Modeling Archival Repositories for Digital Libraries*
Extended Version

Arturo Crespo, Hector Garcia-Molina
Computer Science Department
Stanford University
{crespo,hector}@db.stanford.edu

Abstract

This paper studies the *archival problem*: how a digital library can preserve electronic documents over long periods of time. We analyze how an archival repository can fail and we present different strategies that help solve the problem. We introduce *ArchSim*, a simulation tool that for evaluating an implementation of an archival repository system and compare options such as different disk reliabilities, error detection and correction algorithms, preventive maintenance, etc. We use ArchSim to analyze a case study of an Archival Repository for Computer Science Technical Reports.

KEYWORDS: digital archiving, digital preservation, archival repository, models for archival repositories, performance of archival repositories, simulation of archival repositories.

1 Introduction

A continual threat to digital libraries is the loss of documents. Digital information can be lost not just through magnetic decay in storage devices, but also because of format and device obsolescence. This problem will only get worse as more and more information is provided only in digital form. The solution is an *Archival Repository* (AR), a system capable of storing and preserving digital objects (e.g., movies, technical reports) as technologies and organizations evolve [5]. An AR must preserve not only the bits, but also the "meaning" of its documents [10]. For instance, to preserve a "postscript" technical report, the AR needs to maintain the postscript bits, metadata indicating that this document is postscript, a program that can interpret and render postscript, and an environment that can execute the rendering program.

One of the main challenges in designing an archival repository is how to configure the repository to achieve some target "preservation guarantee" while minimizing the cost and effort involved in

*This work was partially supported by NSF.

running the repository. For example, the AR designer may have to decide how may sites to use, what types of disks or tape units to use, what and how many formats to store documents in, how frequently to check existing documents for errors, what strategy to use for error recovery, how often to migrate documents to a more modern format, and so on. Each AR configuration leads to different levels of assurance, e.g., on the average a document will not be lost for 1000 years, or in 1000 years we expect to still have access to 99% of our documents. Each configuration has an associated cost, e.g., disk hardware involved, computer cycles used to check for errors, or staff running each site.

The number of options and choices is daunting, and the AR designer has few good tools to help. The traditional fault tolerance models and techniques, of the type used to evaluate hardware, are a helpful starting point, but they do not capture the unique complexities of ARs. For example, traditional models may have difficulty capturing different document loss scenarios (e.g., missing interpreter, missing bits, missing metadata) and they frequently assume failure distributions (e.g., exponential) that are too simplistic.

In this paper we present a powerful modeling and simulation tool, *ArchSim*, for helping in AR design. ArchSim can model important details, such as multiple formats, preventive maintenance, and realistic failure distribution functions. ArchSim is capable of evaluating a large number of components over very long time periods. ArchSim uses specialized techniques in order to run comprehensive simulations in a time frame that allows the exploration and testing of different policies.

Of course, no model can be absolutely complete: there is an intrinsic tradeoff between how detailed the model is and the complexity (even feasibility) of its study. In our case, we have chosen to ignore (at least in this initial study) information loss due to format conversion (migration to a new format is always successful and does not introduce any loss). We also do not model partial failures (a failure where we can salvage part, but not all, of the information in a device). As we will see, we also rely on an "expert" that can provide failure distributions for the components of the systems. For instance, if format obsolescence is an issue, an expert needs to give us the probabilities of different format lifetimes.

In summary, our contributions are:

- A comprehensive model for an AR, including options for the most common recovery and preventive maintenance techniques (Sections 2, 3 and 4).
- A powerful simulation tool, ArchSim, for evaluating ARs and for studying available archival strategies (Section 5).
- A detailed case study for a hypothetical Technical Report repository operated between two universities. Through this case study, we evaluate AR factors such as disk reliability, handling of format failures, and preventive maintenance.

2 Archival Problems and Solutions

As a first step in modeling and evaluating archival repositories (ARs), it is important to understand how information can be lost, and what techniques can reduce the likelihood of loss. Because of space limitations, we limit ourselves to a brief review of failure sources and avoidance techniques.

2.1 Sources of Failures

We define an *Archival Repository* (AR) as a data store that gives a guarantee of long-term survivability of its collection. We will refer to each unit of information in the AR as a *Archival Document* or just document for short.

The archival guarantees of an AR should be declared in a precise way. One common way of expressing the archival guarantee is through the *mean time to failure* (MTTF), this is, the mean time before the archival repository loses a document.

An AR fails to meet its guarantee when it loses information. Such a loss may be caused by a variety of undesired events, such as the failure of a disk, on an operator error. A document is lost, not just if the bits that represent it are lost, but if the necessary components to give meaning to those bits are lost. For example, to render a document in human-readable form, we may need both the bits and a particular viewer. The AR will fail if the repository does not have a copy of the bits or of the viewer. We define the *components* of a document to be all the resources that are needed to support access to the document. Examples of document components are the bits that represent the document, the disk that stores the bits, and the viewer that interprets the bits.

An undesired event does not necessarily cause information loss. For instance, if the AR keeps two copies of a document, and the disk holding one of the copies fails, then the document is not lost. It would take a second undesired event affecting the second copy to cause information loss. When the survivability guarantee of a document is decreased due to an undesired event, we say that the document has become *damaged*. Damaged documents should be *repaired* by the system and have their archival guarantees restored.

The most common undesired events that may lead to the loss of a document can be broken down into the following categories. (We do not consider transient failures (e.g., a power failure) that do not lead to a permanent loss.)

Media decay and failure: Natural decay processes may make media unreadable. A well known example is magnetic decay, occurring when media loses its ability to hold the "bits" that encode a document. In Figure 1, we summarize typical decay times for electronic media. Numbers in the figure are from the National Media Lab [3]. They all assume "good" storage and operating conditions, and infrequent accesses. Note that decay typically does not affect all parts of a tape or disk simultaneously. For example, a few bits may first decay, rendering a document (or part of a document unreadable), followed by other decays. Figure 1 gives times to the first decay.

Media may also fail when it is being accessed, e.g., a tape may break when in a reader, or a disk

Media	Decay Time
Magnetic Tapes	10-20 years
CD-ROM	5-50 years
Newspaper	10-20 years
Microfilm	100-200 years

Figure 1: Typical Media Decay Times

head may crash into a platter. The likelihood of failure increases the more frequently the media is accessed. The expected lifetime of a hard disk that starts and stops spinning frequently may be reduced to 5 years, from the 10 to 20 years of a inactive disk. Magnetic tape wears out with usage much more than a disk because it actually touches the reading mechanism.

Component Obsolescence: A document may be lost if we do not have one of the components necessary for supporting access to it. To illustrate, we describe two important forms of component obsolescence: *media obsolescence* and *format obsolescence*. Media obsolescence occurs when we do not have a machine that can read the media (even if the media is still readable). For example, if we have information stored in a 8 inch diskette and we do not have a drive and corresponding software drivers, then we have a media obsolescence fault. Format obsolescence happens when the system does not know how to interpret the bits that are encoded in the media. This kind of failure is the result of the non-self-describing nature of information stored in electronic media. That is, the stored bits must be transformed into a human comprehensible format by a process that cannot be inferred from the original bits. For example, say we have a postscript document. To view it, we must (a) know that its format is some particular version of postscript, (b) have a program that can interpret that version of postscript and display it on a screen (or printer), and (c) have a computer that can execute the code for the program. If we do not have any of these components, then we have format obsolescence.

Human and Software Errors: Documents can be damaged by human errors or by software bugs. For example, a person or a program may delete a document (e.g., by reformatting the disk that holds the document), or may improperly modify the files and data structures that represent the document (e.g., by writing random characters in the middle of the document). Reference [12] suggests that humans and software are the most serious sources of failures. Better technologies and designs are reducing media decay and component failure rates, but there is no sign of a drop in the rate of faults created by software and human errors.

External Events: Events external to the AR, such as fires, earthquakes and wars, may of course also cause damage. Such events may cause several AR components to fail simultaneously. For example, a flood can destroy a collection of disks at one site. If a document was replicated using those disks, all copies will be lost.

2.2 Component Failure Avoidance Techniques

There are two well-studied types of techniques for avoiding an AR failure. We can either attempt to reduce the probability of component faults, or, we can try to design the AR so those faults do not result in a document loss. In this subsection, we review techniques in the first category, while the next subsection will cover the second category. We organize the presentation in this subsection using the taxonomy of faults presented in Section 2.1.

2.2.1 Media decay and failure

A fundamental decision on designing an AR is which media will be used. In this subsection, we first outline how can we choose the appropriate media for an AR. Then, we present strategies that allow reducing the rate of decay and failure for that media.

The question of which media to use involves two decisions. First, we need to decide on the access properties that will be needed from the media. Second, we need to find the most cost-effective technology that provides those access properties. By access properties, we mean the characteristics of the media such as: how fast can we start reading the media, how fast can we find a document, if we can rewrite a document without having to rewrite the whole medium, etc.

The reason we need to consider the access properties is that they will affect how we design our system. For example, the time that it takes to start accessing a medium will have an impact on the available strategies to reduce Media Decay. Specifically, on-line media (i.e., media with very little initial delay) allows for easy detection of decay, as it is cheap to access all the information periodically. On the other hand, detecting decay on off-line media is more complicated because it is expensive to load and read the media. Nevertheless, we need to recognize the tradeoffs between different kinds of media, as off-line media tends to have lower cost of maintenance and a longer expected life than on-line media.

In terms of underlying technologies, there are many alternatives for archival media. In our work, we have concentrated on "digital" media; however, the boundary between digital media and paper has been made diffused with some interesting new media that have been developed recently. For example, the "PaperDisk" technology uses 2D barcodes to print electronic information on paper [1]. The barcodes can then be scanned back into the computer and transformed into digital information. By using special algorithms, this technology can save up to 1Mb of information on a sheet of paper.

The most important factor in choosing media technology is its lifetime. Evaluating the lifetime of media technologies is a complicated process. Media Lifetime is nearly impossible to find in practice (cannot wait "n" years for media to fail). Therefore, estimates are based on the physical properties of the medium. The challenge is to find which are the relevant properties (so we can find the weakest link) and what is the appropriate model of decay for those characteristics. For example, a CD-ROM is basically a reflective layer sandwiched between a clear substrate and a lacquer film. A CD-ROM can fail because of either a deterioration of the reflective layer or because of the lost of optical clarity of the substrate. Deterioration of the reflective layer can happen because of oxidation, corruption, or delamination. As these three sources of deterioration are well known process for the materials used in the reflective layer, then we can predict when they will deteriorate and eventually fail. Similarly, we can study the causes for losing optical clarity of the substrate and try to predict when this failure will occur.

Beyond the choice of which media to use, we need to ensure that the media is stored in the best conditions to ensure its preservation. In the case of tapes, a controlled environment with low temperature and low humidity may increase the life of the tape in an order of magnitude [4].

2.2.2 Component Obsolescence

Avoiding component obsolescence is a complex problem as an accurate prediction of what operating systems, document formats, and devices will be used in the future is impossible. However, there are some techniques that help avoid the component obsolescence problem. In this subsection, we will briefly explore three of these techniques: usage of standards, self-contained media, and preservation of equipment. Later in the next section, we will study system techniques that can help with the problem of component obsolescence.

One proposed solution to the component obsolescence problem is the usage of well-established standards. Standards are (usually) well documented, which will allow a future party that is interested in a document formatted in that standard to re-create a reader for it. There are some problems with standards. First, there is a bootstrap problem; a person re-creating a reader needs to be able to read the documentation for the standard. Second, an important risk when using a standard is that many documents that claim to follow a standard are not really 100% compliant. Take, for instance, HTML documents, most of them do not implement HTML correctly. A browser that is built to accept only correct HTML will not be able to display most documents on the web.

Another solution to the obsolescence problem is the preservation of equipment. This approach has been attempted by the National Media Lab, but is a difficult and costly task, especially when parts are not longer available to repair preserved equipment.

Another interesting proposal for solving this problem is using self-contained media such as the electronic tablet [13]. The electronic tablet is a portable computer that contains all the necessary hardware and software (as well as data) to make a document readable. The only needed input for this tablet is a power source. The author claims that the electronic tablet also solves the component obsolescence problem (as everything needed to read the document is in the tablet). However, no tablets have been built and it is yet to be determined if the high cost of the tablet justifies the added benefits.

2.2.3 Human/Software Errors

As stated before, human and software errors are the fastest growing source of faults in computer systems. Beyond traditional techniques such as program validation, and user training, some specific guidelines are useful in an Archival Repository. These guidelines include:

- *Define interfaces that minimize the amount of damage that can be done:* For example, by not allowing deletions or updates in the Archival Repository, users can not make mistakes that will result in the deletion of a file. In addition, by defining a restricted interface, it is easy to distinguish between faults and the normal operation of the system. For instance, if a file disappears in a system that does not allow deletions, then we know that a fault has happened.
- *Minimize the amount of code that can make damage:* The assumption is that the less the amount of code, the better the chance that it can be extensively tested and verified. For example, if we can restrict access to a storage device only to calls made through a small piece of code. This code can implement a clean and safe interface for other programs and its behavior can be verified.
- *Use hardware safeguards to prevent software errors:* Hardware can help prevent damage to documents. For example, by opening the "read-only tab" in a diskette, we instruct the hardware to disallow any writes to it. Another commonly used hardware protection is virtual memory that can be used to prevent unauthorized writes to some memory addresses.

2.2.4 External Events

A range of techniques can be used here to prevent damage from external or environmental events. These techniques include, for example, fire-proof walls, earthquake proof building, etc. These avoidance techniques are beyond the scope of this paper and will not be discussed further.

2.3 System-Level Techniques

2.3.1 Migration

Migration is the movement of documents from and old component to a new one before failure occurs in the old component. For example, if we know that, let us say, that "Postscript" readers are becoming scarce, and that a new format, let us say "Postscript II" is taking the place of Postscript, then we may decide to migrate Postscript documents into the new format before Postscript readers are not available any more.

Migration techniques can be classified in a continuous, ranging from passive strategies to active strategies. In Figure 2 we show the spectrum in the context of reducing media decay. In the figure, we can see that passive strategies include reducing the media decay time by either using better media or storing media in a "controlled" environment. Active strategies attempt to reduce media decay by reading and copying the data periodically. When the copies are done over the same media, the strategy is called *periodic refreshing.* When the copy are done to different media, then the strategy is usually called *periodic* migration. There are many tradeoffs between active and passive strategies. In general, passive strategies have the advantage of lower use of resources than active strategies, but, unfortunately, they also provide lower probability of preservation. Active strategies have a high probability of preservation, but they consume more resources. This last point is one of the reasons why archival models are needed; designers need to take strategy decisions that achieve their *target* archival guarantee while reducing the total amount of resources consumed.

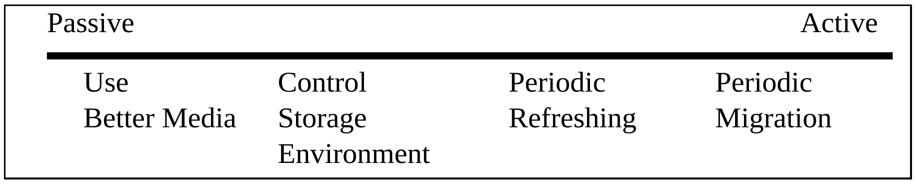

Figure 2: Reducing Media Decay

Finally, migration can also be used to avoid component obsolescence. However, performing a migration process that does not lose some information or formatting is frequently impossible. For example, consider translating an ancient text from Greek to Latin, and then from Latin to English. Most people will agree that if the Greek version is available, a direct translation from Greek to English will be better.

2.3.2 Replication

Replication allows survival in the presence of faults by making extra copies (or duplicating the resources) needed to access an archival document. For example, to prevent media failure, we can write the document in multiple tapes. If one of the tapes is damaged, then we can use the remaining tapes to access the document. Replication can also be used to survive component obsolescence. For example, a document can be written in multiple formats; in this way, if a format becomes obsolete, we can still access the document in another format.

Unfortunately, replication by itself does not improve the MTTF of a system. In [12], it is shown that the MTTF of some systems *decreases* when using replication, making failures more likely. The solution to this dilemma is to add *repairs* to the replication system. Specifically, after a failure has been detected, we should, as fast as possible, repair the failure by either fixing the problem in the failed replica or by generating a new one. For example, a failvote system with three replicas, each of them with a 1 year MTTF, will have a MTTF of 0.8 years without repairs. But if failed replicas are repaired within 4 hours, the MTTF of the system increases to more than 10^6 years [12].

An additional advantage of replication is that it simplifies failure detection. If we keep three copies of a document in three different tapes and one of the copies is different, then we can assume that the copy that is different has failed and should be discarded.

Finally, when using replication, we have to be especially careful with "correlated" faults. Correlated faults are especially common with faults from software, human, and external factors. For example, if all copies are stored in the same location, a natural disaster may destroy all of them.

2.3.3 Emulation

The emulation solution to the Archival Problem involves re-creating all the components needed to read an archival document in a new platform. Emulation can be done at several levels. We can emulate the hardware, the operating system, the software, or the formats, in a new system. The computer game community has done an interesting use of emulation. Many "platform" emulators are available that allow to run old computer games in modern computers.

The main advantage of emulation is that it can potentially improve the scalability of archival systems. With all other system solutions, we need to concern ourselves with each possible component independently of each other. Emulation recognizes that components are frequently stacked on top of each other (e.g., to read a document, we need to run a software, that needs some Operating System, which in turn needs some specific hardware). By choosing an appropriate level, let us say the Operating System level, we do not need to worry about components higher than this level. In other words, if we produce an emulator of Windows 98, we do not need to do anything special in order to use all the programs currently available for that platform. In addition to this benefit, emulation better preserves the "look and feel" of the original application, which may be important in some fields.

The major disadvantage of emulation is its feasibility. Emulating a piece of software/hardware can be a tricky and difficult task that requires a lot of engineering and collaboration from the current producers of the component. Additionally, the task may infringe industrial property and copyright laws. In addition, the benefits of emulation may not materialized in many situations. For example, if we have a relatively uniform collection of documents, the cost of producing an emulation system will probably be much higher than the cost of using replication. In addition, preserving the "look and feel," while important in some context, it is negative in others. We do not want to preserve the "feeling" of waiting a couple of days for the answer to a query when accessing an archival database.

2.3.4 Fast Failure Detection and Repair

In order to enhance most system techniques, we need to check periodically for failed documents. A fast failure detection and repair allows for higher preservation guarantees. For example, if we detect that one of the replicas has failed, we may decide to create a new replica to maintain the preservation guarantees of the system.

In general, it is impractical (and sometimes impossible) to detect exactly when a fault in a component has occurred. This is because the manifestation of the fault, an error, might only happen long after the fault happened. For example, a tape can deteriorate and lose information (a fault), but we will not know that the tape has failed until we attempt to read the data and we get an error. Testing for errors may be a difficult task. A typical system has a very large number of components and it would take a long time to check all of them looking for errors. In order to reduce the time spent in finding errors, we can use sampling (by time or component type). We can also use partial checks or predictors to select the components that have a high probability of failure.

- *Time-based samples:* Documents are probed a specific intervals of time. The goal is that the probability between probes of losing a document due to multiple failures is small. Obviously, the problem with this technique is that we need an accurate model of the faults, so probes are frequent enough so the assumption is true, but not too frequent that we are wasting resources. In the experiments section of this paper, we will see, that time-based samples need to be done much more often than the MTTF of the component in order to achieve a good MTTF for the Archival Repository.

- *Component-based samples:* This technique probes components instead of individual documents. This approach is especially useful when the number of components is much smaller than the number of documents. Specifically, the system will perform a test on the component, under the assumption behind that if the component passes the test, then it is most likely that the component will successfully work for all documents. A way to perform the test is to select one document (or for some small number of documents) that use a component and try to access them; if we succeed, then the component is considered operational, if not, then we have detected a fault. Specifically, if we are testing the postscript renderer, we can select a document and attempt to display it, if we succeed, then, when using this technique, the system will assume that it will be able to render all documents stored in that format.

- *Partial checks:* This technique probes a small part of the documents. Here we are assuming that if "part" of the document is not damaged, then the whole document is undamaged. For example, if we have a Postscript document, we can just try to print the first page; if we succeed, then we can assume that the rest of the pages can be printed too. Similarly, to test a disk, we only check a few blocks; if those blocks are undamaged, then we assume that the whole disk is undamaged.

- *Predictors:* We can use "indirect" information to predict the likelihood of a fault in a component. For example, if a new document format has been released as a replacement of a format in which we have documents stored; we can use this information as a predictor of the obsolescence of the old format. Similarly, some hardware systems provide warnings that we are approaching the end of the life of the device. For example, Compaq's Intellisafe [8], and Seagate's S.M.A.R.T. [15] measure several disk attributes, including head flying height, to predict failures. The disk drive, upon sensing degradation of an attribute, such as flying height, sends a notice to the host that a failure may occur. Upon receiving notice, users can take steps to protect their data.

After we have detected that a failure has happened or if we are predicting an imminent failure, we should repair the component as soon as possible. The longer we wait for the component to be repaired, the higher the risk that another failure will lead to information loss. There are many standard techniques for fast repair, such as having spares available, having parts available and training of repair personnel.

3 Architecture of an Archival Repository

Our goal in this section is to identify the elements of a typical archival repository (AR), so we can model each element and study how it impacts reliability. A typical AR stores documents in a *data store* that can fail. The AR can still achieve long-term survivability by enhancing the data store with an *archival system* (AS) that implements some of the techniques of Section 2.3. We present our AR model in Figure 3. The figure shows the AS modules (in solid-line boxes), the non-fault-tolerant store (in a dashed-line box), and the archival documents. The arrows represent the runtime interactions between the elements.

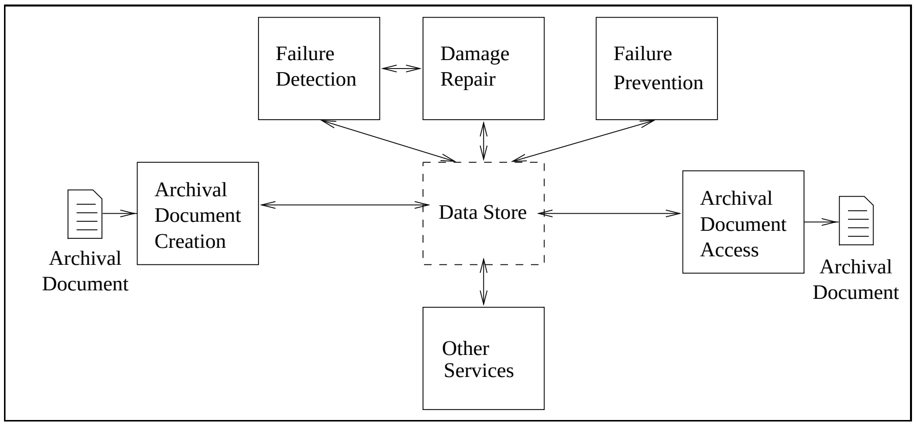

Figure 3: Archival Repository Model

3.1 Archival Documents

In our architecture, an archival document embodies information. An archival document cannot be just a bag of bits, but it must also include all the components necessary to transform the bits into a human comprehensible form.

An archival document is an abstract entity. The connection between document and access to it is achieved through *materializations.* A materialization is the set of all the components necessary to provide some sort of human access to a document. For example, a materialization may include the bits, disks, and format interpreters necessary to display a technical report. The same technical report may be accessible through a different materialization, that may include a different format interpreter to print the technical report.

We illustrate materializations in Figure 4. In the figure, there are two archival documents. Each of those documents has two different materializations. For example, Materialization 1 requires the following components to be available: File 1, Site A, Disk 1, and a ASCII printer. Incidentally, notice that File 1 is stored on Disk 1, which in turn is at Site A. Such component interdependencies will be discussed later, when we model materialization failures.

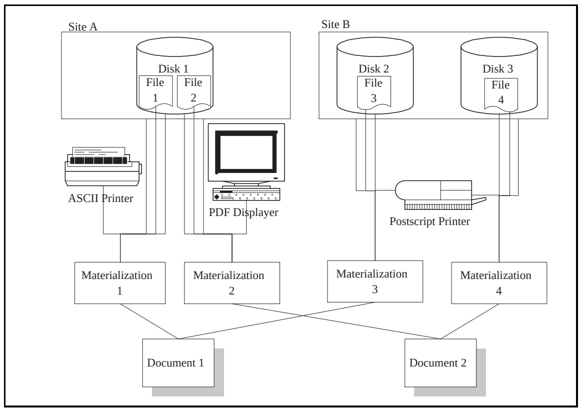

Figure 4: Materializations

The AR is able to preserve documents by preserving the materializations and their components. In this paper, we treat the documents as “black boxes.” We do not attempt to take advantage of document structure (e.g., chapters in a book).

3.2 Architecture of the Non-Fault-Tolerance Store

The *Store* encompasses the set of components, such as sites, disks, or format interpreters that make materializations accessible. Because the store is not fault tolerant, materializations may be lost. A materialization is considered lost when *any* of its components has failed. If *all* of the materializations of a document are lost, then the document is considered lost.

To create a materialization, first we must ensure that the necessary components exist in the store. Then we create a record that links together the components as a materialization. For example, say we want to create a document materialization that requires the bits in file “doc.ps”, which are located in $disk_2$ in $site_1$ and requires the postscript interpreter. Any components that do not exist already (e.g., the file doc.ps) are created. In some cases, “creating” a component may require a physical action, e.g., adding a new printer or disk to the store. Once the components exist, a metadata record containing references to the components is created (e.g., $\langle site_1, doc.ps, disk_2, postscript\rangle$). This record serves as the identifier for the materialization.

A document metadata record includes the records for all available materializations. Note that a document is not accessible if its metadata record is corrupted or lost. Therefore, the record must be one of the required components for any materialization.

3.3 Architecture of the Archival System

The AS provides fault tolerance by managing multiple materializations for each document. The AS monitors these materializations, and when a failure is detected, attempts to repair them. The AS can improve fault tolerance further by taking preventive actions to avoid failures. The AS provides the user the ability to create and retrieve archival documents. It also provides miscellaneous services such as indexing, security, and document retirement, among others. When a user requests a document, the AS uses its metadata to find all the available materializations of that document, selects one and returns it to the user. In this section, we describe the six modules that make up at AS (see Figure 3).

The *Archival Document Creation module* (ADC) generates new documents, implementing policies on the number and types of materializations that are needed. For example, say an administrator has decided that documents should be materialized as illustrated by Document 1 in Figure 4. Then, for each new document, the ADC will create (on the data store) the appropriate "File 1" and "File 2," the document metadata record, and will check that the other components exist. The main objective of this paper is to provide a framework that allows a system administrator to choose the best materialization policies to achieve a desired level of reliability.

The *Archival Document Access module* (ADA) services request for documents. Basically, the module translates the request for a document into a request for the appropriate components of one of the document materializations.

The *Failure detection module* (FD) scans the store looking for damaged or lost materializations. When a damaged materialization is found, the failure detection module informs the Damage Repair module (described below) about the problem.

The *Damage Repair module* (DR) attempts to repair damaged documents. There are many strategies to repair a damaged document, as discussed in Section 2.3. The input of the DR module is a signal from the FD module.

The *Failure Prevention module* (FP) scans the store and takes preventive measures so materializations are less likely to be damaged. For example, the FP module may copy components that are stored on a disk that is close to the end of its expected life, into a newer disk.

Finally, the *Other Services module* (OS) provides miscellaneous services such as indexing, security, and document retirement. Retiring a document involves removing from the store any components that are no longer needed, even by other materializations.

4 Failure and Recovery Modeling

In this section, we will model the failure and recovery characteristics of an AR, based on the architecture presented in the previous section. First, we will explore how to model a non-fault-tolerance store, and then the archival system (AS). Later, we will combine these two models into an archival document model.

4.1 Modeling a Non-Fault-Tolerance Store

To model the failure characteristics of a store, we start with an abstract representation of materializations and components. We model a materialization as an n-tuple $\langle mat_{id}$, doc_{id}, $comp_1$, $...comp_n\rangle$; where mat_{id} is the materialization identifier, doc_{id} is the document identifier, and $comp_1...comp_n$ are the components required to provide the required document access. The identifiers mat_{id} and doc_{id} together, form a unique id for the materialization. For example: $\langle M1$, $TR1233$, *doc.ps*, $site_1$, $disk_2$, $postscript\rangle$ means that the materialization $M1$ contains the document identified by $TR1233$ that needs the bits in file *doc.ps*, $disk_2$, $site_1$ and the postscript interpreter in order to be readable. A document can have more than one materialization. For example, Technical Report 1233 can also have the materialization $\langle M2$, $TR1233$, *doc.ps*, $site_1$, $disk_3$, $postscript\rangle$, which would be a copy of M_1 but on a different disk ($disk_3$).

We model components by a tuple $\langle component_{id}$, $type\rangle$, where $component_{id}$ is a unique identifier for the component instance and *type* is the class (e.g., file, disk, interpreter) to which the instance belongs.

To further model components, we need to describe:

- How many component instances and types are present in the system: this is, how many disks, formats, etc., are available.
- Failure distribution of each component type. Many components have two different failure distributions, one during archival and another during access. For example, a tape is more likely to fail when it is being manipulated and mounted on a reader than when it is stored. Therefore, each component will have two failure distributions: during archival (i.e., time to next failure when the component is not used) and during access. For some components, such as disks or sites, the access and archival distributions will be the same; but for other components, such as tapes or diskettes, these distributions can be very different.
- Time distribution for performing a component check. This distribution describes how long it takes to discover a failure (or to determine that a component is good), from the time the check process starts. For example, consider checking a tape. This may involve getting the tape from the shelf, mounting the tape, and scanning the tape for errors.
- Time distribution for repairing a component failure. This distribution describes how long it takes to repair a component. This distribution may be deterministic (if the component can be repaired in a fixed amount of time). Repair time may be "infinite" if the component cannot be fixed.

In addition, there is an important interdependency between components. Specifically, the failure of one component may cause the failure of another component. For example, if a site fails (e.g., because it was destroyed by a fire), then all the disks at the site will also fail. As pointed out earlier, we are only taking into account permanent failures; transient failures (e.g., the site was temporarily disconnected from the network) are ignored. This failure dependency is captured by a directed graph. For example, an arrow between "Site A" and "Disk 1" in the interdependency graph means that if "Site A" fails, then "Disk 1" will also fail.

We close this subsection with two comments. First, we do not claim that the model presented for the store is complete. For instance, we have not included policies for handling concurrent access. There is always a tradeoff between complexity of the model and our ability to analyze it. We believe that our model strikes a good balance in this respect, and captures the essential features of a store. Second, the reliability predictions we make are only valid for the *current* configuration of the repository. Over time, the repository will change (e.g., as new devices are introduced), so we may need to change our repository model. As the model changes, we may need to revisit our predictions.

4.2 Modeling an Archival System

In this section, we describe how to model the behavior of the modules of the archival system. We do not include failure distributions for these modules as we are assuming that the AS itself does not fail. We recognize that this is a strong assumption, but in this paper, we have chosen to concentrate on the failure of components, instead of on the failure of the system that provides fault-tolerance. How to develop error-tolerant robust software design have been study in [16].

Because of space constraints, we cannot describe each module separately. In general, we model the input of the modules with probability distribution functions and their behavior by algorithms. For example, consider the document creation (ADC) module. Its input distributions tell us how frequently requests for document creation arrive, how many materializations each new document will have, and which components will be selected to participate in a given materialization. The algorithms for the failure detection (FD) module spell out what policies are implemented, e.g., if all components are checked on a regular basis or not.

The probability distributions that drive the model can be obtained in different ways. If we have data from a real system, we can use the data directly (trace driven), or we can define an empirical distribution, or we can fit the data onto a theoretical distribution [2]. If we do not have real data, we need to choose a theoretical distribution that matches our intuition. A sensible distribution to choose (when requests are generated independently) is a Poisson distribution for event inter-arrival times [7].

We summarize the model parameters in Figure 5. The figure is divided in three parts. At the top are the parameters that describe the AR: the number of components and their types, and the failure dependency graph. Then, we list all the distributions needed for the model with the units being modeled in parenthesis. Finally, we list all the policies and algorithms that must be defined to model the archival system.

- **AR Description**
  - Number of components and types
  - Failure dependency graph
- **Distributions**
  - For each component type:
    - Failure distribution during access (time)
    - Failure distribution during archival (time)
    - Failure detection distribution (time)
    - Repair distribution (time)
  - Document creation distribution (time)
  - Document access distribution (time)
  - Access duration distribution (time)
  - Document selection distribution (document)
- **Policies**
  - Document Creation policy
  - Document to materializations policy
  - Failure detection algorithm
  - Damage Repair algorithm
  - Failure prevention algorithm

Figure 5: Archival Repository Model Parameters

4.3 Modeling Archival Documents

In this section, we combine the models for the data store and the AS, in order to describe the life of an archival document. In Figure 6, we depict our model for the life of an archival document. The life of a document starts when its materializations are created by the ADC module and handed to the store. The creation of a document may not be an instantaneous process. For example, if long-term survival is achieved by keeping multiple copies, the document is not considered archived until all the copies are generated. Once the ADC module has taken all the actions that ensures the long-term survival of the document, then the document has full protection, and we say that the document is in the *Archived* state.

When any of the document component fails (based on the distributions in Figure 5), the document is considered to be in the *Damaged* state and becomes temporarily unprotected. For example, if we keep two copies of a document and one of the copies is lost, then the document would be damaged. As explained earlier, the AS will not know that a document is damaged until the FD module detects the failure. When the failure is detected, the document goes to the *Damaged Detected* state.

When damage to a document is detected (by the policies summarized in Figure 5), the AS starts actions to restore the document and, hopefully, return it to the Archived state. For example, if the document is damaged because one of its copies was lost, the repository can just replace the damaged copy by creating a fresh one from one of the good copies. However, if the repair is not successful, then the document may be *Lost.* This later state is the one that we want to avoid in an archival system.

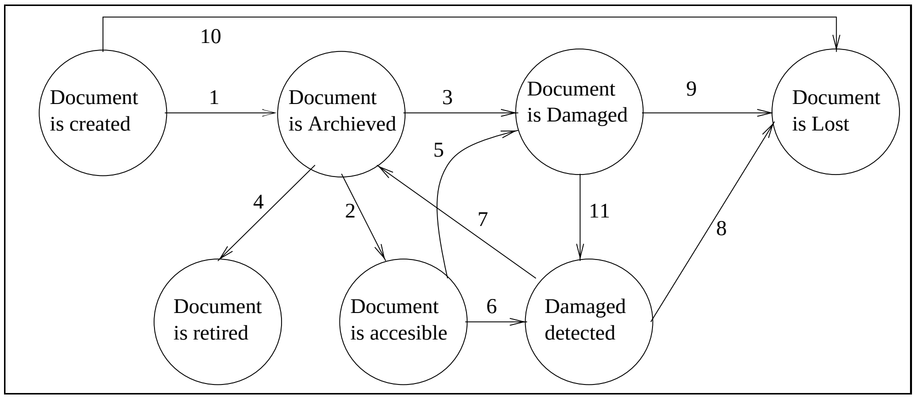

Figure 6: Archival Document Model

We also distinguish two additional states: Accessible and Retired. When a document is in the *Accessible* state, it can be accessed (e.g., read, printed) by users, which is not the case for the Archived state. For example, if some the document's components are stored on a tape which is kept in a safe, we need to take the tape out and mount it in a reading device to make it accessible. When the tape is stored the document is in the Archived state; when the tape is mounted, it is in the Accessible state. As we discussed in the previous section, when making the document accessible, in general, we are increasing the chances of damaging the document; so the probability of transition 7 is, in general, greater than the probability of transition 5.

The Retired state allows users to mark the document that are not needed anymore. In this case, the document is retired from the archival system and the system does not provide any survivability guarantees. It is important to note that retiring a document may eventually result in removing all materializations from the store. This action is different than taking a document "out of circulation," in which case the document is not longer available to regular users, but it is still preserved for historical reasons.

In the Archival Document Model, transitions between states are due to deliberate actions taken by the AR or due to external events. An example of a deliberate action is the creation of enough copies that lead to the transition between the Created state and the Archived state. An example of an external event is the failure a disk, causing the transition of a document from Archived to the Damaged. Most external events have a probabilistic nature.

The archival document model presented only contains the *main* states for a document. In some particular scenarios some of the states may merge or may not exist at all. For example, in an on-line archival system, there may not be any difference between the "Archived" and "Accessible" states. Similarly, depending on the actual policies of the AR, some of the states can be expanded. For example, if a system has a "Document to Materialization" policy that requires the creation of three copies of the document, the Damaged state expands into six states, one for each possible combination of one or two copies being lost. It is important to note that for analyzing an archival system we do not need to expand the nodes unless we are interested in the specific probability of being in those expanded nodes.

Also note that we have simplified our document model by assuming some transitions do not occur. For example, we do not show a transition directly from the Archived state to the Lost state, as we assume that documents are always damaged before they are lost. We also do not show a transition between the Damaged state and the Archived state, as we assume that damage does not repair itself.

4.4 Example: An Archival System based on Tape Backups

Let us describe the archival document model in more detail by using an specific example (Figure 8). This archival system saves documents in two tapes. The tapes are stored in a controlled environment (i.e., a safe) to improve preservation. We summarize the parameters for the model of this system in Figure 7.

- **AR Description**
  - Initial collection: empty.
  - Number of components and types: 2 tape drives
  - Failure dependency graph: empty
- **Distributions**
  - Tape Failure distribution during access: exponential with rate D
  - Tape Failure distribution during archival: exponential with rate L
  - Tape Failure Detection distribution : exponential with rate R
  - Tape Repair distribution: uniform with probability 1.
  - Document creation rate: n documents at startup, then no documents are created.
  - Document access period: A
  - Access duration period: S
  - Document selection: uniform over the n documents.
- **Policies**
  - Document Creation policy: write in both tapes.
  - Document to Materialization: read from any working tape.
  - Failure detection algorithm: complete scan of both tapes taking a negligible amount of time.
  - Damage Repair algorithm: discard bad tape and replace with new copy taking a negligible amount of time.
  - Failure prevention algorithm: none

Figure 7: A model for a dual tape system

The operation of the system starts with the ADC saving n documents in each of the tapes. For simplicity in this example, we will assume that this operation always succeeds. Let us assume that the time between access requests to the ADA module is distributed exponential with rate A. The duration of a successful tape access is also exponential with a rate $1 - S$. The FD module scans the tapes and finds all failures at a rate R. The DR module can repair all failures as long as one of the tapes is working and the repair is done instantaneously. Therefore, the system will lose the documents if both tapes fails in between the scans of the FD module. The system does not have an FP nor a OS module.

Additionally, let us assume that while the tape is in the safe environment, they become unreadable under a exponential distribution with rate L. When the tapes are being read, the rate of damage increases to D ($D > L$).

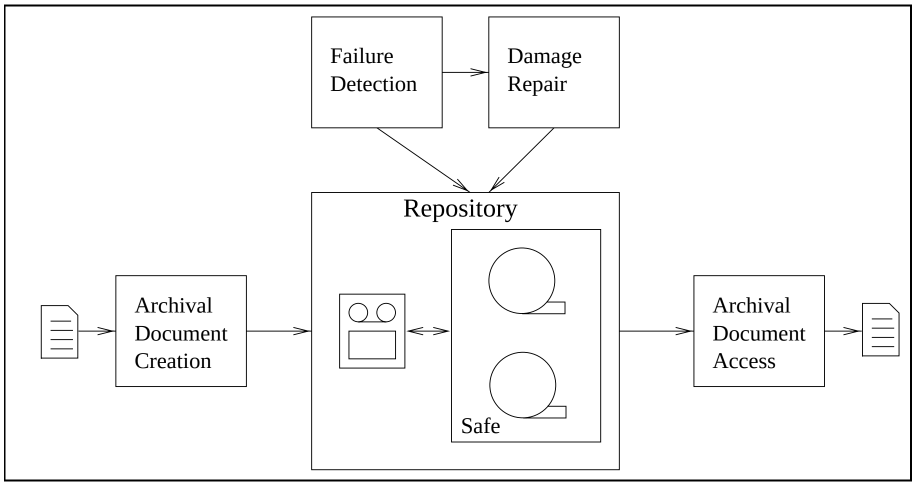

Figure 8: A Dual-tape Archival System

Using the information from the archival model, we can model the life of a document as it is presented in Figure 9. The document is considered Archived after we finished copying the document into the tapes and the tapes are in a safe place. As this process always succeeds, the transition to Archived happens with probability 1.

Given that we have two tapes, then the distribution for transition from Archived to Damaged will be exponential with rate $2L$ (if we assume independent failures). When making the document accessible, we increase the rate of failure of a tape to D, thus, the rate of the transition from Accessible to Damaged will be $D + L$, this is D for the mounted tape and L for the tape that stays in storage. Finally, when a tape is damaged, the system will attempt to recover the tape by using the other tape. We are assuming that this recovery will succeed with probability R, in which case we will return to the Archived state. However, if while in the Damaged state, the second tape becomes unreadable (this will happen with probability L), then the document would be lost. Note, that the rate of access to the document is A (not A/n, where n is the number of documents); this is because when accessing a document, we are actually accessing all documents in the tape.

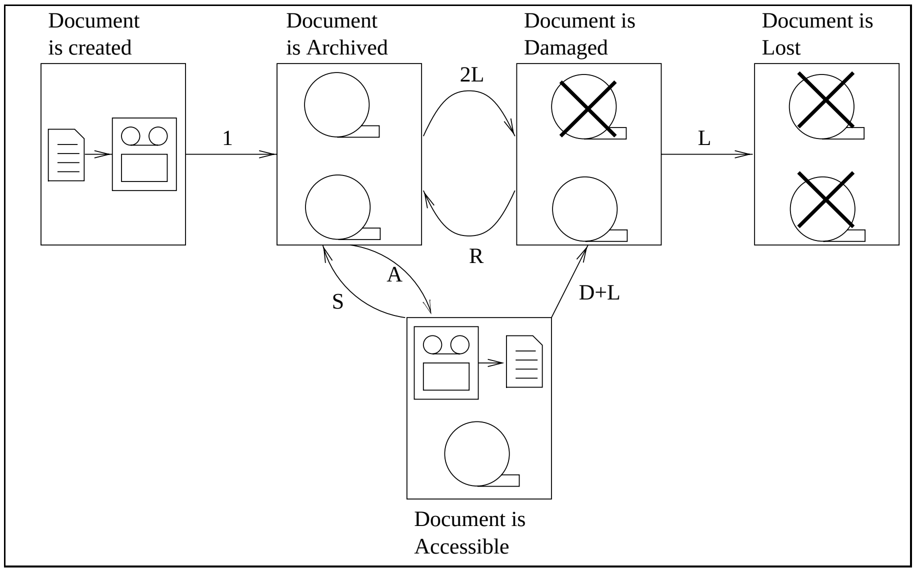

Figure 9: A Dual-tape Archival System

4.4.1 Analyzing the Dual-tape Archival System

The Dual-tape Archival System is simple enough that can be modeled using a Markov Chain and the Mean-Time to failure can be computed exactly. The critical question to answer about the Dual-tape Archival System, is how long would it take for the system to lose documents In other words, what is the mean time to failure (MTTF) of the system. We can use analyze the model of the system to answer that question. The computation of the MTTF can be done in two ways, we can attempt to compute it analytically, or we can infer it statistically by simulating the system. In some simple cases, like the Dual-tape Archival System, the analytical approach is possible. For other cases, the only option is to use a simulation. In the next section, we are introducing a simulation tool that allows the analysis of complex scenarios.

The Dual-tape Archival System is simple enough that can be modeled using a Markov Chain and the Mean-Time to failure can be computed exactly:

$$MTTF = \frac{S \cdot R + D \cdot R + D \cdot L + 2L^2 + 4L \cdot R + 3L \cdot S + 2A \cdot L + 2A \cdot R}{2L^3 + 2L^2 D + A \cdot L^2 + A \cdot D \cdot L + 2L^2 S} \tag{1}$$

A formal modeling allows to take precise allocation decisions. For example, if we use the parameters in Figure 10, then the $MTTF$ of the system will be 9,316 days (or about 25 years). We can also use this result to make resource allocation decisions. Let us suppose that we want to improve the $MTTF$ to 100 years and we can do it by improving the rate of detection and repair of damaged media (i.e., change the value of R). In this case, we can compute that the needed repair frequency to achieve a $MTTF$ of 100 years must be 13 days. This 13 days include the sum of the frequency of checking the tapes plus the number of days that it takes to replace a defective tape with a new copy. Specifically, if it takes 1 day to create a new tape, then we need to check all the tapes every 12 days.

Mean-Time to failure (days)	
Parameter	*value*
Tape in stable storage ($1/L$)	1000
Tape when reading ($1/D$)	250

Failure detection and Repair Rate (avg. days)	
Parameter	*value*
Tape ($1/R$)	60

Access Rates (days)	
Parameter	*value*
Access frequency ($1/A$)	20
Access duration ($1/S$)	2

Figure 10: Parameter values

The computation of Equation 1 is complex, even though we are not considering software errors, sites failure, etc. When trying to consider all those factors, trying to obtain a close result is very hard, and in some cases, impossible. In the next section we will introduce a simulation tool, ArchSim, that will be able to handle more complex cases.

5 ArchSim: A Simulation Tool for Archival Repositories

To evaluate a possible AR configuration, we need to predict how well it protects documents. This prediction can sometimes be done analytically, but as the AR gets more complex, an analytical solution is impractical (and sometimes impossible). Instead, we rely on a specialized simulation engine for archival repositories: ArchSim. We start this section by discussing the specific challenges confronted when simulating an AR. Then we describe ArchSim and its libraries.

5.1 Challenges in Simulating an Archival Repository

ArchSim builds upon existing simulation techniques for fault-tolerant systems. However, the unique characteristics of archival repositories make their simulation challenging:

- *Time Span:* The life of an archival system is measured in hundreds, perhaps thousands of years. This means that simulation runs will be extremely long, so special precautions must be taken to make the simulation very efficient. Furthermore, given these long periods, failure distributions must take into account component "wear-out." (A component is more likely to fail after 50 years that it is when new.) Simple failure distributions (e.g., exponentially distributed time between failures) are frequently used in fault-tolerant studies, but they cannot be used here since they do not capture wear out.
- *Repairs:* In an archival system we cannot in general assume that damaged components can always be replaced by new identical components (another common assumption when studying fault-tolerant systems). For example, after say 100 years, it may be impossible or undesirable to replace a disk with one having the same failure characteristics.

- *Component models:* Component models are fairly rich, compounding the number of states that must be considered. For instance, as we have discussed, a file is not simply correct or corrupted. Instead, it can be corrupted but the error undetected, it can be correct but not accessible for reads, and so on. The failure models in each of these states may be different, e.g., a file is more likely to be lost when being read.
- *Sources of failures:* A document can be lost for many reasons, e.g., a disk fails or a format becomes obsolete. Each of these failures has very different models and probability distributions. The approach of finding the "weak link" and assuming that all other factors can be ignored is not appropriate for ARs.
- *Number of Components:* An AR needs to deal with a large number of components and materializations. The challenge of simulating large number of objects has been studied extensively [11, 14] and ArchSim uses those results.

5.2 The Simulation Engine: ArchSim

ArchSim receives as input an AR model, a stop condition (stop when the first document is lost or when all documents are lost) and a simulation time unit (minutes, hours, days, etc.). ArchSim outputs the mean time to failure (mean time to stop condition), plus a confidence interval for this time. We are currently considering other output metrics, e.g., the fraction of the documents that are available after some fixed amount of time. However, these other metrics are not used in our case study (Section 6).

For defining the AR model, each distribution and policy is implemented as a Java object, so they can be as general as necessary. For example, for component repair, the corresponding Java module can simply use a probability distribution (perhaps one of the library functions described below) to generate the expected repair time. However, that module can easily be replaced by one that first decides if the component can be repaired (using one probability distribution), and then for each cases generates a completion time (when the component is repaired or the repair is declared unsuccessful).

5.3 Library of Failure Distributions

ArchSim makes available a library of pre-defined failure distributions, that can be used to describe AR components. The distributions in the library are: bathtub, infant mortality, historical survival, uniform, and deterministic. In Figure 11(a)-(e) we sketch generic versions of these probability distributions. (The area under these curves, from time 0 to t, represents the probability the component will fail by time t.)

1. Bathtub model: this model is based on the probability function of Figure 11a. This probability failure function is a typical electronic device. In the bathtub distribution, the instantaneous probability of failure early in the component's life (left on the horizontal axis) and late in its

life (to the right) are higher than during the middle years. The parameters of the bathtub model are: how long does the infant mortality phase last, mean failure during infant mortality phase, mean failure during the stable life of the component, when does the end of life phase starts, mean failure during the end of life phase.

2. Infant Mortality model: Figure 11b is the probability failure distribution in which the model is based. This model is good for describing failures due to System or Human errors. At first, when the document is recently created, the probability is high (as there could be errors in its creation and/or initial storage); then the probability drops sharply and remains about constant during the rest of the life of the document. The parameters of the infant mortality model are: how long does the infant mortality phase last, mean of the infant mortality phase, mean of the stable life of the component.

3. Historical survival model: this model attempts to describe the life of a published resource. The model is based on that idea that if a resource survives after a long period, it becomes of "historical significance," and its chance of continual survival increases. Figure 11c shows the probability distribution function for this model. Although there have not been studies modeling the obsolescence of formats, we theorize that this function is a good choice for modeling the obsolescence of a format interpreter. When the format is adopted and for a long time (and we are assuming here that we are adopting a "commercial" format), the probability of not being able to interpret the format is low. As the format ages, the probability of failure increases as knowledge about the format starts disappearing until we reach a point where the probability of losing the format, after having survived that far, starts decreasing. Note that for all the functions the probability of failure eventually drops to zero (perhaps asymptotically), so we do not end with probabilities greater than one. The parameters of the historical significance model are: how long does the adoption phase last, mean and standard deviation of the adoption phase, mean of the historical phase.

4. Uniform Model: This is a very simple model based on the uniform distribution of Figure 11d. Uniform distributions are convenient, as they are easy to generate and analyze. In many fault-tolerance system, uniform distributions are used instead of the distributions of Figures 11a-c under the assumption that components are checked before installing them in the system and that the life of the system is much shorter than the life of the components. However, in archival system, we should exercise extreme caution on using uniform distributions, as these assumptions are many times false. The parameter of the model is the mean of the component and how long is the expected life of the component.

5. Deterministic Model: This is a deterministic model where the event happens (with probability one) at a fixed time (where a spike is shown). It is useful for describing deterministic events such as the start of a failure detection process. The probability distribution function for this model is presented in figures 11e.

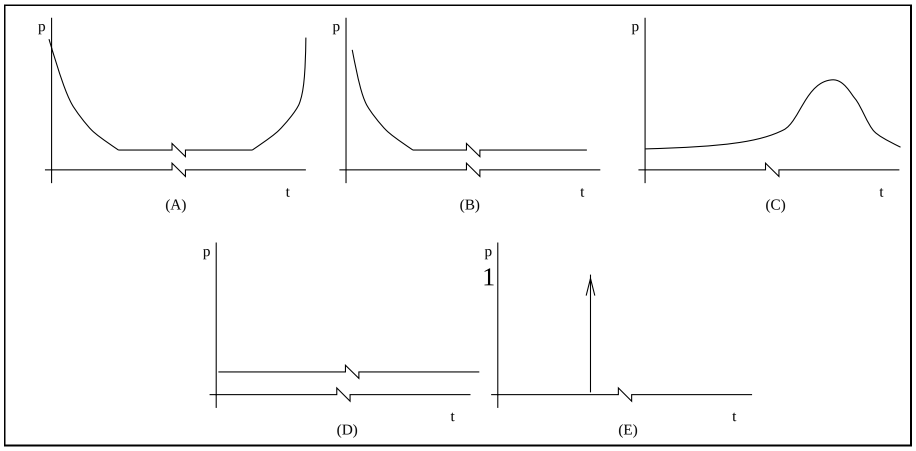

Figure 11: Possible Failure Functions

5.4 ArchSim's Implementation

ArchSim follows the structure of a traditional simulation tool. Each module of the AR model can register future events in a timeline. For example, when a disk is created, the simulation uses the disk failure distribution to compute when the disk will fail; then, it registers the future failure event in the timeline. The simulation engine advances time by calling the module that registered the first event. This module may change the state of the repository and register more events in the timeline. After the module returns, the simulation advances to the next event, in chronological order. The user can choose between two end conditions for the simulation: the simulation can stop after the first document is lost or after all the documents are lost.

ArchSim needs to be very flexible and efficient to meet the challenges of simulating an archival repository. Flexibility is needed to model very different archival conditions and implementations. Speed is needed to cope with many materializations, components, and events. Additionally, each simulation needs to be run many times in order to obtain narrow confidence intervals.

In an AR simulation, many events are inconsequential. For example, suppose that the detection module schedules periodic detection events. If the detection event finds a fault (i.e., there was a failure event before the detection event), then the module starts a component repair; if no failure is detected, then the module does nothing. If a repair module checks a component with mean time to failure ($MTTF$) of 20 years, every 15 days, then, in average, 486 events ($20 * 365/15$) will be fired and the repair module will just return without doing anything; only in the event 487, when a failure of the component has happened, will the repair module perform an action. Given the large number of components and modules that may be part of the model, this large number of inconsequential events represents a significant overhead. To avoid this overhead and to improve efficiency, modules are allowed to register *conditional* events in the timeline. These events will only happen if some other event happened before them. By using conditional events, we can condition the firing of the detection event only if a failure event on an specific component happened before it. A conditional event is not registered directly in the timeline. Instead, it is registered in an index that is part of its triggering event. For example, if event B is conditional to the occurrence of event A, we will put B in the index of A. When a new event A is schedule in the timeline, we look in the index and find that B is conditional to it, so at that point we also schedule B.

To reduce the number of events further, ArchSim also modifies failure distributions. For example, when modeling preventive maintenance, a large number of "replace component" events are generated. We can eliminate all those events, by modifying the failure distribution. Specifically, the original failure distribution is used to generate the time of the next failure of the component. If this time is higher than the prescribed preventive maintenance period, we ignore this time, and we generate a new failure time, gain using the original distribution. We repeat this process until a failure time is lower than the preventive maintenance period. The new distribution then returns the number of iterations minus one, times the PM period, plus the last failure time.

Another challenge for ArchSim is how to deal with a large number of materializations and components. This is done by scheduling only the *next* failure for each component type and associating a trigger with that event. When the failure event happens, the trigger is activated. The trigger computes when the *next* failure of a member of that component type will occur, and adds it to the timeline. Obviously, this approach is only beneficial if we have many components of the same type, which is a reasonable scenario for an AR. In the case when we have N different distributions for N components, this technique does not improve the simulation time, but it does not increase it.

6 Case Study: MIT/Stanford TR Repository

In this section, we use ArchSim to answer some design questions for an hypothetical MIT/Stanford Technical Report Archival Repository. The AR follows loosely the Stanford Archival Vault (SAV) design [9] and implementation [6]. (We actually considered creating such a repository some years ago, when both institutions were participating in the DARPA sponsored CSTR Project.) In this case study, MIT and Stanford preserve their Computer Science Technical reports by replicating the reports at both universities. In this case study we will have to make many assumptions. Our goal here is *not* make any specific predictions, but rather to illustrate the types of evaluations that ArchSim can support, the types of decisions that must be made to model an AR, and the types of comparisons than can be made to support rational decisions among alternatives.

We will assume that the collection has 200,000 documents and that each document is stored in one or more of four available formats. The repository will have two types of components: storage devices (disks) and format interpreters. To ensure preservation, the AR maintains four materializations of each technical report; two materializations at Stanford and two at MIT. Materializations are created by choosing two formats out of the four available formats. Two of the materializations will be in one of the chosen format, while the other two will be in the other. Then, at each site, we place two materializations that are in different formats in two different disks. Figure 12 illustrates the arrangement of the technical reports in this system.

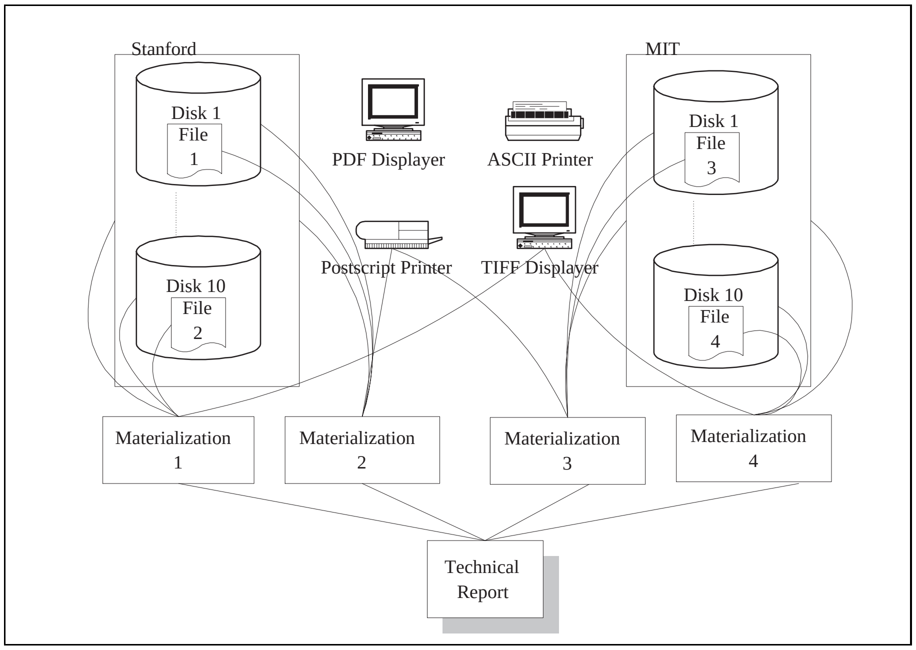

Figure 12: MIT/Stanford CSTR Scenario

Disks, formats, and sites have uniform failure distributions with parameters $1/\phi_{sto}, 1/\phi_{form}, 1/\phi_{site}$. As our base values, we are assuming ϕ_{sto}, ϕ_{form}, and ϕ_{site}, the mean time to failure (MTTF) for disks, formats and sites, to be 3, 20, and 45 years respectively.

These values are our best guess for a typical archival system. We chose 3 years as the disk MTTF, as this is the normal period under which a hard drive is under manufacturer guarantee. We chose 20 years for format MTTF as we theorize that it will take that long for a well-known format to be replaced by a new format and for all displayers and transformers for the old format to be lost. We chose 45 years for the site MTTF as we assign a 50% probability to the event of loosing the Stanford site due to a high intensity earthquake in the San Francisco Bay Area (which is predicted to happen in a 45 year period).

The archival system checks for faults periodically. When a fault is detected in one of the storage devices, the bad device is retired, and a new device is set online. Then, the system regenerates the bad device by making a copy of the lost materializations from the other site. Similarly, when a format becomes obsolete, a new format is selected and a new set of materializations (transformed from a non-obsolete format) is created in the new format. In addition, in case of site failure, the site is recreated from the other site. If all sites, formats and devices that support all the materializations of a technical report are lost, then the technical report is lost and the simulation stops. Disks, formats, and sites are checked and repaired (if needed) every ρ_{sto}, ρ_{form}, and ρ_{site} days. As our base values, we are assuming ρ_{sto}, ρ_{form}, and ρ_{site} to be 60, 60, and 7 days.

Parameter	*Symbol*	*value*
Number of disks	n_{sto}	100 per site
Number of formats	n_{form}	4
Number of documents	num_{doc}	200,000
Mean Time to Disk Failure	ϕ_{sto}	3 years
Mean Time to Format Failure	ϕ_{form}	20 years
Mean Time to Site Failure	ϕ_{site}	45 years
Disk Failure Detection/Repair time	ρ_{sto}	60 days
Format Failure Detection/Repair time	ρ_{form}	60 days
Site Failure Detection/Repair time	ρ_{site}	7 days

Figure 13: Base values

To justify these values, we need to describe what is involved in the detection and repair of component failures. Detecting a failure in a disk involves scanning the whole disk and checking for lost data. When we find lost data, we need to order a new disk and then copy all the data that was in the damaged disk from other sources into the new disk. Assuming that the repair time is 60 days means that we need to dedicate only 3% of the disk bandwidth to scan all materializations in order to detect failures. We did not chose a quicker repair because the scanning overhead would be too high in our opinion. For example, 25% of the disk bandwidth is required to detect failures in 7 days.

For formats, the detection/repair times imply that we are able to realize that a format is obsolete and that we can create a new copy of the document from a non-obsolete format within a 60 day period. In the case of site failures, we are assuming that the detection/repair time for the site is much lower than for formats and disk. The detection itself should be rather fast in this case (the entire site is down), and the 7 days could be the time it takes to find a backup site to take over.

In this case study, we are assuming that failures are total. This is, we cannot partially repair a component and salvage some of the materializations. The failure distribution during access will be assumed to be the same as the failure distribution during archival. The simulation parameters are summarized in Appendix I and the base values for our simulation are in Figure 13.

In our first experiment, we evaluate the effect of the failure MTTF and repair times of storage devices on the system MTTF. In Figure 14, we show the system MTTF for different disk MTTFs, given a detection/repair time of 60 days. To single-out the influence of storage device failures, we are assuming in this experiment that formats and sites never fail. The dotted lines in the figure represent the 99% confidence interval for the simulation, while the solid line is the average of all the simulation runs. As expected, the system MTTF increases when the disk MTTF increases (when the disk failure rate decreases). The exponential shape of the curve is the result of the constant repair time. As we keep increasing the disk MTTF, it is much more improbable that another device will also fail before 60 days have passed. This graph allows us to select a good storage device for a target system MTTF. If the library targets a 10-year MTTF, then a disk with a failure MTTF of 3 years will suffice. However, if the library requires a MTTF of 100 years, then we will need disks

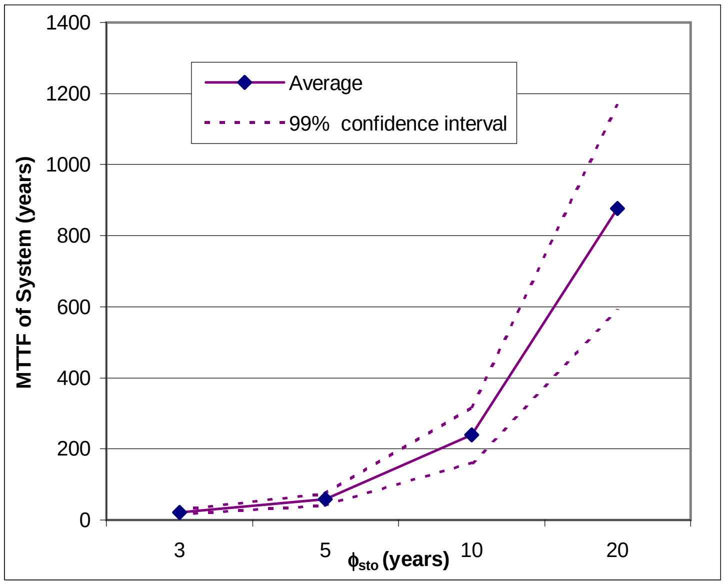

Figure 14: MTTF vs. ϕ_{sto} with ρ_{sto} of 60 days

with a MTTF of about 6 years. Most manufacturers guarantee their disks for three years and very few guarantee them beyond five years. Therefore, a 6-year disk MTTF requirement will be hard (or very expensive) to meet. Nevertheless, we can still achieve our target system MTTF by changing other parameters in our system, as we will see in the next experiment.

We now evaluate the sensitivity of the system MTTF to the repair time (ρ_{sto}). In Figure 15, we show the MTTF of the AR for different disk MTTF (ϕ_{sto}) values, and for different expected detection/repair times. (Note that the graph has a logarithmic scale. Confidence intervals are not shown to avoid clutter.) First, let us concentrate in the curve for a ϕ_{sto} of 3 years. As expected, the MTTF of the system decreases when the ρ_{sto} of the storage devices increases. However, the shape of the curve is more interesting. At low detection and repair times, there is a high positive impact on the MTTF of the system. However, as we increase the repair times, the system MTTF drops sharply. Interestingly, for a repair times greater than 120 days, about 1/9 of the MTTF of the storage device, the effect on the system MTTF of the detection/repair module is small. Thus, the detection and repair times must be much lower than the storage device failure MTTF to have a significant effect on the MTTF of the Archival System.

What is the optimal solution for a given MTTF with respect to disk MTTF repair times? The answer depends on the cost assigned to those two factors. By looking to all the curves of Figure 15, we can observe that we can achieve similar MTTF by using better media or by reducing the detection/repair times. For instance, a system that uses a storage device with ϕ_{sto} of 5 years (i.e., a low quality storage device) and has a detection/repair time of 30 days, is as good as a system that uses a high quality storage device with ϕ_{sto} of 20 years, but is only checked and repaired every 360 days. The decision of which alternative to choose will depend on the cost of the storage device

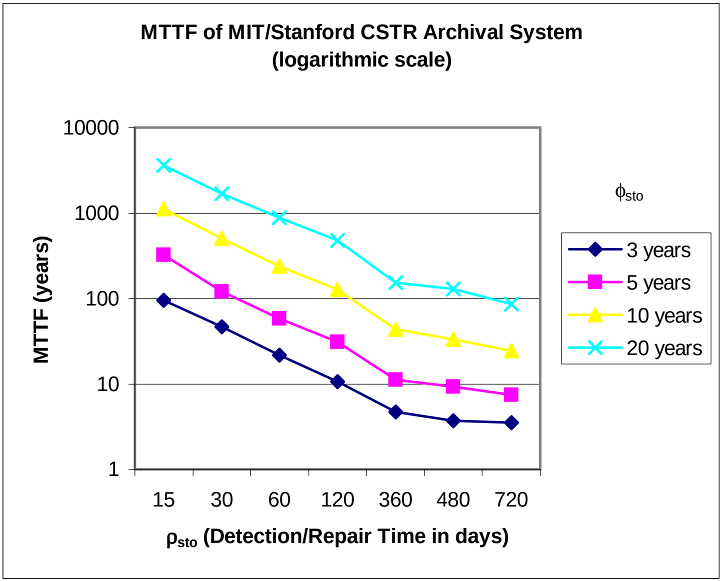

Figure 15: System MTTF (logarithmic scale)

versus the cost of more frequent detections.

We now expand our experiments by allowing formats to fail. In Figure 16 we show system MTTF as a function of ϕ_{sto} when formats can fail. We fix ρ_{sto} at 60 days, and now formats can fail with ϕ_{form} = 20 years. To avoid introducing additional factors in the analysis, we assume site failures still cannot happen. Format failures are detected and repaired with ρ_{form} = 60 days. For comparison purposes, we have included in Figure 16, the results presented in Figure 14. The important conclusion that we can derive from the figure is that in an archival repository we cannot focus on single component types. It is surprising that even though the format MTTF is much larger than the disk's MTTF, the failure of formats still has a significant impact. This is because a document is lost if there is a disk failure *or* a format failure. The result is that we are taking the "worst" of those two failures, resulting in a system with a low MTTF. The result of this experiment shows that we need a comprehensive model, like the one proposed in this paper, to realize the interactions between the components and their effects on system MTTF.

Our model can be used to explore other possibilities that may improve reliability. For example, we now consider what happens if we are able to increase the number of copies maintained in the sites from two to three. In Figure 17, we maintain the same parameters at the same base values, but now each site has three copies in three different formats. In the figure, we can see that by increasing the number of copies to three, the MTTF increases from 34 years to 2101 when ϕ_{sto} is equal to 3 years. Although increasing the number of copies to three will undoubtedly increase the cost (as we need an additional 33% disk space and the need to handle an extra format), we have achieved an important improvement in the MTTF of the system.

As we stated earlier, a comprehensive model is important to get an accurate picture of the

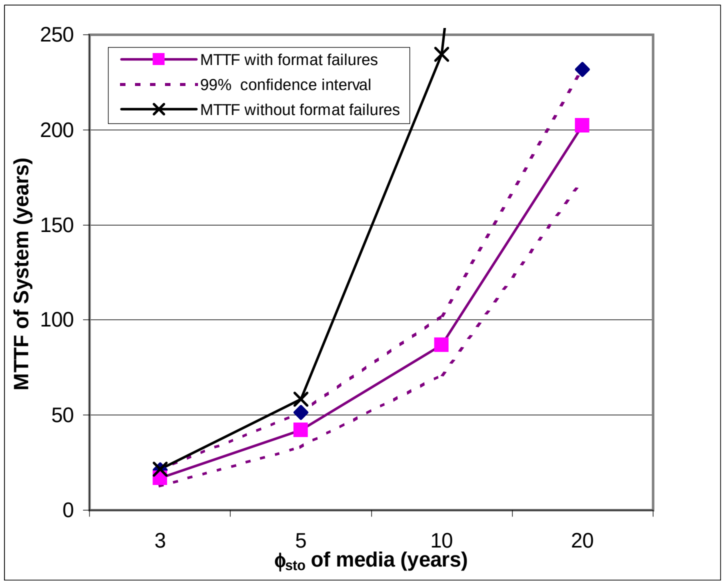

Figure 16: MTTF ϕ_{form}=20 years, ρ_{sto}=ρ_{form}=60 days

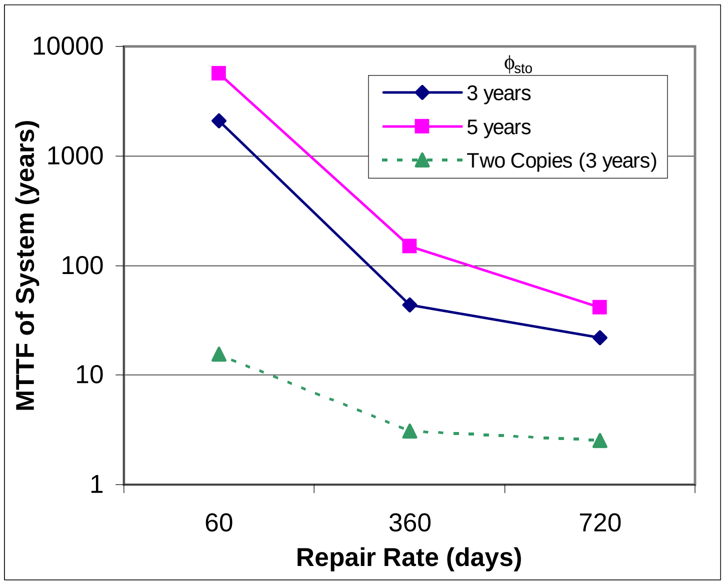

Figure 17: MTTF with 3 Copies (logarithmic scale)

reliability of the system. In the next experiment, we use our system to explore a different, more complex failure distribution for storage devices. In this new distribution, we want to include the issue of "infant mortality."

When the failure distribution includes infant mortality, storage devices have a higher failure rate in the beginning of their life (in our case, 30 days) than in the rest of their lives. This can be expressed as a distribution that has a low MTTF within the first 30 days and a higher MTTF after that. The MTTF after the 30-day period will be 5 years. We will vary the early MTTF between 44 and 285 days. In this experiments we will assume that formats cannot fail. All other assumptions and repair procedures of the previous experiments are maintained (see Figure 13). In Figure 19 we show the system MTTF at given percentages of storage devices that fail in the first 30 days. For example, with an early MTTF of 285 days, 10% of the devices will fail within 30 days. With a MTTF of 135 days, 20% will fail. We can convert from one measure to the other by using the formula (conversions shown in Figure 18): $\frac{1}{MTTF} = 1 - \sqrt[30]{1 - pct}$

Percentage failed devices	MTTF first 30 days (days)
0%	N/A
10%	285
20%	135
30%	85
40%	59
50%	44

Figure 18: Conversion between %failed devices and MTTF

As expected, the higher the infant mortality, the lower the MTTF of the system. At a 0% infant mortality, the system MTTF was 65 years, dropping to 48 years when the infant mortality is 10% and dropping to only 11 years at 50% infant mortality level. Given this, when using components that suffer from infant mortality, a way to increase the MTTF of the system is for the failure detection module to check new components much more often than older components.

We now explore the issue of aging of storage devices. With aging, a storage devices will have a lower MTTF at the end of its life. For example, a disk may have a 5 year MTTF during its initial life and a MTTF of 2 years when it reaches its "aging" phase. In this experiment we will assume that formats cannot fail. All other assumptions and repair procedures of the previous experiments are maintained (see Figure 13). We evaluated the MTTF of the system for different points when aging starts (see Figure 20. As expected, the sooner aging starts, the lower the MTTF of the system. If aging never occurs, the MTTF of the system is 65 years. If aging starts after 1 year, the system MTTF is 17 years, increasing to 42 years when aging starts after 5 years. Given this, when using components that suffer from aging, a way to increase the MTTF of the system is for the failure detection module to check old components much more often than newer components. Moreover, we should consider replacing old components with newer ones before the old components fail.

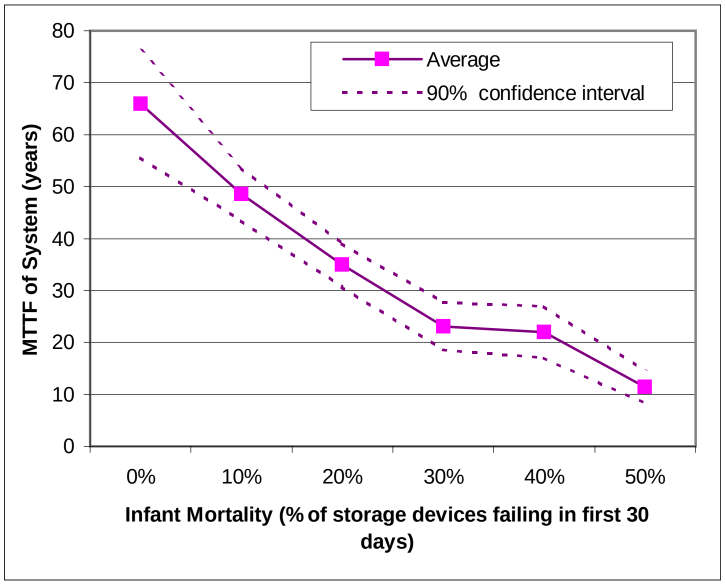

Figure 19: MTTF with Infant Mortality

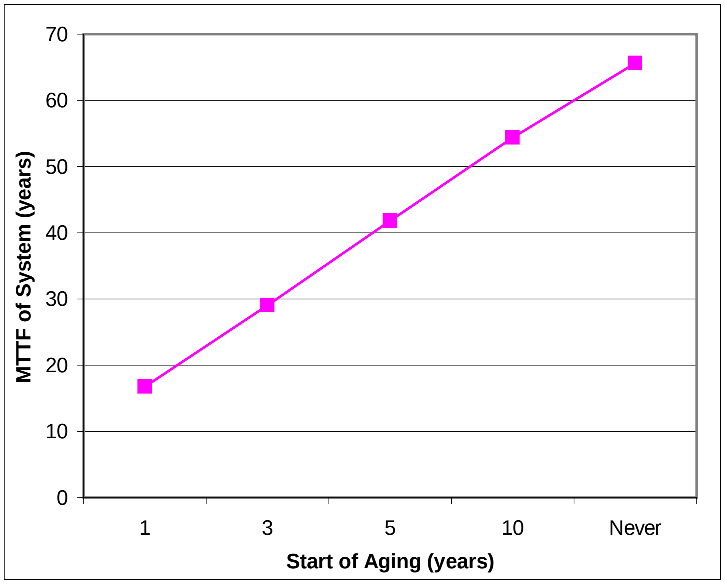

Figure 20: System MTTF with Aging and ρ_{sto} of 60 days

As a final experiment, we will evaluate the impact of Preventive Maintenance (PM) on a system with aging disks. Specifically, we will replace old disks with new ones before the old disks are expected to fail. This is done by copying (instantaneously) all documents from the old disk into a new disk, and then removing the old disk. In this experiment, we are assuming that disks do not have infant mortality and that disks have a 5 year MTTF during their initial life and a MTTF of 2 years when they reach their "aging" phase. Figure 21 shows five PM schedules for disks that age at different points. From the figure we can see that the most efficient PM schedule is one that matches the start of the aging period of the disk. For example, when we use a 10-year PM plan a system with disks that age after 5 years, will have a MTTF of 42 years. When we never perform PM, the system MTTF does not increase significantly. However, when we use a 5-year PM plan, the MTTF of the system increases to 63 years. If we keep increasing the frequency of the PM plan, the MTTF does not improve much more.

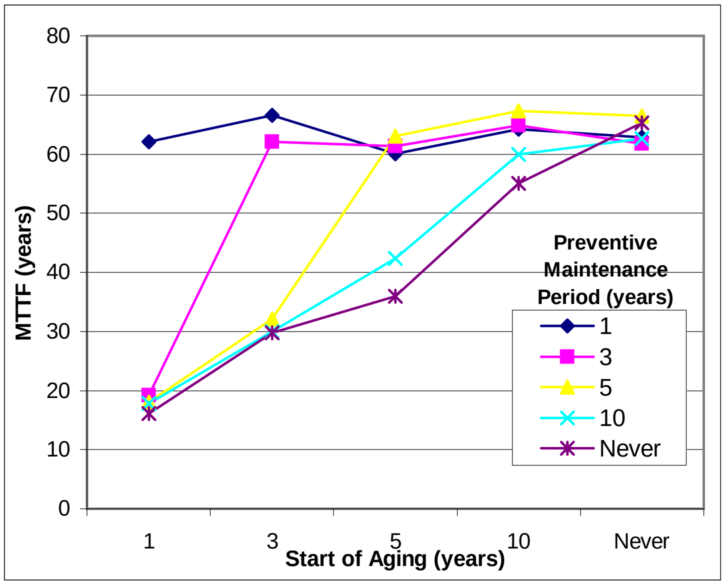

Figure 21: System MTTF with PM and Aging

7 Conclusions

In this paper, we have studied the *archival problem*. We studied the different options for recovery and preventive maintenance, developing a comprehensive model for an AR. We described a powerful simulation tool, ArchSim, for evaluating ARs and the available archival strategies. We described how ArchSim can efficiently perform large simulations many components and very long durations. We demonstrated the use of ArchSim with a case study for a hypothetical Technical Report repository operated between Stanford and MIT. We considered options such as disks with different reliability, number of copies, format failure handling, and preventive maintenance. We believe ArchSim can help librarians and computer scientists make rational decisions about preservation, and help achieve better archival repositories.

References

[1] Thomas Antognini and Walter Antognini. A flexibly configurable 2d bar code. In *Information Based Indicia Program Technology Symposium*, 1996.

[2] W. David Kelton Averill M. Law. *Simulation Modeling & Analysis*. McGraw-Hill, 1991.

[3] John W.C. Van Bogart. *NML Storage Technology Assessment: Final Report*. National Media Lab, 1994.

[4] John W.C. Van Bogart. *Magnetic Tape Storage and Handling*. National Media Lab, 1995.

[5] C. Borgman, S. Chen, H. Garcia-Monlina, K. Thibodeau, , and G. Wiederhold. *NSF Workshop on Data Archival and Information Preservation*. National Science Foundation, March 1999. At http://cecssrv1.cecs.missouri.edu/NSFWorkshop/.

[6] Arturo Crespo Brian Cooper and Hector Garcia-Molina. Implementing a reliable digital object archive, 1999. Submitted for publication to ACM DL 2000.

[7] E. Çinlar. *Introduction to Stochastic Processes*. Prentice-Hall, 1975.

[8] Compaq. Intellisafe. Technical Report SSF-8035, Smal Form Committee, January 1995.

[9] Arturo Crespo and Hector Garcia-Molina. Archival storage for digital libraries. In *Proceedings of the Third ACM International Conference on Digital Libraries*, 1998. Accessible at http://-www-diglib.stanford.edu/cgi-bin/WP/get/SIDL-WP-1998-0082.

[10] John Garrett and Donald Waters. Preserving digital information: Report of the Task Force on Archiving of Digital Information, May 1996. Accessible at http://www.rlg.org/ArchTF/.

[11] G. Gordon. *The Application of GPSS V to discrete System Simulation*. Prentice-Hall, 1975.

[12] Andreas Reuter Jim Gray. *Transaction Processing: Concepts and Techniques*. Morgan Kaufmann Publisher, 1993.

[13] Kranch. Preserving electronic documents. In *ACM Digital Library Conference*, 1998.

[14] T.I. Oren. Application of system theoretic concepts to the simulation of large scale adaptive systems. In *Proceedings of the 6th Hawaii International Conference on Systems Sciences*, pages 435–437, 1973.

[15] Erik Ottem and Judy Plummer. Playing it S.M.A.R.T.: The emergence of reliability prediction technology. Technical report, Seagate Technology Paper, June 1995.

[16] Dhiraj K. Pradhan. *Fault-Tolerant Computer System Design*. Prentice Hall PTR, 1995.

Appendix I: MIT/Stanford CSTR Scenario Parameters

- **AR Description**
  - Initial collection: num_{doc} documents. Each document, d, will have the following four materializations:
    - $\langle d, MIT, disk_i, form_j \rangle$,
    - $\langle d, MIT, disk_k, form_l \rangle$,
    - $\langle d, Stanford, disk_x, form_j \rangle$,
    - $\langle d, Stanford, disk_y, form_l \rangle$.

    Where MIT and $Stanford$ are the two sites; $disk_i$, $disk_k$, $disk_x$, and $disk_y$ are different storage devices; and, $form_j$ and $form_l$ are two different formats.
  - Number of components and types: n_{sto} storage devices, n_{form} formats, 2 sites.
  - Failure dependency graph: $site \rightarrow disk$, when the disk is in the given site.
- **Distributions**
  - Disk Failure distribution during access: $U(1/\phi_{sto})$
  - Format Failure distribution during access: $U(1/\phi_{form})$
  - Site Failure distribution during access: $U(1/\phi_{site})$
  - Disk Failure distribution during archival: $U(1/\phi_{sto})$
  - Format Failure distribution during archival: $U(1/\phi_{form})$
  - Site Failure distribution during archival: $U(1/\phi_{site})$
  - Disk Failure Detection distribution: instantaneous
  - Format Failure Detection distribution: instantaneous
  - Site Failure Detection distribution: instantaneous
  - Disk Repair distribution: instantaneous
  - Format Repair distribution: instantaneous
  - Site Repair distribution: instantaneous
  - Document creation: num_{doc} documents at startup, then no documents are created.
  - Document access rate: irrelevant as failure distribution during access is the same as during archival.
  - Access duration rate: irrelevant as failure distribution during access is the same as during archival.
  - Document selection: uniform over the num_{doc} documents.
- **Policies**
  - Document Creation policy: for each document, four materializations are created, 2 in each site. In each site, each materialization is created in a different disk and in a different format.
  - Document to Materialization: read from any materialization.
  - Failure detection algorithm: complete scan of all disk, formats, and sites every τ_{sto}, τ_{form}, and τ_{site} days, respectively.
  - Damage Repair algorithm: discard bad component and replace with new component taking δ_{sto}, δ_{form}, and δ_{site} days, for disks, formats, and sites respectively.
  - Failure prevention algorithm: none
- **ArchSim Parameters**
  - Stop Condition: when losing the first document.
  - Simulation time unit: days.