The relations tend to be relatively constant, as shown above.

#### Administration Cost

The cost situation changes when considering the total cost of ownership (TCO). IDC reported that an enterprise spends an average of \$3 administrating storage for every \$1 spent on hardware (Nuygen 2005). Additionally Gartner Group speaks of \$3.5 being spent for managing every \$1 spent for storage hardware (Paquet 2004). This relation between administration cost and hardware cost has to be considered in the cost model, too.

### Cost Model

The cost model developed consists of the storage costs and administration cost:

$$COST_{total} = c(t) = \int_{t_0}^t COST_{storage\ capacity} dt + \int_{t_0}^t COST_{administration} dt$$

$$\text{with } COST_{storage\ capacity} = AMOUNT \cdot PRICE =: a(t) \cdot p(t) \text{ and } COST_{administration} = 3 \cdot COST_{storage\ capacity}$$

<sup>1</sup> Fibre Channel (FC), Small Computer System Interface (SCSI), Fibre Connection (FICON), Enterprise Systems Connection (ESCON)

<sup>2</sup> Serial Advanced Technology Attachment (S-ATA)

$$\text{Hence } COST_{total} = c(t) = 4 \cdot \int_{t_0}^t a(t) \cdot p(t) dt$$

An ILM system consists of different storage hierarchies with different cost structures. Therefore the accompanying cost model must also support different hierarchies.

$${}^n c(t) := \sum_{i=1}^n 4 \int_{t_0}^t a_i(t) \cdot p_i(t) dt$$

The highest level has the most expensive cost function. The hierarchies thereafter are arranged in descending order of their cost function.

Under this prerequisite there is a price relation  $\beta_i \leq 1$  between two successive hierarchy levels.

$$p_i(t) = \beta_i \cdot p_{i-1}(t)$$

Furthermore the storage demand of hierarchy  $i$  can be expressed by the relative need  $\alpha_i \leq 1$  according to the entire capacity need  $a(t)$ :

$$a_i(t) = \alpha_i \cdot a(t)$$

The cost function for an  $n$ -dimensional ILM scenario can be expressed as follows:

$${}^n c(t) = \sum_{i=1}^n 4 \int_{t_0}^t \alpha_i a(t) \cdot \beta_i p_{i-1}(t) dt$$

In order to examine the cost reduction potentials generally, we consider the relation between two designs. The  $n$ -dimensional ILM system is compared with the design without ILM (no tiered storage, i.e. 1-dimensional costs). This relationship  ${}^n r$  arises from the relation of the respective cost functions:

$${}^n r := \frac{{}^n c(t)}{{}^1 c(t)} = \frac{\sum_{i=1}^n 4 \int_{t_0}^t \alpha_i a_i(t) \cdot \beta_i p_{i-1}(t) dt}{4 \int_{t_0}^t a(t) \cdot p(t) dt}$$

The relation allows canceling the factor “4”.

The relative prices  $\beta_i$  are subject to negligible fluctuations so that they can be regarded as temporarily constant (see section on expense factors).

Hence

$${}^n r = \frac{\sum_{i=1}^n \int_{t_0}^t a_i(t) \cdot p_i(t) dt}{\int_{t_0}^t a(t) \cdot p(t) dt} = \frac{\sum_{i=1}^n \alpha_i \prod_{j=1}^i \beta_j \int_{t_0}^t a(t) \cdot p(t) dt}{\int_{t_0}^t a(t) \cdot p(t) dt} = \sum_{i=1}^n (\alpha_i \prod_{j=1}^i \beta_j)$$

The cost relation  ${}^n r$  only depends on the storage demand per hierarchy  $\alpha_i$  and the relative prices per hierarchy  $\beta_i$ .

Again, analysts say the relative prices remain almost constant for all enterprises (Moore 2003 and Sun 2005).

But what about the  $\alpha_i$ ? They are dynamic. Without consolidated values the model is obsolete.

To answer this question and to derive the values for the  $\alpha_i$ , we now implement a simulator and proceed with the decision/design process.

## APPLICATION OF THE MODEL

### Example

We implemented a simulator to derive the  $\alpha_i$ . For applying the cost model we have designed an instance for examining the cost reduction potentials of ILM.

The aim of the instance is comparison of TCO of two storage architectures. One storage solution works with 3-dimensional ILM the other without ILM (i.e. 1-dimensional ILM). We try to derive general design guidelines from the comparison.

We attempt to demonstrate with the example on which scale ILM can typically contribute to a reduction in the storage cost.

Therefore a scenario as realistic as possible is feigned. From the simulation data the relative storage requirements  $\alpha_i$  of the individual hierarchies are derived and the results are then applied to the cost model.

The basis for the cost comparison is an ILM scenario with three storage hierarchies. The simulator feigns the migration behaviour of office files over a period of 2,000 days. We assume that the amount of data rises by about 60% every year.

To be able to make use of the potential of ILM completely, files must be moved automatically between the storage hierarchies. The simulator uses migration algorithms which calculate the access probability from access history, file type and file age. If this probability is above 10%, then the file remains at the first storage hierarchy. If the probability falls to under 10%, the file is migrated to the second hierarchy. If the access probability is less than 2.5%, the file is migrated to the third hierarchy.

The simulator feigns the capacity need  $\alpha_i$  for each hierarchy (see figure 1). Since the migration method looks at the access history it needs some time to stabilize the capacity-need in an ILM system. The stationary storage requirements for the cost calculation are then simply applied to the cost model.

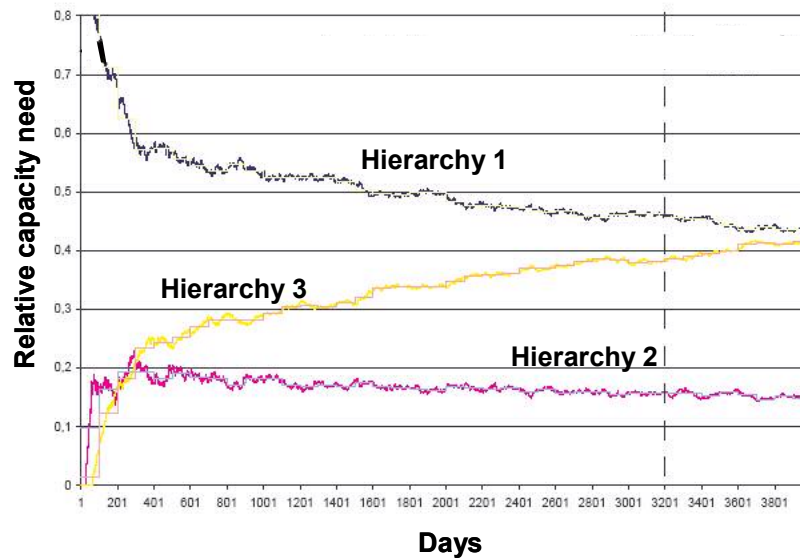


Figure 1: Determination of alpha 1, alpha2 and alpha 3 by simulation

The determined  $\alpha_i$  measured after 3200 days are:

$\alpha_1 = 47\%$ ,  $\alpha_2 = 38\%$  and  $\alpha_3 = 15\%$ .

Now we define  $\beta_i$ . A cost-reducing effect can be achieved by ILM if the additional hierarchies are cheaper than their predecessor hierarchies. In our instance we assume that the first storage hierarchy consists of high-quality FibreChannel hard disks.

The first hierarchy has a relative price  $\beta_1=1$  by definition. The second hierarchy is a SATA hard disk level. The relative price is determined by the market and is in proportion to the FibreChannel technology at about 1/7,5 ( $\beta_2 = 1/7,5$ ) (Moore 2003).

The third hierarchy is supported by a tape memory technology. According to market analyses SATA hard disk is approximately six times more expensive than tape memory. Therefore the relative price is  $\beta_3 = 1/6$ .

We have  $\beta_1=1$ ,  $\beta_2 = 1/7,5$  and  $\beta_3 = 1/6$ .

Inserting the feigned alphas and derived betas in the cost model, then one can simply calculate the relationship of the costs or the cost saving.

$$r^3 = \sum_{i=1}^n (\alpha_i \prod_{j=1}^i \beta_j) = 0,52$$

This means that about 48% (1-0,52) lower costs arise in our instance for data storage with ILM.

Below we look more generally at the cost reduction effects of ILM.

### Generalization of the example

The relative capacity need is dynamic for each storage hierarchy. Alphas are derived from simulations.

The betas are also variable and depend on the choice of storage technologies.

Using the cost model we now provide a graphic representation of the cost reduction potentials of 3-dimensional ILM dependent on alpha and beta.

Therefore the following graphics are based on a solution with three storage hierarchies.

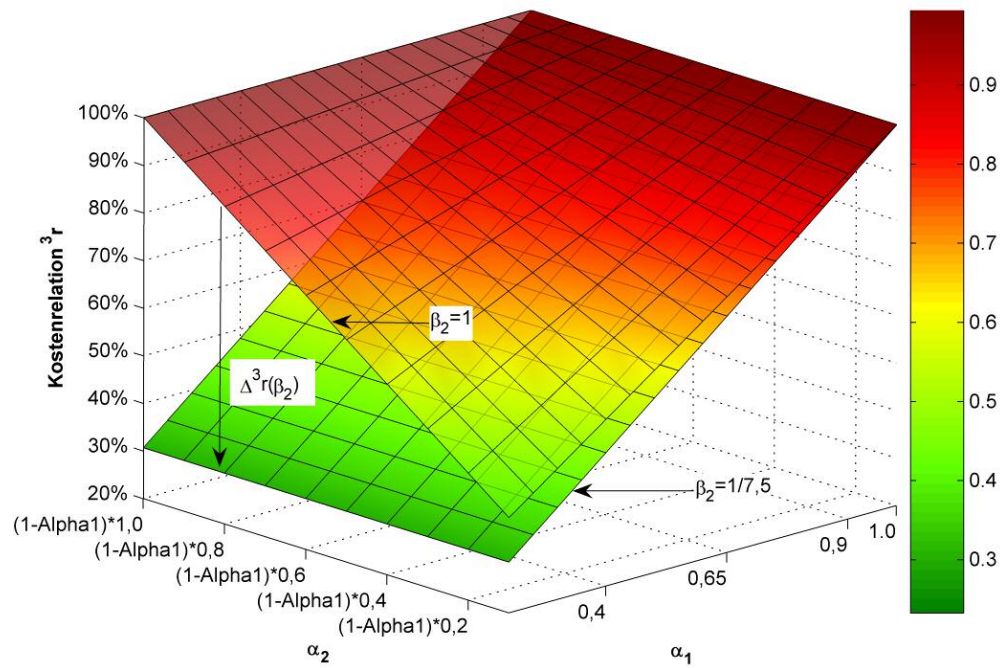
Figure 2 shows the relation of storage costs of a 3-dimensional ILM solution compared to a solution without ILM. We see  $r^3$  dependent on relative storage requirements of the upper two hierarchies  $\alpha_1$  or rather  $\alpha_2$  and the relative price  $\beta_2$ .

$\beta_3 = 1$  is fixed.

The lower of the two distorted areas represents the result area if  $\beta_2 = 1/7,5$ . A point in this area represents  $r^3$  with a permitted combination of  $\alpha_i$ .

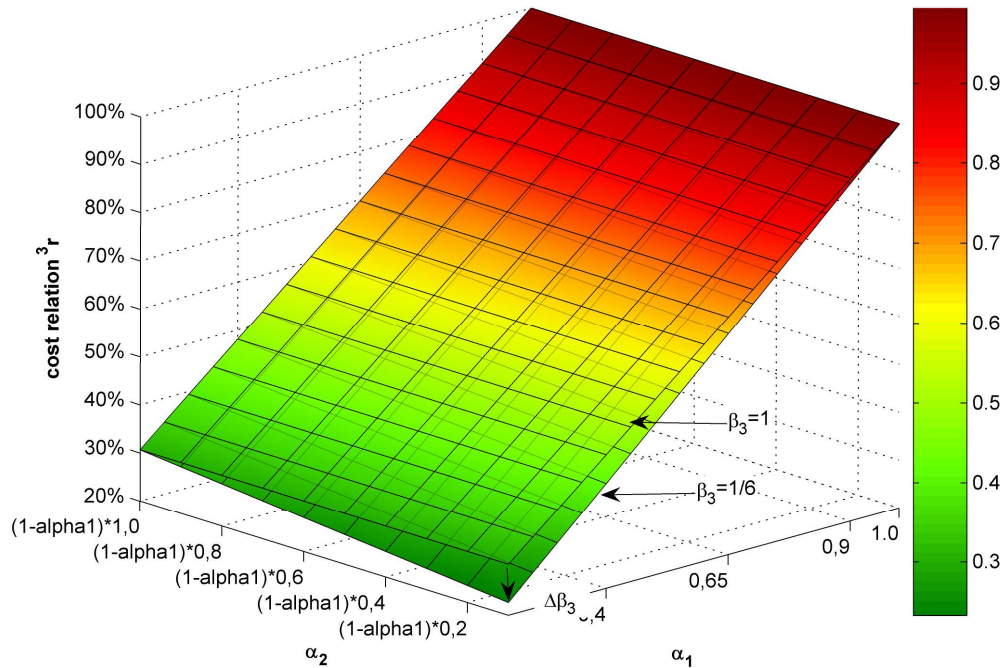
$\Delta^3 r(\beta_2)$  is the plumb line between the two areas and the diminution of the relative storage costs if the relative price  $\beta_2$  is lowered from 1 to 1/7,5. If  $\beta_2$  is between these two values, the resulting cost relation is located between the two highlighted areas.





**Figure 2: Cost relation dependent on beta<sub>2</sub> with beta<sub>3</sub> = 1 fix**

A similar representation is included in figure 3. Unlike the previous graph, the cost relation in figure 3 is shown dependent on  $\alpha_1$ ,  $\alpha_2$  and  $\beta_3$ .  $\beta_2$  is set to be 1/7,5. The lower colored area contains the cost relation  ${}^3r$  at a relative price  $\beta_3 = 1/6$ .



**Figure 3: Cost relation dependent on beta<sub>3</sub> with beta<sub>2</sub> = 1/7,5 fix**

At a price of  $\beta_3 = 1$ , the price of the second and third hierarchies remains the same. For economic reasons such an order makes no sense.  $\Delta^3 r(\beta_3)$  represents the cost saving at a reduction in the relative price from  $\beta_3 = 1$  to  $\beta_3 = 1/6$ .

The cost reduction effect by a reduction of  $\beta_3$  will turn out all the greater, the higher  $\beta_2$  is. The effect in the price constellation shown is quite small since a great saving is already obtained by the low relative price of the second level  $\beta_2$ .

An essential part of the storage costs can apparently be saved by a sufficiently small relative price of the second level. This follows from the restriction that the sum of the saving is smaller than 1. Under this prerequisite the savings by adding more hierarchies will still only be marginal.

$$\Delta^n r = \sum_{i=2}^n \Delta^n r(\beta_i) \leq 1$$

$$\Rightarrow \Delta^3 r = \Delta^3 r(\beta_2) + \Delta^3 r(\beta_3)$$

Hence ILM per definition needs at least two different hierarchies, but when assuming that admin cost increase with each hierarchy, too many hierarchies will counteract the ILM-cost- effect.

These results can helpfully be applied to the design process of an ILM system.

## SUMMARY AND OUTLOOK

Information Lifecycle Management (ILM) is a strategic concept for storage of information and documents. ILM is based on the idea that in an enterprise information have different values. Information with different values are stored on different storage hierarchies.

90% of decision makers consider implementing ILM (Linden 2006). Nonetheless, there are too few experience reports and experimenting and researching in real systems are too expensive. In addition, 66% of IT managers do not have time to put together a basic cost model or a data value model for ILM projects. (Foskett, 2006). This paper addresses these issues and contributes to supporting and assisting IT managers in their decision-making process.

We presented a cost model for ILM in an enterprise. The model was applied to ILM simulations. This led to design guidelines for ILM environments. The results raise the awareness of the required number of hierarchies and the choice of storage technologies.

One example showed a cost potential of 43%. The generalization of the example identified dependencies between the hierarchies. These dependencies influence the design guidelines.

The results were explicitly:

Result 1: The number of storage hierarchies chosen should be as low as possible. The more hierarchies used, the lower the cost reduction effect of additional hierarchies. In addition, this increases the administrative expenditure.

Result 2: Technologies of different storage hierarchies shall differ sufficiently in price (e.g. 1:6).

For tapping cost potentials it is necessary to obtain a broader knowledge about ILM procedures and scenarios. In this respect ILM nowadays lacks methods and tools.

The model presented helps to gain experience on cost potentials. This can help to avoid mismanagement and poor investments in advance.

Further work will be done in analyzing the real storage environment on their individual ILM cost potentials.

## REFERENCES

1. Beigi, M., Devarakonda, M., Jain, R., Kaplan, M., Pease, D., Rubas, J., Sharma, U., Verma, (2005) A., Policy-Based Information Lifecycle Management in a Large-Scale File System. In Proceedings of the Sixth IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY'05) pp. 139-148
2. Chen, Y. (2005) Information valuation for Information Lifecycle Management. In: Proceedings of the Second International Conference on Autonomic Computing (ICAC'05), pages 135-146.

3. Foskett, S. (2006) Best Practices – Belt tightening continues to shift budgets away from security and information management, TECHTARGET  
[http://searchwebservices.techtarget.com/magPrintFriendly/0,293813,sid35\\_gci1179438,00.html](http://searchwebservices.techtarget.com/magPrintFriendly/0,293813,sid35_gci1179438,00.html)
4. Gibson, T. and Miller E. (1999) An Improved Long-Term File-Usage Prediction Algorithm. 25<sup>th</sup> Annual International Conference on Computer Measurement and Performance (CMG 99), Reno, NV, USA, pages 639-648.
5. Gostner, R., Turczyk, L., Berbner, R., Heckmann, O. and Steinmetz, R. (2005) „Analyse von Datei-Zugriffen zur Potentialermittlung fuer Information Lifecycle Management“, TU Darmstadt KOM Technical Report 01/2005
6. Groepl, M., Turczyk, L., Heckmann, O. and Steinmetz, R. (2006) „Statistische Datenanalyse von langfristigem Dateizugriffsverhalten in einer Unternehmensdatenbank fuer ILM“, TU Darmstadt KOM Technical Report 01/2006
7. Kraemer, T. (2001) The Storage Report - The customer Perspectives & Industry Evolution, McKinsey & Company,
8. Linden, L. (2006) Storage Budget Survey 2006, Glasshouse Technologies Inc. <http://www.glasshouse.com/storage-budget-survey/registration.shtml>
9. Lyman, A., Varian, B. (2003) How Much Information? , University of California <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>
10. Merrin, R., Harnetty, D. (2005) Compliance Confusion Set to Drive Market in 2005, Engenio Information Technologies, <http://www.engenio.com/news/uk/2005/02-23-700.html>
11. Merrill, D. (2003) Storage Economics Identifying and Reducing Operating Expenses in the Storage Infrastructure Hiatchi Data Systems, WHP-153-00, November 2003
12. META Group, (2005) META Trends 2005/2006 — Infrastructure, Trends #2059
13. Miller, E. L., Katz, R. H. (1993) An Analysis of File Migration in a UNIX Supercomputing Environment. In: USENIX Winter, pages 421–434, 1993.
14. Moore, F. (2003) Storage - New Game New Rules, Horison Information Strategies,
15. Nguyen, H. (2005) Worldwide Disk Storage Systems Quarterly Tracker, IDC's March 2005  
<http://www.energy-insights.com/getdoc.jsp?containerId=prUS20079206>
16. Page, L., Brin, S., Motwani, R. and Winograd. T. (1999) The pagerank citation ranking: Bringing order to the web.  
<http://dbpubs.stanford.edu:8090/pub/1999-66>.
17. Paquet, R. (2004) Why You Need a Storage Department, Gartner Research, June 2004
18. Paquet, R. (2005) Poll Confirms Companies Aren't Ready for ILM. Gartner Inc. #G00125956
19. Patel, C. D., Shah, A. J. (2005) Cost Model for Planning, Development and Operation of a Data Center. Internet Systems and Storage Laboratory HP Laboratories Palo Alto, HPL-2005-107(R.1), June 9, 2005
20. Peterson, M (2004)., ILM Definition and Scope - An ILM Framework, SNIA Data Management Forum, Version 2.3, July 2004
21. Reiner, D. Press, G., Lenaghan, M., Barta, D., Urmston, R. (2004) Information Lifecycle Management: The EMC Perspective. Proceedings of the 20th International Conference on Data Engineering (ICDE'04) 1063-6382/04
22. Short, J. E. (2006) ILM Survey: What Storage, IT and Records Managers Say. ISIC Research Report Vol. 06, No. RB06-02, Information Storage Research Center, University of California.
23. Strange, S. (1992) Analysis of Long-Term UNIX File Access Patterns for Application to Automatic File Migration Strategies. Technical Report UCB/CSD-92-700, EECS Department, University of California, Berkeley.
24. Sun Microsystems (2005) Storage Optimization – ILM. March 2006”
25. Tanaka, T., Ueda, R., Aizono, T., Ushijima, K., Naitih, I., Komoda, N. (2005) Proposal and Evaluation of Policy Description for Information Lifecycle Management Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'05) 0-7695-2504-0/05
26. Turczyk, L., Berbner, R., Heckmann, O., Steinmetz, R. (2006) A formal approach to Information Lifecycle Management. In Proceedings of 17th Annual IRMA International Conference, Washington D.C.