Standards in the Data Storage Industry: Emergence, Sustainability, and the Battle for Platform Leadership

by

Jean-Claude Jacques Saghbini

M.S., Mechanical Engineering
Massachusetts Institute of Technology, 1996

B.S., Mechanical Engineering
University of Massachusetts Dartmouth, 1994

Submitted to the System Design and Management program
in partial fulfillment of the requirements for the degree of

Master of Science in Engineering and Management

at the

Massachusetts Institute of Technology

June 2005

Signature of Author: ________________________
Jean-Claude Jacques Saghbini
System Design and Management
February 2005

Certified by: ________________________
Michael Cusumano
Sloan Management Review Distinguished Professor of Management

To my parents Jacques and Leila
and my brother Christian

This page is intentionally left blank

Standards in the Data Storage Industry: Emergence, Sustainability, and the Battle for Platform Leadership

By

Jean-Claude J. Saghbini

Submitted to the System Design and Management Program
in partial fulfillment of the requirements for the degree of
Master of Science in Engineering and Management

Abstract

In order to cope with the continuous increase of magnetically stored data and of mission critical dependence on this data, storage systems must similarly increase in their functionality offerings, and with that, their complexity. The efficient management of the heterogeneous and complex aggregation of these systems is becoming one of the major challenges to IT customers. At the same time, hardware is becoming commoditized, and the industry is looking towards software for additional revenue generation. This document examines proprietary as well as open-standards attempts at solving the interoperability problem. The first attempt was made by EMC when it developed WideSky, a middleware software layer that would be able to manage third party hardware. It is shown that the aim was to eventually transform this middleware into a de facto standard and with that establish platform leadership in the industry. The WideSky effort failed, and the analysis of this failure attributes it to a lack of industry support and inability at establishing a sustainable value chain. Meanwhile, the industry players rallied around the SNIA body and adopted the SMI specification (SMI-S) as a standard. SMI-S adoption is on the rise, but although it has the formal backing of most of the storage industry firms, it has not yet fulfilled its promise of enabling centralized management of heterogeneous systems. This is partially because of the fact that the functionality that it provides is still lagging behind the functionality that native APIs provide. Moreover, client adoption and the availability of client products that can be directly used by IT customers are still very limited. However, an examination of the dynamics surrounding this standard show how SMI-S will benefit greatly from learning effects and network externalities as it continues to grow, and although lagging in traditional functionality, it offers an ancillary functionality of interoperability that is missing from current non-standardized software interfaces. The adoption tipping point is highly dependant on whether or not the value chain can be established before vendors start dropping support for the specification. It is proposed that a positive tipping of the market will make SMI-S a disruptive technology that has the potential of becoming the dominant design for storage management interfaces.

Thesis Supervisor: Michael Cusumano
Title: Sloan Management Review Distinguished Professor of Management

This page is intentionally left blank

ACKNOWLEDGEMENT

I would like to present my first "thank you" to my advisor Prof. Michael Cusumano for his guidance and support during this research effort. In addition, I am greatly indebted to my managers Rony Lipetz and Gil Shapira for providing me with the opportunity and support to pursue my SDM studies. My research has also benefited from very insightful discussions with Joseph Saleh and Richard Roscoe, and from joint projects with Badari Kommandur, Prithvi Bannerji, Sashi Somovarapu, and Rafael Holtzhacker.

On a personal note, I would like to thank Rania, for her support and love and for putting up with all the stress of work, school, and a long distance relationship.

Above all, I would like to thank my parents Jacques and Leila, and my brother Christian for everything they have done to allow me to be at this point in my life.

This page is intentionally left blank

Table of Contents

Table of Figures

Table of Tables

1 Introduction

1.1 Background

With the advent of the internet revolution, it is no surprise that the volume of stored digital data has been experiencing a very rapid increase. A study conducted in 2003 has estimated that between 1999 and 2002, the amount of magnetically recorded data has risen by 73% from 2,428 Petabytes/year to 4,207 Petabytes/year[1]. In parallel, over the last 10 years, the data storage industry has also seen a fast paced increase in its diversity of product offerings: storage arrays, servers, tape libraries, switches, host bus adapters, and directors. The products themselves are also increasing in their sophistication and functional capabilities and with that their complexity. Moreover, the popularity of combining these various products into networked solutions such as Storage Area Networks (SAN) and Network Attached Storage (NAS) is also growing. Estimates are that most enterprises have either adopted SAN in their IT infrastructure or are going to do so in the near future[2]. This is largely due to the fact that the Direct Attached Storage (DAS) approach that was so popular in the nineties can no longer cope with today's enterprise level IT requirements. Amidst this proliferation of data and of systems to handle this data, comes also the need for storage management software to manage the data and the various storage systems. The problem is that the management software for the storage systems is not standardized. When a full storage solution, consisting of many heterogeneous storage systems, is assembled; its centralized management becomes quasi-impossible because of the lack of compatibility between the various software interfaces.

IT customers are forced to rely on individual component-level management instead of a solution-level management; and this comes at a high cost, both in financial terms as well as in complexity.

Currently, there is a movement towards standardization throughout the entire data storage industry. It is being done under the umbrella of the Storage Networking Industry Association (SNIA), and has the backing and contribution of the majority of the vendors. The output of this standardization effort is the Storage Management Initiative Specification, or SMI-S. This specification aims at standardizing the interfaces to the various components of a storage solution, and has been touted as the answer to the problem of heterogeneous storage system management.

1.2 Research Objectives

There are various factors influencing the evolution of management software standardization in the data storage industry; and although embarked on a growth path, the future of the SMI specification is still unclear. The adoption rate could continue to climb, in which case the specification has the potential of becoming the standard of the industry and the dominant design for storage management interfaces. On the other hand, the adoption rate could tip the other way, resulting in a reversion to a pre-standardization era of native API-based management and struggles over proprietary standards.

This research first establishes a foundation for the study by presenting an overview of the data storage industry from a technology as well as a business point of view. This knowledge is instrumental at providing a context for the rest of the analysis. The core of the thesis can then be divided into two main objectives. The first objective is to understand the dynamics of the struggle between open standards and proprietary standards efforts; and analyze the external factors and firms' strategic decisions that led to the demise of the latter. The second objective is to understand the factors influencing the growth and sustainability of SMI-S; and to offer insights as to the future of the specification and its potential at becoming a disruptive technology.

1.3 Structure of Thesis

Chapter 1 gives the background to the research and presents a general overview of the current state of the data storage industry. It also states the objectives that the research aims at accomplishing.

Chapter 2 is the first of two background chapters. It provides a technology overview of the data storage industry, both from hardware as well as software point of view. It also provides an overview of the construct and characteristics of the most widely deployed storage solutions such as DAS, SAN, and NAS; along with the challenges of managing each one of them.

Chapter 3 is the second background chapter. It addresses the business aspect of the industry. It lists the major players and highlights the competitive and complementary dynamics between them. Finally, it discusses the trends of shift from hardware to software, explaining their causes and their implications.

Chapter 4 discusses the rise and fall of EMC's WideSky, and the struggle between proprietary standards and open standards at establishing design dominance. In addition, the chapter traces the history of the SNIA body and of the SMI specification (SMI-S), and of the factors that led to their industry-wide support and adoption.

Chapter 5 provides an overview of the SMI-S technology. It starts with the hierarchical structure of the specification and it dependence on other standards. Next, an overview of the SMI-S architecture is presented; followed by the future plans for SNIA and SMI-S. The knowledge that this chapter provides in understanding the capabilities and limitations of the specification is instrumental to the analysis to follow, regarding sustainability dynamics.

Chapter 6 lists and elaborates on the major factors that are fueling or dampening the adoption of SMI-S. This is followed by an analysis that provides insight regarding the 'tipping point' of the SMI-S technology, in order to help answering the question: when and how will SMI-S adoption tip? Finally, SMI-S is presented as a disruptive technology, in an aim at providing yet more insight as to its possible course of evolution and adoption.

Chapter 7 is the conclusion chapter; it offers a summation of the entire research effort in an executive summary format. In addition, it lays out the foundation of future research through a series of questions related to the effects of standardization in general and SMIS in particular on the construct of the industry, on the value chain, on architecture, and on innovation.

Appendix 1 lists and defines all the acronyms and storage capacity units used in this thesis.

Appendix 2 lists the literature references of the research.

2 Data Storage Technologies

This chapter provides an overview of the various hardware and software storage technologies, and of the commonly used enterprise storage architectures.

2.1 Hardware

2.1.1 Storage Arrays

Storage arrays are the backbone of any storage solution. A storage array is a hardware appliance that is populated with disk drives. In its most basic usage, it is connected to a server and exposes the combined storage capacity of its disk drives as available storage for the server. Storage arrays are normally classified under three main product categories, but since there are no strict guidelines to the classifications, the following is just meant to give the reader a feel for the storage capacities of each category.

- The low end or entry-level systems can house 10 to 20 disk drives, and have a total capacity of less than 10 Terabytes (TB).
- The medium-range systems have a maximum capacity that can reach about 50 Terabytes.
- The high end or enterprise-level storage systems have a maximum capacity of 200 to 300 Terabytes.

In a storage array, the devices (also called, drives, disks, or volumes) are mapped internally to the ports of the array and can be accessed by external systems that are connected to the ports. In addition to storage capability, arrays provide the end user with various functionalities that can be performed on the stored data, such as mirroring and data replication for example.

2.1.2 Servers

The server is the actual *computer* that processes the information. It is also sometimes referred to as the host. In a storage solution, the server is connected to the storage array via Fiber Channel (FC), SCSI, or any other type of connectivity; the most important of which will be explored in a later section. In more complex networked solutions, servers are indirectly connected to the storage through an intermediate appliance called a switch. The component that handles the physical connection inside the server is the Host Bus Adapter, most commonly referred to as HBA. A cable is usually run between the port of the HBA and a port on the storage array. The server then becomes capable of accessing the devices of the array that are mapped to that port.

The major competitors in the servers market are firms such as IBM, Sun Microsystems, HP, and Dell. Servers are responsible for two main functionalities. The first one is the handling of the data I/O (Input/Output) to and from the storage array. The second and more important functionality – for the purpose of this research – is the running of the software. For example, servers run filesystems, database software, volume managers, and storage array management. The management part can be done either in-band or out-of-band. In-band management means that the server that is running the management software (for an array or switch) is also handling I/O. The communication path through which the commands are sent to the array is the same as the one I/O is being sent on.

Commands are therefore interleaved with *writes* and *reads*, which has the potential disadvantage of performance degradation of the I/O or delays in sending and receiving the commands. With out-of-band management on the other hand, the server that is running the management software is a different server than the one doing the I/O (see Figure 1).

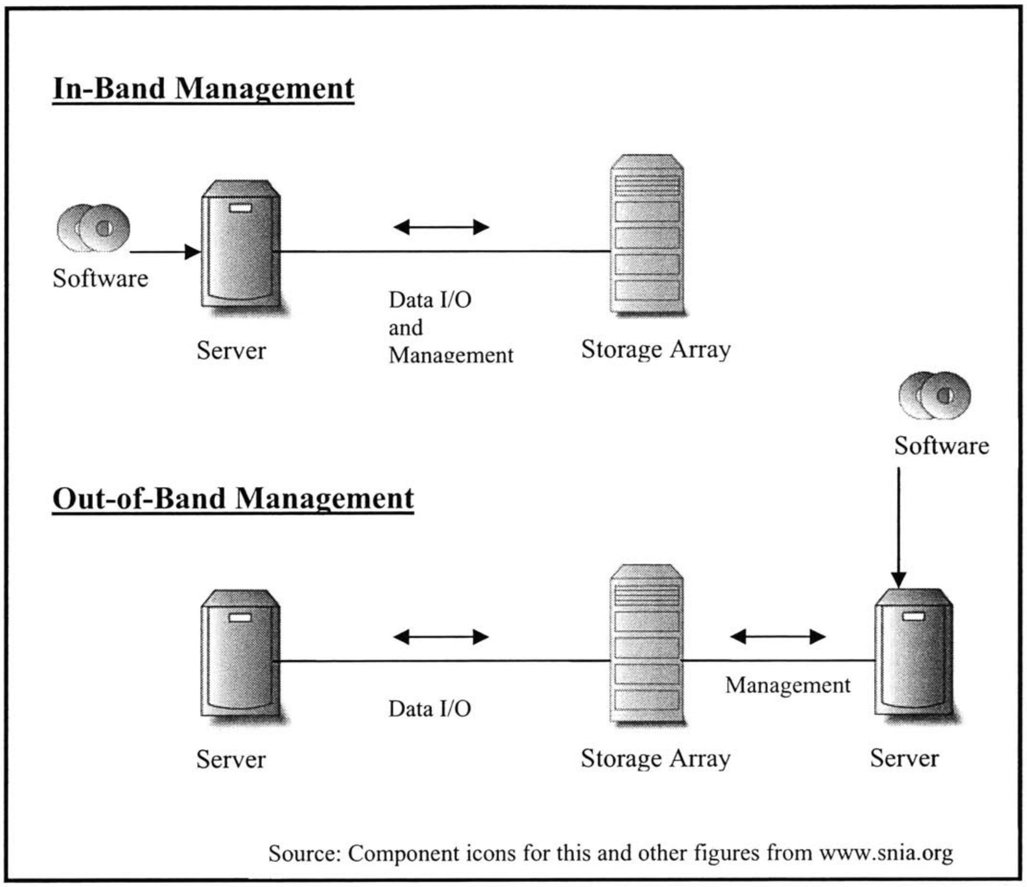


Figure 1: In-band versus out-of-band management.

2.1.3 Host Bus Adapters

Host Bus Adapters (HBAs) are circuit boards that plug into the backend of a server. They handle the interface between the CPU of the server and external peripherals such as storage arrays for example. The technology embedded in an HBA uses a standard communication protocol to transform the electrical or optical signals coming from the connection cable into digital data. Small software modules called *drivers* are installed on the servers and handle the interface with the HBA. Players in this market include companies such as Qlogic, Emulex, and AMCC.

2.1.4 Switches

Switches are appliances that reside between the servers and the storage arrays. They allow multiplicity of connections between the systems on each end (see Figure 2 and Figure 3). For example, without a switch, one server HBA port is limited to seeing one storage array port, and through that, the internal data devices mapped to that port. When a switch is present, the server connects to a port of that switch, and internally, the switch can then fan that connection out to many ports, and out of each port a connection is made to a port of the array. This results in the server being able to access more than one port through a single HBA, and through that all the devices that are mapped to all these ports, as shown in Figure 2. Vice versa, multiple servers can connect to multiple ports of a switch which then fans in all these ports to a single connection to an array port (Figure 3). Through this scheme, multiple servers can access the same array port and by doing that, share the devices that are mapped to that port, and in turn, the data that is present on these devices. Such a configuration is commonly implemented in clustered environments.

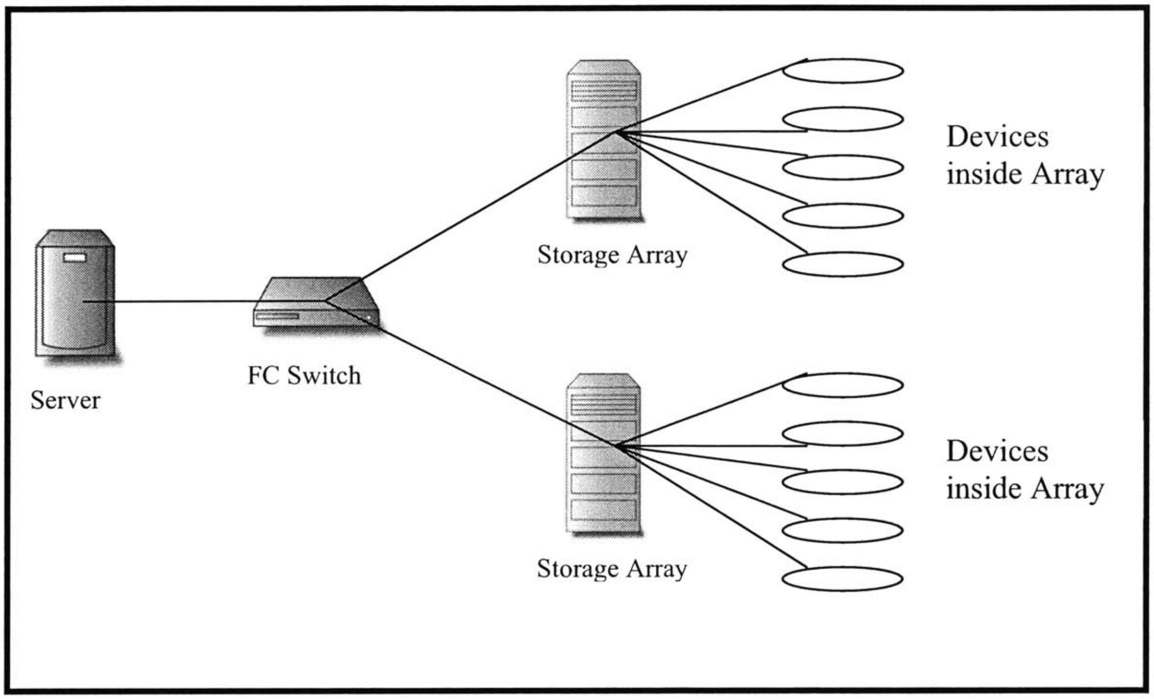


Figure 2: Example of a fanned out server-switch-array connection.

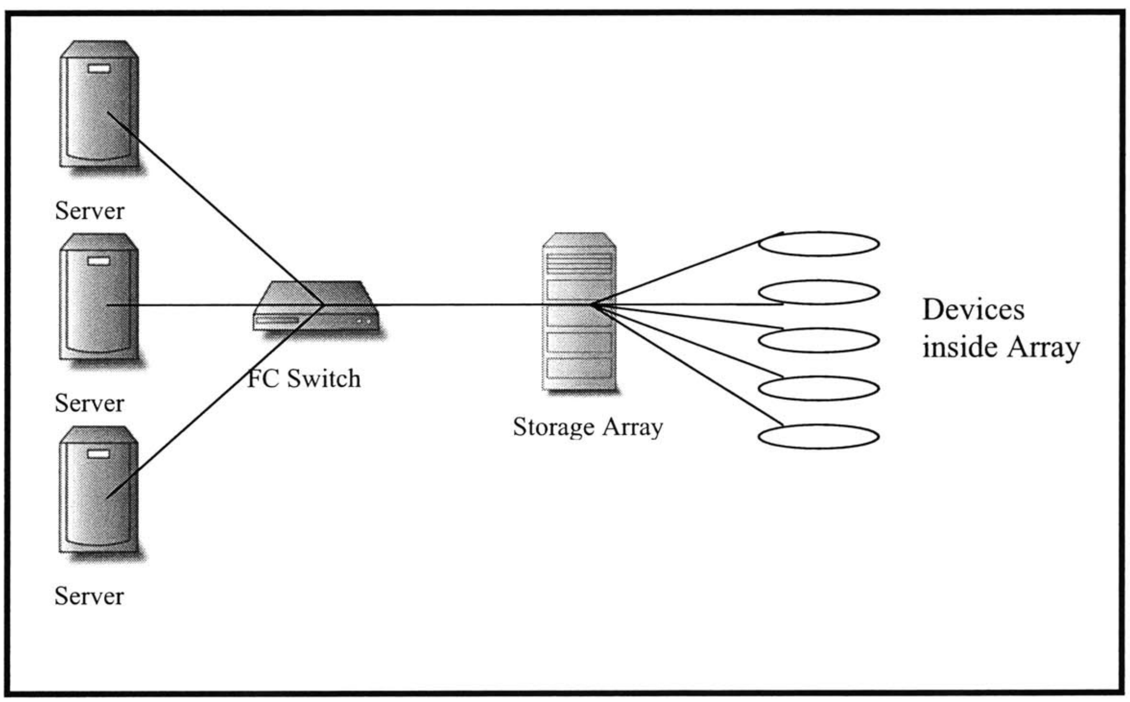


Figure 3: Example of a fanned in server-switch-array connection.

2.1.5 Tape Libraries

In most deployed IT solution, data backup and archiving is still mainly done on magnetic tapes. Tapes libraries are therefore the component of a storage solution that handles this functionality. Tapes libraries vary in size, capacity, and performance; and just like all the other products of the industry, they can be classified in three categories, high-end, medium range, and low-end. In essence, they are appliances that are an aggregation of multiple tape drives along with hardware and software to manage them. In a SAN, they are attached to the fiber fabric, and handle two major operations:

- Backup of data from array disk storage to tape
- Recovery of data from tape to array disk storage

2.2 Connectivity

There are various types of connections that can be made between servers, switches, and storage arrays. Each type is characterized by the physical medium through which the connection is made, as well as a driver that implement the protocol for the communication. This section will briefly discuss the most prevalent types of connections, namely SCSI, FC, ESCON, FICON, and iSCSI.

2.2.1 SCSI to Fiber Channel

Throughout the nineties, SCSI (Small Computer System Interfaces) was the most popular type of connectivity between storage arrays and servers. Currently, Fiber Channel (FC) connectivity dominates the market. Most storage array vendors do not ship their products with SCSI support anymore, and there are hardly any HBA manufacturers that still produce SCSI HBAs. The uptake of FC was greatly helped by the fact that although the connection requires a different medium (fiber optic cable instead of SCSI wire cable), the FC protocol is the same as the SCSI protocol. This allowed software written to operate with the SCSI protocol and SCSI connections to seamlessly operate with FC connections. Fiber Channel's advantages over SCSI are the following[3]:

- Speed: 320 MB/sec for SCSI versus 400 MB/sec for FC
- Distance: Perhaps this is one of the biggest advantages of FC. SCSI cable length is limited to 12 meters, while FC can extend to 10,000 meters.
- Number of connections: SCSI can support 15 devices per channel, while FC supports 127 devices per arbitrated loop

2.2.2 ESCON to FICON

ESCON is a connection mechanism between storage arrays and IBM mainframe computers. It's use is currently on the decline and being replaced by FICON (Fiber Connectivity). Although some manufacturers still support ESCON for backward compatibility purposes, this support is slowly diminishing, and all new products that are expected to be compatible with the IBM mainframe are now supporting FICON.

The advantages of FICON over ESCON are highlighted by a Brocade white paper[4], and fall in the following areas:

- Speed: ESCON allows half-duplex 17MB/sec compared with FICON's full-duplex of up to 400MB/sec
- Distance: 43 kilometers for ESCON versus 100 kilometers for FICON
- Number of connections: ESCON allows 16 channels on a physical connection. FICON allows 16,000 addresses per channel and 4,000 channels per logical unit. This enables the creation of SANs with fewer physical connections.

2.2.3 iSCSI

iSCSI stands for *internet* SCSI. It is a recent technology that is based on the premise of connecting storage arrays to servers via Ethernet fabric. Instead of the point to point direct connection of SCSI and FC, in an iSCSI connection, both sides connect to an IP network (see Figure 4). Similar to FC, iSCSI implements the same protocol as SCSI, which also facilitated its uptake because of compatibility with existing software.

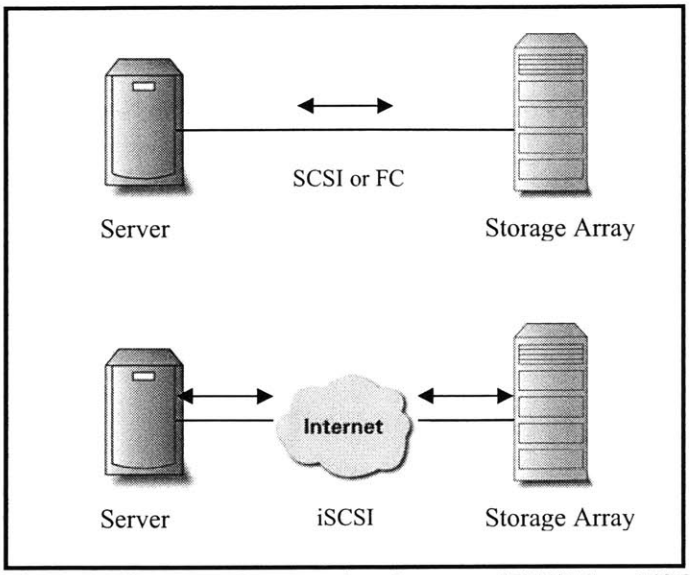


Figure 4: Conceptual comparison between SCSI/FC and iSCSI connections.

2.3 Software

Software products in the data storage industry can be subdivided into two types, storage management software, and content management software. These two types have been slowly converging towards the self-actualization of the Information Lifecycle Management (ILM) vision. However, since the focus of standardization in the industry has been mainly around storage management, it is this type of software that will be the focus of the following analysis.

This section will provide a technical overview of storage management software. Paquet (2005) presents a seven-level hierarchy of software categories[5], some of which –the major ones- will be elaborated upon in this analysis.

2.3.1 Backup, Recovery, and Replication

This constitutes the largest sector of storage management software. Backup and recovery applications control the backup of the data from storage arrays to tapes (or other disk based technologies), and the recovery of the data back to the storage array from the tapes. Given that in most enterprise firms, data loss and data unavailability are unacceptable and prohibitively costly; this makes backup and recovery the most critical aspect of IT storage management, because it revolves around the protection of data. Replication is also done with the intent of backup and protection. There are various types of replication, such as SAN replication, point-in-time replication, and remote replication.

2.3.2 SAN Management

SAN management is the next largest category. It includes all aspects of managing a SAN networked solution consisting of servers, storage arrays, switches, and directors. SAN management functionality includes the following: path configuration, zoning, monitoring, and path failover recovery.

2.3.3 Storage Resource Management

Storage resource management (SRM) software is mainly used for the monitoring of resource usage in a storage network. This information is used to optimize the overall structure and data distribution over the entire solution

2.3.4 Higher Order Software

By higher order software, I refer to categories such as Hierarchical Storage Management (HSM) and policy based management. These markets are still the smallest, but they bring the industry closer to the sought-after goal of automated abstracted management.

2.4 Integrated Storage Solutions

A storage solution refers to the architecting and combining of several hardware storage components into an IT deployable system. As it was previously mentioned, there are three major types of storage solutions deployed in IT shops: the Direct Attached Storage (DAS), the Storage Area Network (SAN), and the Network Attached Storage (NAS). The three solutions vary in complexity and capabilities, and although the DAS form of storage was prevalent in the nineties, it has seen a considerable decline of popularity in enterprise settings in recent years, yielding the way to the other two more complex solutions. Figure 5 shows these trends starting from 1999 with a projection to 2007. Both NAS and SAN are seen as gaining market share from DAS implementations, and SAN is expected to be more popular that NAS. This latter trend could also be confirmed anecdotally since discussions in literature and market research tend to mention and focus more on SAN than NAS.

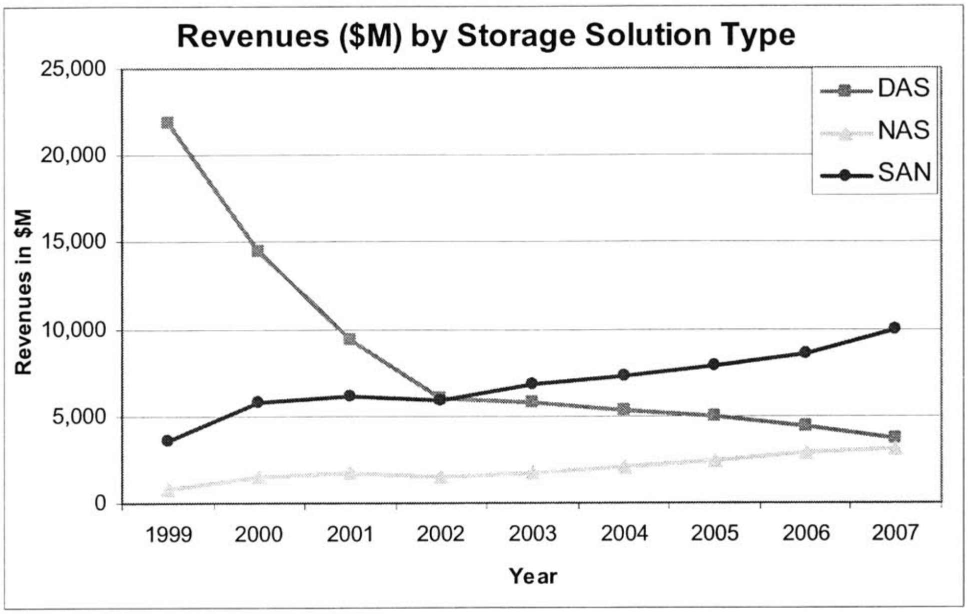


Figure 5: Comparison of annual revenues by storage solution type (Source: IDC, 2003)

2.4.1 Direct Attached Storage

Direct attached storage is the simplest, and up until 2001, the most commonly used storage model. Although it still retains some market share, DAS has become less prevalent in new enterprise-level implementations, and is currently limited to small to medium type implementations. In a typical DAS configuration, the server is directly connected to one or more hard disk drives (HDD) or disk arrays (Figure 6). Historically, this connection was mainly done through a SCSI interface, although recently, Fiber Channel (FC) has all but replaced SCSI interfaces in most products. In fact, new storage arrays from EMC, HDS, and HP have dropped support for SCSI. DAS solutions are well suited for small implementations with a limited number of storage systems and servers. As these numbers start to increase, and with that the complexity of managing and locating the data, the DAS approach starts to lose efficiency. As far as the I/O is concerned, the server-array communication happens at the block I/O level, while the client computer communicates with the servers via higher level File System (FS) I/O protocols. Figure 6 shows a simple example of a DAS implementation. Managing such a solution from a software point of view is relatively simple. The storage management software would normally run on a server that is connected to the array; the server would be used to manage by proxy the array's functionality such as data replication, backup, and recovery.

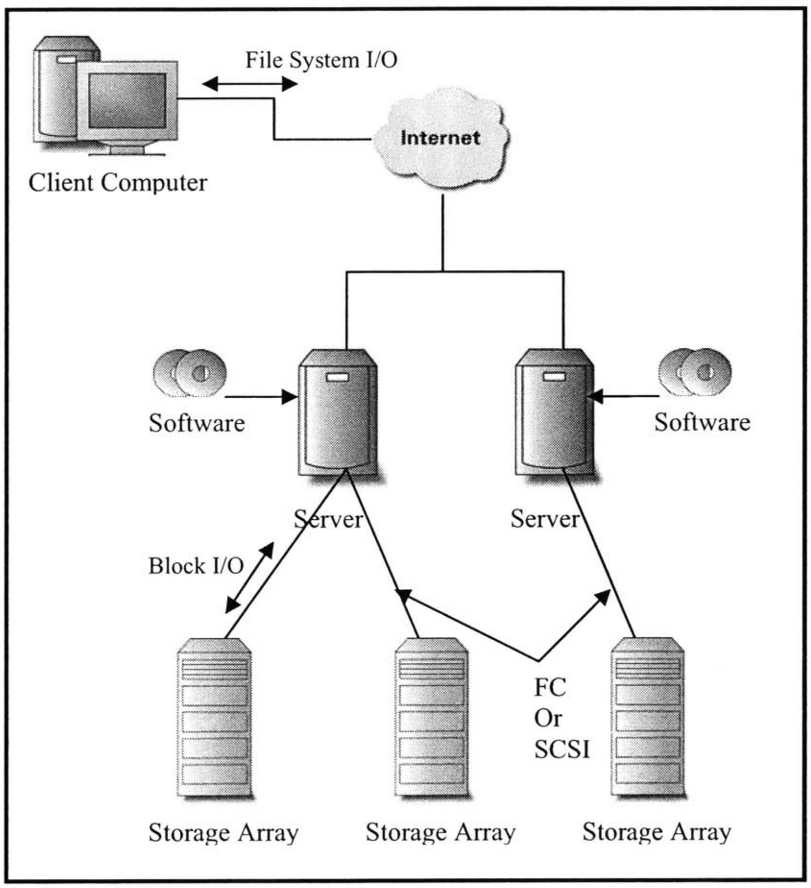


Figure 6: Example of a Direct Attached Storage (DAS) setup.

2.4.2 Network Attached Storage

The core concept behind a NAS solution is that of a storage system attached to the data access points through a network fabric (Figure 7). There are two types of NAS implementations. The first consists of a NAS appliance which combines both the storage disks as well as a server that exposes the storage to the network. The second implementation consists of a storage array or a SAN on the back end, with a separate NAS head or gateway as interface. Once attached to a network, both types of NAS behave in similar fashion, but the main difference is in the internal composition of the server's interaction with the disks at the lowest level. In one case, it is an all in one system; which puts a limit to the growth of the system as far as storage capacity is concerned. While in the other case, the gateway, which is essentially a server with special software running on it, is separate from the disks and can therefore accommodate more expansion in capacity over time. The I/O mechanism between the NAS gateway and storage arrays, and inside the NAS appliance is block level I/O; while the appliance and gateway expose FS I/O capabilities to the attached network on the front end. Figure 7 shows both types of NAS implementations, and how they connect to the network. Note that as it was mentioned previously and also shown in the figure, a SAN, which is discussed in more detail in the next section, can itself be the storage component of a NAS solution.

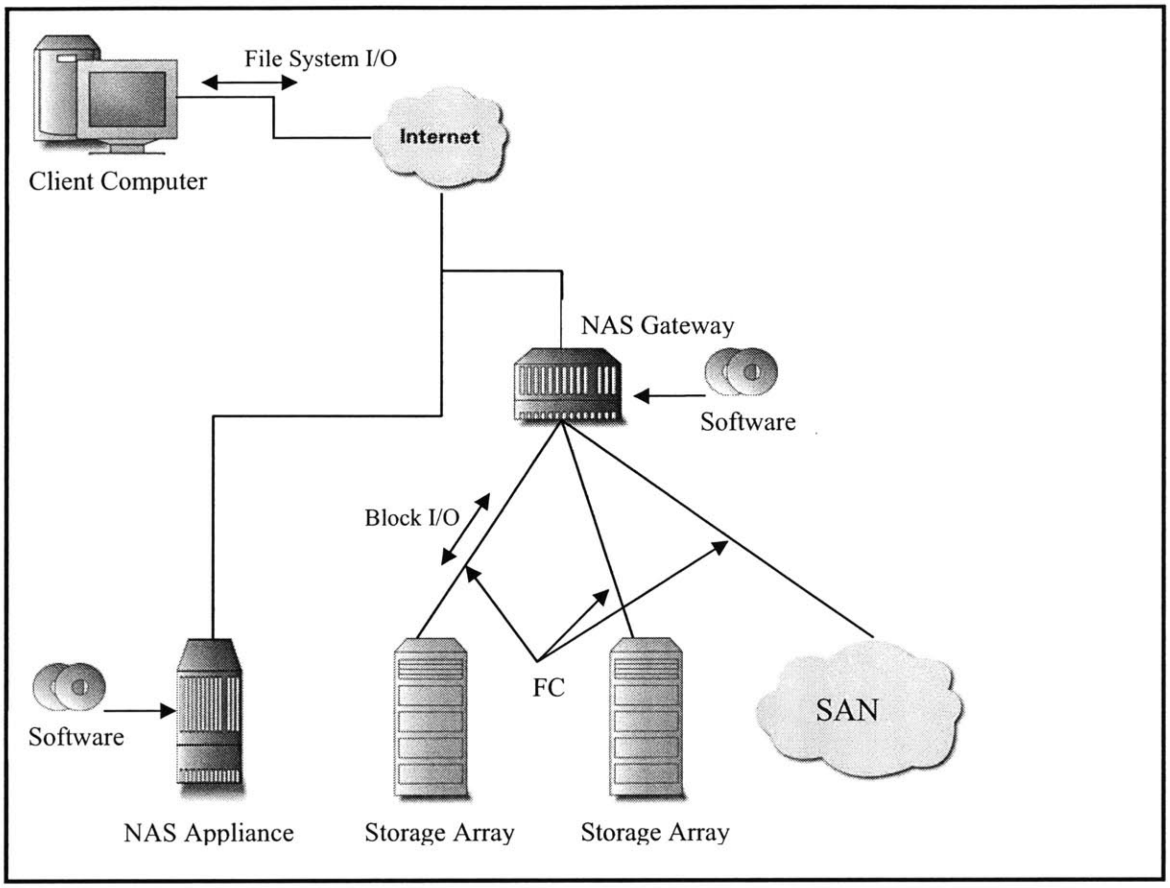


Figure 7: Example of a Network Attached Storage (NAS) setup.

2.4.3 Storage Area Networks

The simplest SAN can be thought of as a system consisting of a storage array connected to a server through a switch. Although this is a broad definition, it helps explain the basic SAN functionality. SANs connect storage and computers through a network of switches, hubs, and routers to facilitate the transmission of raw blocks of data (Figure 8). They can provide compelling advantages: storage consolidation, centralized management, improved data access, flexibility for growth, and security. In a SAN, servers have a high-bandwidth connection between each other and their respective disk storages; moreover, these connections can be modified, without being physically moved, through switch configuration. The I/O format between the storage arrays, the switches, and the servers is all block I/O, and similar to the other two implementations, it becomes FS I/O to the front-end of the server. Another characteristic of SANs is that backup devices such as tape libraries and ATA disk backup systems can also be plugged into the switch fabric, and the connections configured through the switches. With the complexity of connections comes the need for elaborate software to make all the components of a SAN work together. This software is usually vendor specific and targeted towards the control of the individual subsystems as shown in Figure 8. As mentioned previously and shown in the trends of Figure 5, SAN solutions are currently experiencing the fastest growth when compared to NAS solutions which are growing at a slower rate, and DAS solutions which are actually on a rapid decline.

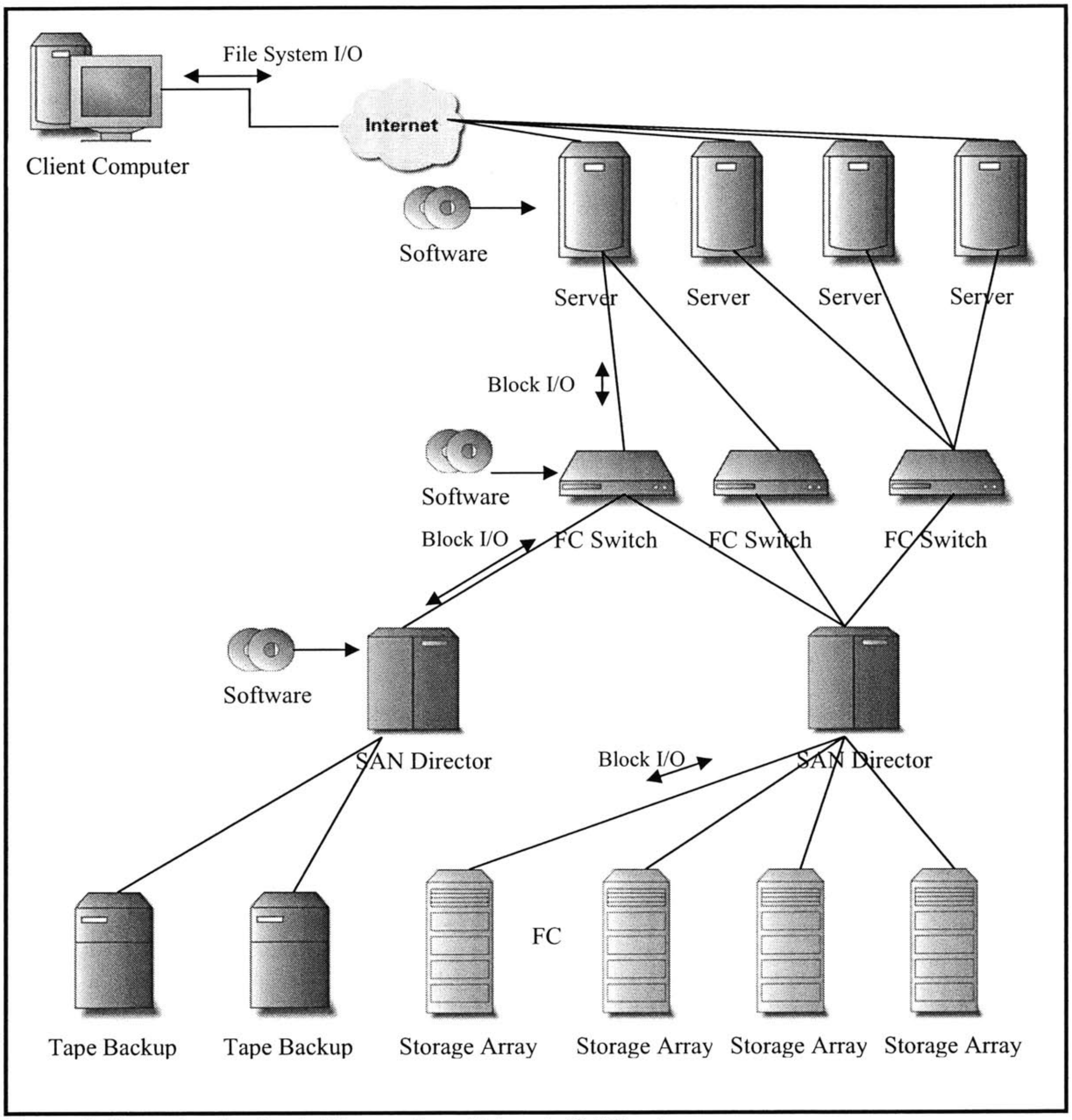


Figure 8: Example of a simple Storage Area Network (SAN) setup.

2.4.4 iSCSI Networks

Although they are not as prevalent as the other three (DAS, SAN, and NAS), iSCSI storage solutions are becoming more and more widespread. iSCSI offers one of the same benefits that NAS offers, namely that it uses an existing IP network and does not require a specific fiber fabric. Contrary to NAS though, it does not require a NAS gateway nor that the storage array be a NAS appliance. Instead, the storage array is a normal high-end or midrange array with iSCSI ports that are connected to an IP network. This exposes the devices that are mapped to the iSCSI port to the network. The server's HBA is an iSCSI HBA that is also connected to the network and configured to access the IP address of the storage iSCSI connection. Given that iSCSI implements the SCSI protocol, when the configuration is setup, the end result is that the server can access the storage devices as if they were directly connected to it via a SCSI (or FC) cable. Figure 9 shows what such a setup would look like.

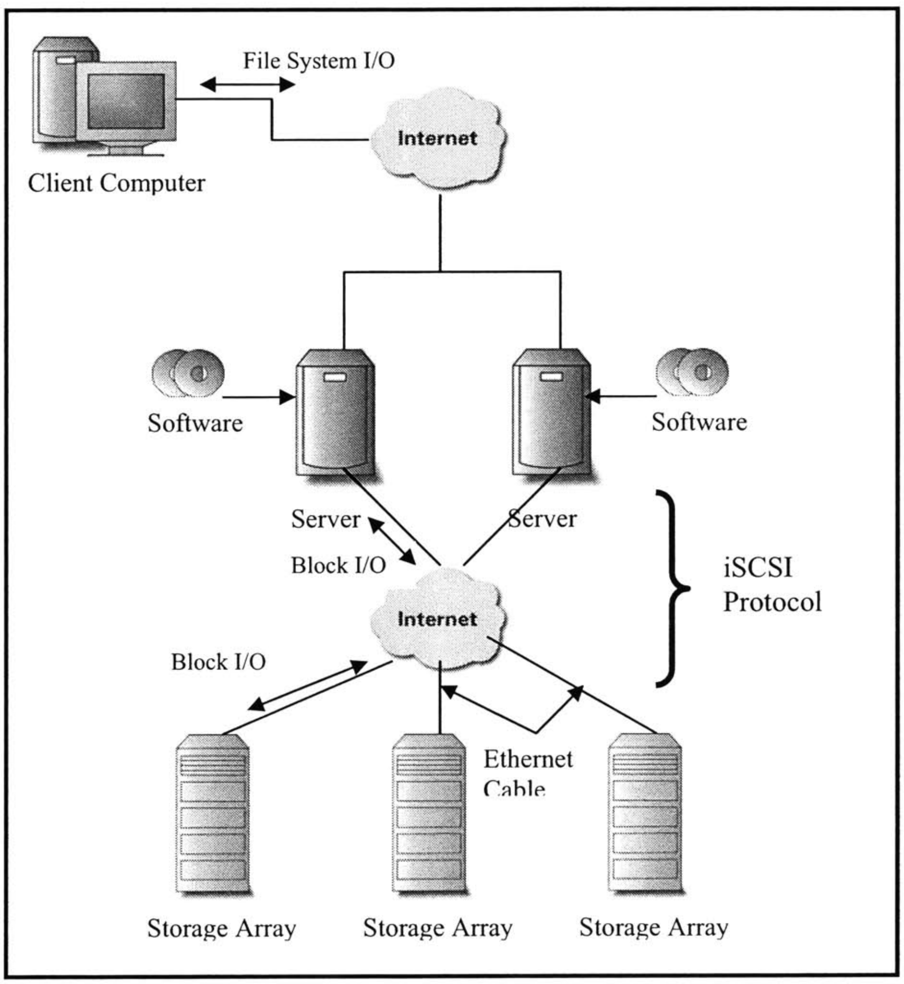


Figure 9: Example of a simple iSCSI setup.

2.5 The IT Challenges

The challenges faced by data storage IT customers are the result of two major trends.

Growth of Magnetically Stored Data

The first of these trends is the continual growth of magnetically stored data. According to a study conducted by researchers at the University of California at Berkley, hard disk storage, which is the medium of interest for the data storage industry, accounts for about 38% of the magnetic data stored annually. Furthermore, HDD-based storage is experiencing a high rate of increase as clearly shown in Figure 10 (see Table 10 for capacity unit definitions).

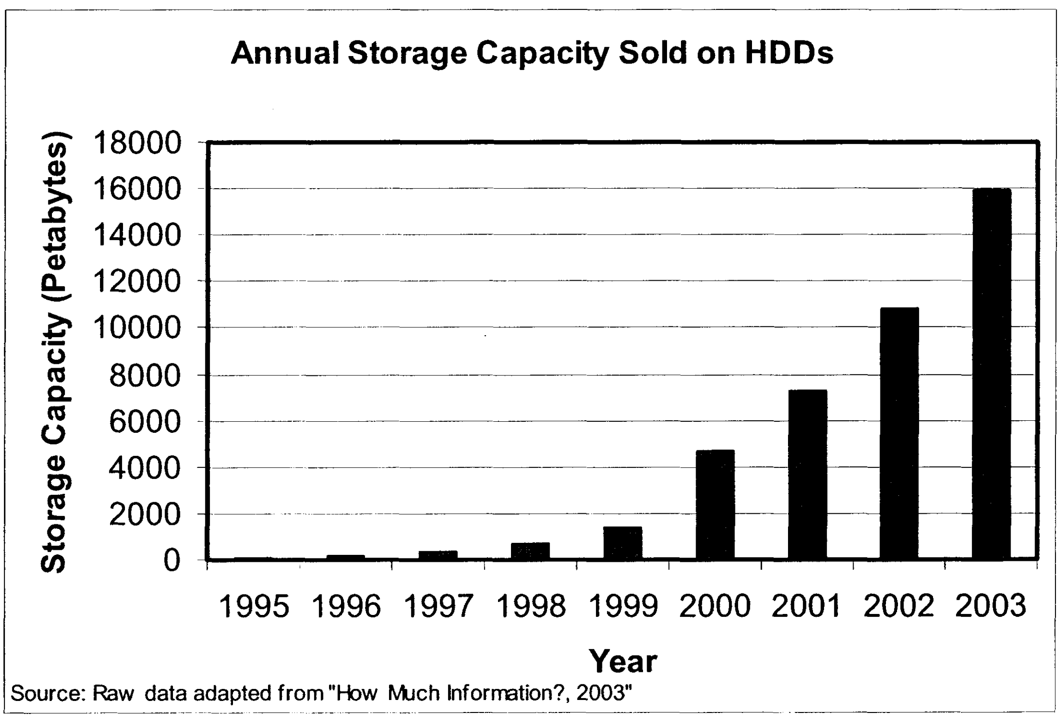


Figure 10: Trends of annual HDD storage capacity.

Several factors are contributing to this growth, some of which are listed below:

- The increase in size of the World Wide Web (WWW).
- More and more historically paper-based transactions are being performed electronically.
- Government regulations regarding the mandatory archiving of emails, phone conversations and online chat messages, especially for financial institutions.
- The increase of magnetically stored data increases its criticality, and creates a need for backups for this data, which is also stored magnetically.

Increase in System Functionality

Another trend is the increase in functionality of the hardware storage components provided by the storage industry vendors. In each new product, new functionality is added, while the old functionality is usually maintained. Moreover, in order to be able to deal with the increase in data, these systems need to be aggregated, which then yields meta-systems with increasing complexity and cost of management. The resulting dynamics are shown in the simplified causal loops diagram of Figure 11. Note that this model is simplified to only represent the three major loops of interest, and is intended to be purely an insight model rather than a simulation one.

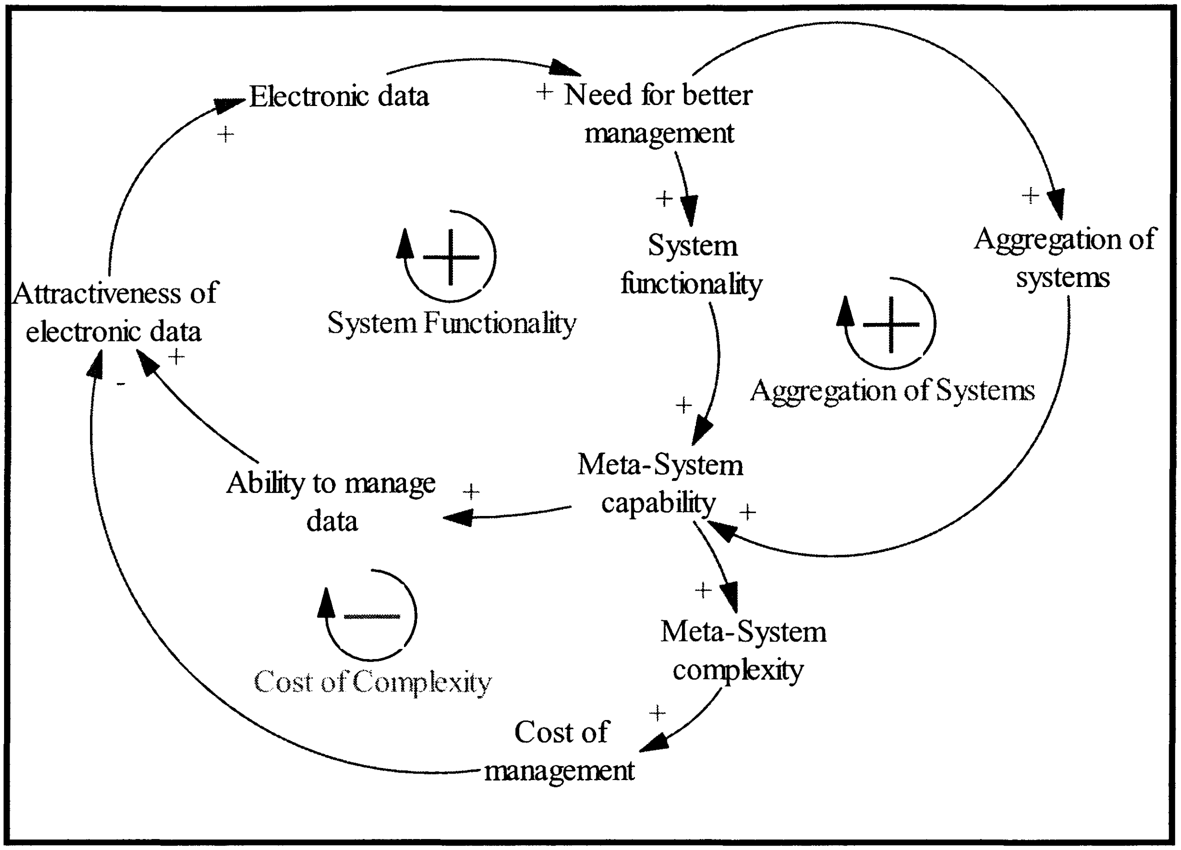


Figure 11: Simplified causal loops diagram of storage functionality and complexity.

The *System Functionality* loop is a positive feedback loop. It indicates that as the amount of electronic data increases, new functionality is added to the storage systems to facilitate the management of this data. In turn, this increases the attractiveness of electronic data, because it is now manageable.

The second positive loop is the *Aggregation of Systems* loop. It operates based on the same premise as the previous one, except that instead of new functionality being added to the systems in order to deal with data proliferation; the systems are aggregated to create more capable meta-systems.

The third loop is the *Cost of Complexity*. It is a negative loop that attenuates the growth that is driven by the previous two. With the aggregation and increase in functionality to deal with the data, comes an increase in complexity and cost of the management solutions. This reduces the attractiveness of adding more data and therefore dampens the growth of the underlying economy. Reduction of this complexity and cost of management is the main request of IT clients; and it is on this problem that the industry players should focus in order to fuel the overall industry growth.

3 The Data Storage Business

This chapter is aimed at providing the reader an understanding of the business side of the data storage industry. It is not meant to provide a comprehensive coverage of all aspects of the business. Instead, the aim is to provide the information necessary to be able to put later discussions of struggles over standards and platform leadership in perspective. In order for the business dynamics to be fully understood, a basic knowledge of the underlying technology, such as the one presented in chapter 2, is necessary. The flow of this chapter is as follows. First, it starts with an overview of the industry leading firms and their products. Next, it provides an explanation of the competitive and complementary dynamics between these various firms and lays out some dynamics of the value chain. Finally, it offers an explanation of the current trends of shift from hardware to software, and the resulting impact on the value chain.

3.1 The Industry and the Industry Leaders

Up to this point, the term *'Storage Industry'* has been mentioned in a manner that might have given the impression that the boundaries of this industry are well defined. In reality, this is not at all the case. The boundaries of the industry are highly dependent on the points of interest of a particular analysis or discussion. In fact, these boundaries actually mimic the boundaries of the technical solutions, in that they have a high degree of intersection. For example, in an enterprise IT architecture, one cannot accurately define where the storage part of the architecture ends and the networking part begins. The analysis of the industry will therefore be limited to the firms that produce the core system of a storage solution: the storage array.

3.1.1 Storage Arrays

The discussion of the storage array business can be broken down into three separate categories: high-end arrays, midrange arrays, and networked storage.

High-End Arrays

The high-end array market was dominated by IBM until the advent of then-newcomer EMC in the mid-nineties. EMC had introduced a superior product, Symmetrix, with which they were able to capture much of the market share that IBM and other companies held, and establish themselves as market leader. This momentum of increased market share and revenues continued its tilt in EMC's way until the market downturn of 2001. It was fueled by a growing storage market, a superior product, and the considerable hardware margins they were able to charge for it. In 2001 however, they entered into a price battle with IBM who had introduced the Shark system. This battle led to the erosion of margins and revenues. HDS then introduced a product that was able to capture market share from both of its competitors.

In 2003, the size of the high-end array market reached $5.417 Billion[6]. It is dominated by the most prominent storage firm, EMC, which controls 34.9 percent of market share with its flagship DMX product (see Figure 12). The other major players in this market are IBM, Hitachi Data Systems (HDS), HP, and Sun.

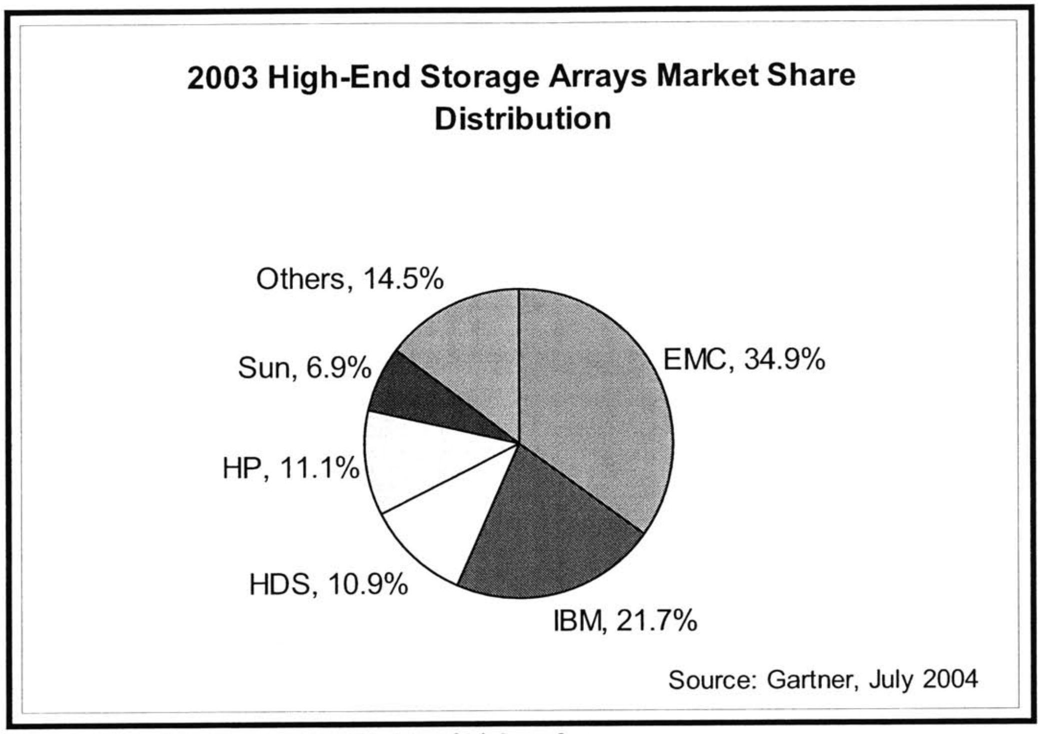


Figure 12: Market share distribution of high-end storage arrays.

The competition variables in this market are the following:

- Price
- Capacity
- Performance
- Reliability
- Functionality
- Interoperability with other products

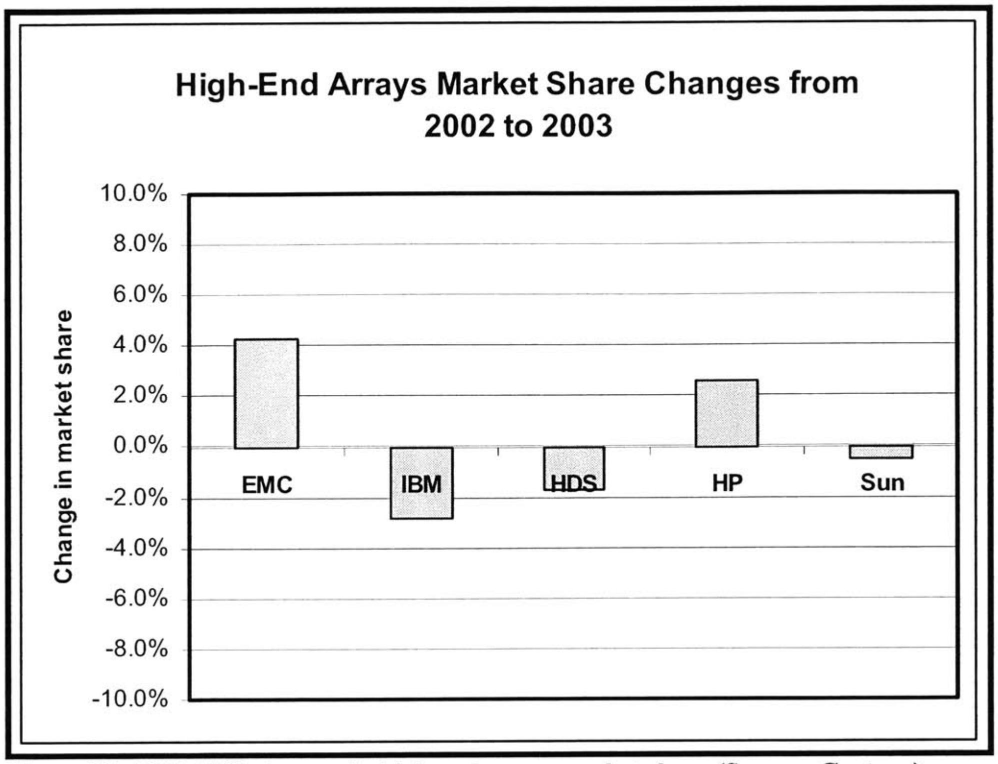


Figure 13: 2002-2003 changes in high-end array market share (Source: Gartner).

Although fierce competition takes place; year to year shifts in market share remain low (see Figure 13) because of two major factors. The first factor is the implicit lock-in created by the array vendors. The high price of storage arrays, as well as the criticality of interoperability and the lack of standardization create a lock-in effect with the customers when time comes for them to expand and make new purchases. The second factor that limits the market share shifts is the leapfrog nature of the competition. Products introduced by these firms continually leapfrog each other on the basis of the aforementioned competition variables.

The main source of revenues for firms in this sector comes from the hardware sales of the physical arrays. In addition, these firms develop software that allows access to the internal functionality of the arrays such as various types of data replication and data backups. This complementary software (API) is licensed based on functionality, which results in additional sources of revenue.

Midrange Arrays

Midrange arrays are continuously expanding upward and downward in performance, capacity, and functionality; which is leading to the blurring of the lines of division between the markets. This expansion is promoting the entrance and sustainability of many players in the market. The criteria for competition are similar to the ones listed above for the high-end arrays, except that the target customers are generally smaller companies, instead of large enterprise IT shops.

In general, the midrange arrays market is more crowded than the high-end one. Table 1 is a list of the competitors in this market; they range from recent startups to well established veterans. The list includes all the firms listed in the previous section, which indicates that firms that compete in the high-end market also compete in the midrange. This certainly was not the case a few years back. EMC for example did not have a presence in this market until it acquired Data General in 2000. In addition, as highlighted in Table 1, there are many alliances, partnerships, and OEM relationships in this market: Dell and EMC, StorageTek and Engenio, and Sun Microsystems and Dot Hill Systems.

Midrange firms	Relationships with other Firms (color coded)
EMC	Co-branding agreement with Dell
HDS	
IBM	OEM FastT from Engenio
HP	
Network Appliances	
Dell	Manufactures EMC's Clariion and co-brands it
Engenio	OEM to IBM OEM to StorageTek
Sun Microsystems	Some joint development with Dot Hill Systems
StorageTek	OEM from Engenio
XIOTech	
3PARdata	
Dot Hill Systems	Some joint development with Sun Microsystems

Table 1: Competitors in the midrange array market (Source: Gartner)[7].

Network Attached Storage

The Network Attached Storage (NAS) market is a much smaller market than the previous two; it is however experiencing the fastest growth. There are few competitors in this market, all of them also competing in the high-end and midrange markets. The most dominant of these competitors is Network Appliances with 48% market share[8]. In second place comes EMC, followed by HP and Dell.

3.2 Competitors and Complementors

There are no fully vertically integrated firms in the data storage industry; which means that an IT customer cannot build a SAN environment from a single source vendor. Storage solutions are therefore composed of heterogeneous systems, supplied by multiple vendors, who produce complementary products. There are several factors that are necessary for these products to successfully interoperate:

- Existence of standards for communication between the various hardware and software products.
- The adherence of the various firms to these standards.
- The implementation of integration qualification tests to ensure interoperability and eliminate potential bug from final integrated systems.

On the other hand, in addition to their attempt at gaining market share, most firms in this industry are also trying to diversify their product portfolios. The logical result is an increase in competition when the portfolios intersect. Hence, the industry is experiencing a combination of complementary dynamics and competitive dynamics. The effects of this duality will become more interesting in later discussions of standardization and platform leadership.

3.3 The Shift from Hardware to Software

Hardware prices in the data storage industry have experienced a considerable decline and are continuing to do so. In the storage array market for example, traditional metrics have ceased to be the main points of competition. This is due to the fact that the performance and functionality of competing products are very close and are hardly differentiating factors anymore. Moreover, the market downturn and the bursting of the dot-com bubble have resulted in major cuts in IT spending, which has forced customers to search for the cheapest solution to meet their need. At the same time, storage solutions complexity has been rising. The number of systems in a SAN environment and the complexity and functionality of each system has increased the overall complexity of the storage network and generated a need for better overall management capability. The *'bigger and faster'* traditional requirement has been replaced with the *'easier and cheaper to manage'* requirement, and demand for better management tools has topped customer surveys. These dynamics of falling hardware prices and rising system complexity have forged the way for a booming software industry in the data storage arena. Traditional hardware vendors were looking for new markets to continue their growth and compensate for falling prices and the commoditization of their hardware. Customers were looking for solutions to their complexity problem. Software was the answer to both.

In essence, the business model is starting to morph into one that is comparable with the Polaroid and Gillette models. Polaroid sold cameras at a loss in order to increase its install base, and thus increase the demand for instant film. Film sales generated most of the revenue for the company. Gillette has a somewhat similar model with razors and razor blades. The razors install base enables the sales of the razor blades, which generate most of the revenues. This is what Katz and Shapiro (1994) refer to as the "hardware/software" paradigm; and they list many examples of such dynamics in various industries including credit cards, durable goods, and typewriters[9]. In the data storage industry, the market is starting to exhibit similar characteristics. The hardware sales are ceasing to be the main -and sometime only- source of revenue. They are expected to establish the install base for the firm. This install base in turn will create a network that promotes the sales of software to support the hardware, for additional revenue generation.

4 Proprietary Standards versus Open Standards

This chapter is a discussion of the struggle between proprietary standards and open standards, and is divided in to two sections. The first section outlines the proprietary standards efforts and their failure, while the second section traces the history of open standards in data storage management.

4.1 Proprietary Standards and their Failure

This section will start with an overview of the EMC WideSky initiative. Next, a discussion of EMC's quest for establishing a dominant design is presented; followed by a discussion of the challenges of establishing platform leadership in general, and of EMC's attempt at achieving that in particular. Finally, I review EMC's eventual abandonment of the WideSky initiative in favor of open standards.

4.1.1 EMC's WideSky

The most notable attempt at proprietary standards came about in late 2001, when EMC announced its plans for WideSky. Up until then, storage hardware vendors offered storage management software APIs (Application Programming Interface) that allow the management of their own hardware products. WideSky was a middleware that was poised to go beyond providing an interface to the management of EMC's Symmetrix and Clariion storage systems, and allow users to control third party storage systems such as those of Hitachi, HP, and IBM. Figure 14 depicts where the WideSky middleware fits in the storage solution stack.

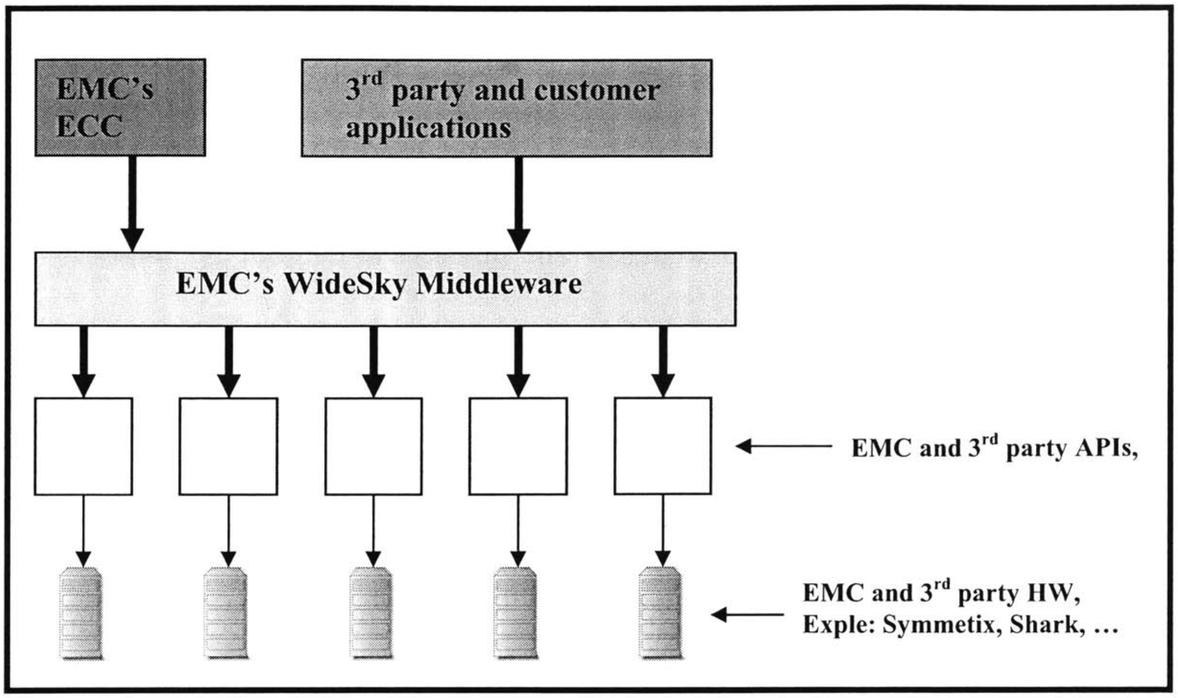


Figure 14: The WideSky middleware in the storage solution stack.

The following is an excerpt from the news release announcement on October 29, 2001:

> *"EMC Corporation today announced WideSky™, the industry's first storage management middleware technology to address the complex issues of managing a multivendor storage infrastructure. Through a new access mechanism, WideSky masks the underlying complexity of multiple vendors' products, including storage systems, network devices and host storage resources. By using WideSky, EMC and its partners will create powerful, integrated storage management applications that address customers' entire multivendor storage environments..."*[10]

The implications of the announcement reverberated accross the industry. While EMC hailed it as the solution IT customers have been longing for, competitors feared that such a product would become the de facto standard in the industry and would give EMC a considerable competitive advantage.

EMC's approach to managing third party storage arrays was two fold. Their primary and more preferable model was that of cooperation, in which vendors would give EMC access to their APIs. This would allow EMC to wrap these APIs with the WideSky interface and expose that to end users. In exchange, EMC would also provide its API to those vendors. This is what is called API-swap, and although not everyone agreed to it, EMC was nonetheless successful in convincing some vendors at adopting this approach. The second and less preferable model is a more technically challenging one. In instances were EMC was unable to secure the cooperation of array vendors, a more brute force methodology was used. The solution was to create wrappers around the command line interface (CLI) that these vendors provided with their product. This approach is more difficult and less efficient, but nonetheless, it allowed EMC to go past the non-cooperation hurdle. Doug Fierro, EMC's director of enterprise storage networking described it as follows:

> *"It's easier to cooperate, but we can do it on our own without the help of our competitors. We'll go the reverse engineering route if we have to."* [11]

4.1.2 The WideSky Release and the Reaction of Competitors

In March 2002, EMC finally announced the availability of WideSky. The press release statement provided a list of companies that supported the WideSky initiative.

"...Companies supporting the WideSky initiative include: BMC Software, Brocade Communications Systems, Inc., Citrix Systems, CommVault Systems, Computer Associates, Legato Systems, McDATA Corp., Microsoft, Novell, Oracle Corp., Precise Software Solutions, QLogic Corp., Quest Software, SAP, SAS and Sybase, Inc..."[12]

This list spanned a wide range of storage related industries including switches, host bus adapters, backup management software, and databases. It did not include however any storage array vendors and companies such as IBM, HDS, HP, and Compaq were clearly missing from it.

Being the market leader in high end storage arrays, this release was EMC's first step at establishing itself as the platform leader in data storage management by making WideSky the storage management platform product. The supporting companies provided complementary products to this platform-to-be, and everyone realized that with the advent of standardization, although proprietary, the networked storage solutions market would begin to experience considerable growth and everyone would stand to profit from that. It is clear that the limiting factor to such growth is the lack of standardization in storage management. This prevents interoperability between storage arrays, and makes the management of heterogeneous storage solutions comprised of multiple storage array types extremely difficult.

On the other hand, IBM, HDS, HP, and Compaq were in direct competition with EMC in the storage array market, and the dynamics of this market were changing at the time. The price war that took place in 2001 as these companies fought for dominance had considerably reduced the profitability of storage hardware. Moreover, IT customers started demanding more than additional functionality, additional speed, and additional capacity. These historical metrics could no longer be used as the differentiating factors between products because of the enterprise level IT shift towards management and reduction of complexity; which could only be achieved through software. The storage array vendors were looking for new sources of revenues and having realized that the platform leadership was to be established at the software management level, they were poised to prevent EMC from establishing itself as the leader with WideSky. Steve East, vice president of storage integration as HDS clearly stated that by saying:

> *"We'd love to have EMC manage some features, but we're not going to make their product stronger by giving them access to our arrays."*

4.1.3 The Quest for the Dominant Design

The term 'dominant design' was first introduced by Utterback and Abernathy[13] in 1975. It is the technological path that emerges as the dominant one, among a set of competing technologies. The emergence of such a design among a set of competing technology paths has the effect of changing the competitive landscape of the industry. Suarez and Utterback argue that this will reduce the population density in that area and decrease the chance of survival for new entrants[14]. In the case at hand of storage management, the competing technology paths were not only differentiated by the underlying technology itself but also by the strategy of the companies, or at least for one of them, EMC. To explain, a specific area of storage is chosen: storage arrays. The technology embedded in storage arrays and in the software APIs that allow the management of these arrays greatly differs from company to company. These companies competed on creating better storage arrays, as measured by the conventional storage metrics or performance, capacity, and functionality. The business model was to sell storage arrays, which generate the largest portion of the revenues, and then sell APIs, CLIs and GUIs to manage them. Such a struggle is not one of establishing a dominant design, but rather one of establishing product dominance and increasing market share and profits.

In 2001, through the announcement of WideSky, EMC embarked on adding yet another technology path to the competition mix. The path towards a dominant design was based on three of the four levers that were proposed by Suarez and Utterback:

- Possession of collateral assets
- Existence of bandwagon effects or network externalities in the industry
- Strategic maneuvering at the firm level

Possession of collateral assets and network effects

EMC could draw heavily on its collateral assets such as its sales channels, its brand name in data storage, and its established market share leadership especially in the high end arrays (see Figure 15). It would be able to push WideSky onto at least its existing customer base, and from there, expand to the rest of the market. Through WideSky, EMC would be also solving one of the biggest impediments to the growth of the networked storage industry, the complexity of its management. A survey shown in Figure 16 clearly indicates that management and interoperability are key requirements for

IT customers, more so than the traditional performance and functionality metrics. It is generally accepted that solving the complexity problem would allow the industry as a whole to grow at a much faster pace, and would increase revenues for many other firms in the industry. The prospect of such growth encourages some of these firms to line up in support.

Strategic maneuvering

In addition, the push towards dominance lies in EMC's 'strategic maneuvering'. Cusumano, Mylonadis, and Rosenbloom discussed the strategic maneuvering that a firm undertakes in order to establish a standard in a competitive market[15]. EMC's strategy was to shift the focus from a direct approach of gaining market share through the conventional performance-capacity-functionality front, to an indirect one of establishing WideSky as the dominant design. By providing the capability that allows end users to manage third party hardware through its interface APIs, EMC's aim is to make its interface the one of choice. By controlling the interface, which is becoming the most critical aspect to managing the complexity of storage networks, it would essentially control the market. Although a dominant design is different from an established standard, WideSky would have the potential of morphing into a de facto standard. Such a standard is quite different from a regulated one, in the sense that it is established because of the industry strength of the firm that is behind it, and it is adopted because of the same reasons[16]. A good definition is the one proposed by Hemenway when he describes it as 'that which is accepted for current use through authority, custom, or general consent'[17].

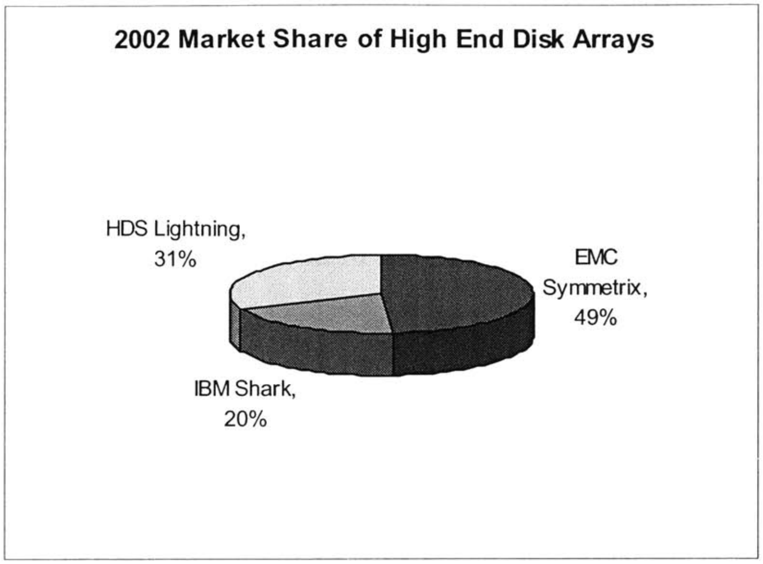


Figure 15: 2002 market share distribution of high end arrays[18].

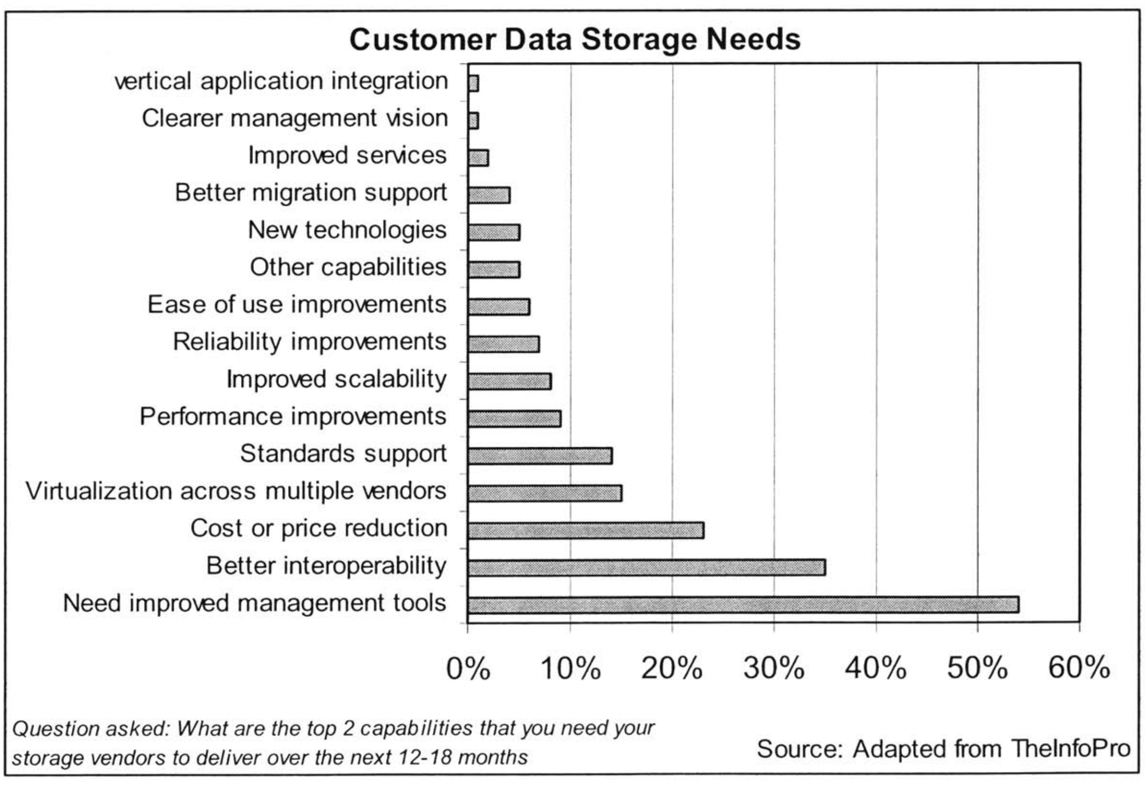


Figure 16: Rating of customer data storage needs.

4.1.4 The Challenge of Establishing Platform Leadership

The Risks

"The game is risky."[19] That is what Gawer and Cusumano said when referring to the game of establishing platform leadership. The following sections attempts to examine this risk from a general perspective. At a glance, one can identify three main components of interest when referring to risk: the magnitude of the positive outcome in case of success, the magnitude of the negative outcome in case of failure, and the probability of success and failure. The assumption to start with is that the firm makes a conscious effort into establishing platform leadership. The risk in this case stems from the number of factors that the firm has to manage and that are in essence outside the full control and sometimes the reach of the firm. But before discussing these factors in more detail, it is important to focus on the negative fallout associated with failure in the attempt to establish this leadership. The cost of this fallout can be broken into the three categories that are listed below in order of complexity:

- Direct loss
- Missed opportunities loss
- Loss of leadership

The *Direct Loss* refers to the wasted resources, both time and human, associated with the effort. These resources are spent on technology development, building partnerships and alliances, developing marketing channels, and public relations work. It is true that the firm could capitalize on some of this effort to be used in other venues, even after failure, but some of it will infallibly go to waste.

The *Missed Opportunities Loss* is associated with wasted opportunities that the firm could have embarked on but didn't because of its focus on platform leadership. An example of such opportunities would be going after niche markets for example, which could prove to be very lucrative. Another example is going after a vertical integration strategy in delivering full solutions, rather than a horizontal platform. Moreover, in the pursuit of a certain technological path that fails, the firm might find itself locked out of others, either because of being late to market, or because of falling behind technically.

The *Loss of Leadership* to another entity, that becomes the platform leader, may or may not be directly related to the efforts of the firm to establish itself as the platform leader, depending on whether or not it is the firm that had started a platform leadership war. In any case, the result is the same, and the firm might find itself operating in a market in which a competitor has claimed the platform leadership spot.

Technology Strategy

The section above details the outcome in case of failure at establishing platform leadership, but the question about what influences success or failure still remains. Embarking on an effort to establish platform leadership usually goes beyond the core competencies of a firm. In order to reach a point where it can consider undergoing such attempts, the firm would have built over time a considerable technology arsenal, and this arsenal is one – if not the most important one – of its core competencies. In addition it is possible for the firm to have acquired other competencies such as product development expertise, manufacturing expertise, marketing and sales channel development, industry wide partnerships, and so on. These competencies allow the firm to compete and excel in the marketplace, and even establish market share leadership, but it is not enough to ensure platform leadership. For the latter to happen, a specific well balanced platform strategy needs to be implemented or else the firm risks the fate of many famous failed attempts, such as Sony's Betamax format for the VCR and Apple's Macintosh PC, to name a few. Gawer and Cusumano propose the 'four levers of platform leadership' as a framework that a company can use to develop a new strategy or enhance its existing platform strategy. These levers need to be adjusted and optimized for a platform approach to successfully materialize. Below is a list of these levers:

1. ***Scope of the firm**: What to do inside the firm, and what to let external firms do.*
2. ***Product technology**: Make decisions regarding system architecture (the degree of modularity), interfaces (the degree of openness of the interfaces to the platform), and intellectual property (how much information about the platform and its interfaces to disclose to outside firms).*
3. ***Relationships with external complementors**: How collaborative versus competitive should relationship with complementors be? How will consensus be created? How will conflicts of interest be handled?*
4. ***Internal organization**: How to organize the firm to support the above three levers.*

I will use some of these levers as lenses to examine EMC's attempt at establishing WideSky as the storage management platform leading interface. The analysis will be limited to looking through two of the four lenses. The two lenses to be excluded are the *'scope of the firm'* lens and the *'internal organization'* lens. The *'scope of the firm'* lens is excluded because not much value can be extracted from this analysis. The decisions regarding what to develop inside the firm and what to develop outside the firm are dictated by the technical specification of the WideSky API. Simply put, third party vendors are expected to continue developing native API interfaces to their products, and EMC develops WideSky as a wrapper interface to the native APIs. The other lens to be excluded is the *'internal organization'* one. The reason for that is the lack of insight into EMC's organizational changes that were implemented to accommodate the platform effort. Without such knowledge, an intelligent and insightful analysis cannot be made. It is worth mentioning though that EMC was not dichotomous in its approach at proprietary versus open standards support. Instead, the approach that was pursed was one of fully backing and influencing the development of open standards all the while developing WideSky.

On the other hand, the two lenses to be used in the analysis are the *'product technology'* lens and the *'relationships with external complementors'* lens. Since the effort is known to have failed, the focus will be on *what went wrong*. From the *'product technology'* view, one could clearly see that the approach still promoted individual innovation in the sense that WideSky did not attempt to control other firms' development efforts but rather interface to and manage their products. This was one of the positive aspects of this approach. The negative aspect however, is that it ends up being a one way street of communication. By not developing or inherently having the capability to influence other companies' products or the interfaces they exposed to their products, each one of these companies was still able to optimize its own products, but the entire system that is composed of these products managed under the WideSky interface was not necessarily optimized. This is a classic case of optimization at the sub-system level at the expense of the whole. It does not mean that WideSky was doomed to fail because it was not technically optimized as a whole, although this did reduce its chances for success when presented to end users.

Compounded to the technical challenges are those that become apparent when one examines WideSky through the *'relationships with external complementors'* lens. One of the key metrics to the success of EMC's initiative was the ability to manage third party storage arrays. Without that ability, the product (WideSky) would be of no great added value to customers. The third party storage arrays producers are therefore complementors, but in this case: ***The complementor IS the competitor***. The relationship with the complementor/competitors was therefore marred by a feud because of the power WideSky would give EMC over them. Moreover, EMC did not take the 'gradual low-key approach' recommended by Gawer and Cusumano when a firm pushes its innovation, and by that, they failed to muster the political rally that is much needed for such an endeavor. In addition to the lack of support by storage arrays, the lack of support was also visible on the software front, when Veritas Software, another major competitor to EMC, also failed at supporting WideSky. The situation could be summed up by this quote from Gary Bloom, the CEO of Veritas Software, in an interview with VARBusiness.

"We don't see any need to support WideSky. We are already an open independent software vendor that supports a broad range of hardware devices. If you really look and dissect their announcements for partners, they essentially don't have any storage suppliers partnering. They don't have Hitachi. They don't have Sun. They don't have IBM. They don't have HP. If you don't have the storage providers, how much industry momentum is there? I'll never say we won't support it. It is just right now there is not enough industry momentum behind WideSky that suggests we need to support it. The real question about WideSky is, will anybody else use it? Is Hitachi going to use it for their storage architecture? Is IBM going to use it? I don't believe they will."[20]

The *'relationships with external complementors'* lens also unveils another problem, which can best be explained by first looking at the WideSky value chain in Figure 17.

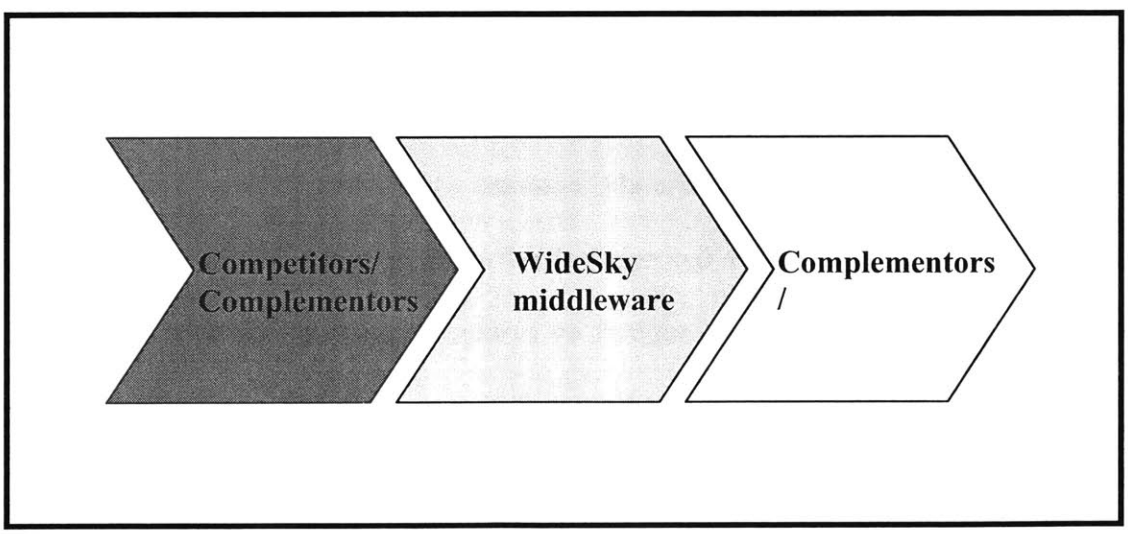


Figure 17: Simplified diagram of WideSky value chain.

It has already been shown how downstream of the WideSky middleware, the value chain breaks because of the '*complementor is the competitor*' situation. In addition, upstream of the middleware the value chain also breaks, but for a different reason. WideSky has the capability of delivering value to a set of upstream customers or partners who are in a position to build complementary products on top of WideSky. These upstream complementors differ from the downstream ones in the sense that rather than fitting within WideSky, such as being a product that WideSky can manage, they are built on top of it and deliver direct value to end customers. The problem is that no such complementors were being included in the strategy and they are clearly missing from the following list:

> *"…Companies supporting the WideSky initiative include: BMC Software, Brocade Communications Systems, Inc., Citrix Systems, CommVault Systems, Computer Associates, Legato Systems, McDATA Corp., Microsoft, Novell, Oracle Corp., Precise Software Solutions, QLogic Corp., Quest Software, SAP, SAS and Sybase, Inc…"*

The analysis clearly shows major issues with managing the competitors and complementors during the WideSky initiative. Whether these problems are related to the approach of the firm and could have been better strategically managed, or whether they are an inherent characteristic of the construct of the industry and thus unavoidable, could be the subject of further study and exploration. For now, it suffices to say that the results of the platform leadership push were unsuccessful.

4.1.5 WideSky is Abandoned

In September 2003, almost two years after the first WideSky related press release, EMC announced that it was abandoning its WideSky initiative. The need for standardization of management and a middleware layer that could manage groups of heterogeneous storage arrays was still there, but EMC opted out of its plan at doing it alone. Two major factors led to this change in strategy. The first factor was related to the technical challenges arising from the reluctance of the other vendors at sharing their interfaces and their approach of restricting EMC's access to their APIs. The second and more important factor was that at that time, another specification for storage management was being developed in parallel. Unlike WideSky, this specification was an open standard, worked on and supported by most vendors, including EMC. The Storage Management Initiative Specification (SMI-S) was being developed under the umbrella of the Storage Networking Industry Association (SNIA). The difference between the two standards can be best accentuated by positioning them on the standards matrix. Henderson[21] proposed this matrix as a two-by-two mapping where one axis is the technology, which can be 'open' or 'closed' and the other axis is the ownership, which can be 'public' or 'private'. SMI-S is positioned as an open technology with public ownership, while WideSky is an open technology with private ownership. The matrix itself and the criteria for each position are presented in Figure 18.

OWNERSHIP	TECHNOLOGY: Open	TECHNOLOGY: Closed
Public	Details of standards are available to all: no single firm has control over how they evolve: no charge for their use. **Eg:** **TCP/IP, HTML, SMI-S**	Standards are owned and controlled by the public sector but are not freely available **Eg:** **Cryptography**
Private	Details of standard are made available to all: but owner has control over how the standard evolves and may charge for use **Eg:** **Palm OS, Nintendo, WideSky**	Technology may be standard, but details are not made available beyond the firm **Eg:** **IBM 360 Arch.**

Source: Original matrix from "Competing in Standards Driven markets", R. Henderson

Figure 18: SMI-S and WideSky on the standards matrix.

One could also argue that WideSky was an important catalyst for the support that SMI-S received, and that without WideSky, such a quasi-unanimous and strong support for an open standard would not have occurred at that point in time. In any case, EMC's competitors rallied behind this open standards effort to fight WideSky, and their efforts proved successful when the latter was finally abandoned. At the same time of announcing the end of the WideSky initiative, EMC also announced that it stands behind SMI-S:

Executive VP Mark Lewis said recently that the company now favors SMI-S, the Storage Networking Industry Association's Storage Management Initiative Specification, and API swaps. The SMI Specification is expected to be adopted later this year.[22]

4.2 History of SMI-S

4.2.1 PDP and the Bluefin Specification

In early 2002, the Partner Development Process (PDP) group presented Bluefin and made a proposal to the Storage Networking Industry Association (SNIA). The PDP was formed by a group of storage involved companies that included notable members such EMC, IBM, HP, Veritas, and Computer Associates amongst others (see Table 2 for a full list of participating companies). The PDP group had been working on a specification codenamed Bluefin. Bluefin was based on the Common Information Model (CIM) and was developed in order to standardize the management of storage components in a storage network and allow interoperability in the management of multi-vendor systems. The initial specification of Bluefin was complete, and the PDP wanted to hand it over to SNIA. Their expectation was that SNIA would own it, and extend its development under its umbrella.

PDP Member Companies	
BMC Software	Hitachi
Brocade	IBM
Compaq	JNI
Computer Associates	Prisa Networks
Dell	QLogic
EMC	StorageTek
Emulex	Sun Microsystems
Gadzoox	Veritas
Hewlett-Packard	

Table 2. List of PDP member companies.

4.2.2 SNIA, SMI, and the SMI-S

SNIA, which stands for Storage Networking Industry Association, was formed in 1997. It is a body whose main objective is the advancement of the adoption of networked storage solutions. It is composed of an ensemble of member companies some of which - the ones with voting power- are listed in Table 3. Among its many activities, SNIA sponsors technical work groups that address various storage areas such as fiber channel, NAS, security, IP storage, and backup. It is also the producer of the Storage Networking World Conference. As far as facilities are concerned, SNIA is the host for a vendor-neutral technology center in Colorado Springs, Colorado.

> *"The Storage Management Initiative (SMI) was created by the Storage Networking Industry Association (SNIA) to develop and standardize interoperable storage management technologies and promote them to the storage, networking and end user communities"*[23]

SNIA adopted the Bluefin specification as the basis for its Storage Management Initiative Specification (SMI-S), which is a standard for the management of storage and is being developed by the Storage Management Initiative (SMI). SMI was created by SNIA in 2002 and is currently comprised of about 50 member companies. The aim of the specification is to allow the interoperability of heterogeneous networked storage products supplied by multiple vendors, and to enable the management of these products.

Large Voting Members	Medium Voting Members	Small Voting Members
Brocade Communications Systems	Adaptec	Arkivio
Cisco Systems	ADIC	Cloverleaf Communications
Computer Associates	AMCC (formerly JNI)	CreekPath Systems
Dell Computer	American Megatrends	Crosswalk, Inc.
EMC Corporation	AppIQ	Cyrca Solutions Inc.
Hewlett-Packard	Atempo, Inc.	DataCenter Technologies
Hitachi Data Systems	ATTO Technology	EqualLogic, Inc.
IBM	BakBone Software, Inc.	Exanet
Intel Corporation	CommVault Systems	HIFN, Inc.
Microsoft Corporation	Computer Network Technology	Intransa
Oracle	Crossroads Systems	iStor Networks
Quantum, Data Protection Division	Dot Hill Systems Corp	Knowledge Transfer
Seagate Technology	Emulex Corporation	Lefthand Networks
StorageTek	FalconStor	MonoSphere, Inc.
Sun Microsystems	Engenio	Neartek, Inc.
VERITAS Software	McDATA	Permabit Inc.
	Network Appliance	Pillar Data Systems
	Nth Generation Computing	Rainfinity
	PANASAS	Red Hat
	QLogic	Revivio
	SANZ SGI	Sandial
	Syncsort	SANRAD
	Vitesse Semiconductor	
	XIOtech Corporation	
	Xyratex	

Table 3. List of voting SNIA member companies (source www.snia.org)

In August 2003, SNIA announced the release of SMI-S version 1.0. This announcement occurred about one month before EMC announced its plans to abandon WideSky in favor or SMI-S. In April 2004, through a well concerted effort across the industry, major vendors such as EMC, IBM, HP, Sun Microsystems, and Hitachi Data Systems made a slew of announcements regarding conformance to the SMI-S (see Table 4). This strong support by these IT giants added to the strength of the prospects of SMI-S.

Date	**Company**	**Announcement**
April 5 2004	EMC	*"EMC Corporation today announced that its flagship EMC ControlCenter and EMC Visual families of storage resource management software support the SNIA Storage Management Initiative Specification (SMI-S)... EMC has successfully demonstrated SMI-S management of storage arrays from Dell, HDS, HP, IBM, LSI Logic, SUN and STK."*[24]
April 6 2004	IBM	*"IBM today announced that the IBM TotalStorage Enterprise Storage Server has been certified with the latest industry storage management standard from the Storage Networking Industry Association (SNIA), known as the Storage Management Initiative Specification (SMI-S) v1.0.2. In a further show of focus on industry standards, IBM announced it plans to have its entire disk storage server portfolio, including the FAStT Storage Servers, SMI-S conformant by the end of 2004."*[25]
April 6 2004	HDS	*"Hitachi Data Systems today announced that the Hitachi Lightning 9900™ V Series, Thunder 9500™V Series storage systems and a prototype of the HiCommand® Device Manager 3.x have been certified as fully compliant with the Storage Management Initiative Specification (SMI-S) following an evaluation using the Storage Networking Industry Association Conformance Testing Program (SNIA-CTP)."*[26]
April 6 2004	SUN	*"Sun Microsystems today announced that its Sun StorEdge products, from enterprise to workgroup offerings, have passed the Storage Networking Industry Association Conformance Testing Program (SNIA-CTP). This achievement validates Sun's efforts to deliver interoperability and the resulting customer benefits of reduced cost and complexity through the implementation of the SMI-S in its storage products."*[27]
April 16 2004	HP	*"An HP interoperability conformance test has become the first tool accepted by the Storage Networking Industry Association (SNIA) for multi-vendor use in assessing the conformance of industry products to the Storage Management Interface Specification (SMI-S)..."*[28]

Table 4. Announcements of conformance to the SMI-S by major vendors.

4.2.3 Interaction between WideSky and SMI-S

The history of WideSky and of SMI-S clearly shows a power struggle between proprietary standards pushed forward by EMC, and open standards backed by the industry as a whole (including EMC). Part of the aim of the open standards support was to fend off EMC's attempt at establishing a de facto standard that could disrupt the business model of EMC's competitors should it succeed at developing a software interface that can manage third party storage hardware. As it was mentioned previously, EMC was pursuing two paths simultaneously. The first path is that of WideSky and of proprietary standards, while the second was that of SMI-S. EMC's strategy of dual pursuit could be justified as follows. When WideSky was first announced, the SMI and for that matter the SMI-S had not come into existence yet. EMC was fully committed to WideSky because of the potential gains it would stand to reap from its success. At the same time, EMC hedged its risk by also allocating some resources and efforts into advancing SMI-S because should SMI-S come succeed, the cost of being left out of it would be huge and would come in three folds: a public relations cost portraying an image of non-conformance to open standards; a technical cost of not influencing the specifications to better interface with EMC's products; and finally, a business cost of being late on the delivery of conforming products and risking being locked out of the market.

As far as the interaction between EMC's plan for WideSky and SNIA's open standards efforts, one could clearly detect reactionary steps taken by each side to the other side's actions. Figure 19 shows a timeline comparison between the evolution of proprietary standards and that of SMI-S. Early on, in 2001 and 2002, it seems that SNIA was reacting to EMC's moves, while two years later towards the end of 2003, as the struggle started coming to and end, the table are turned and it is EMC that is reacting to SNIA's moves.

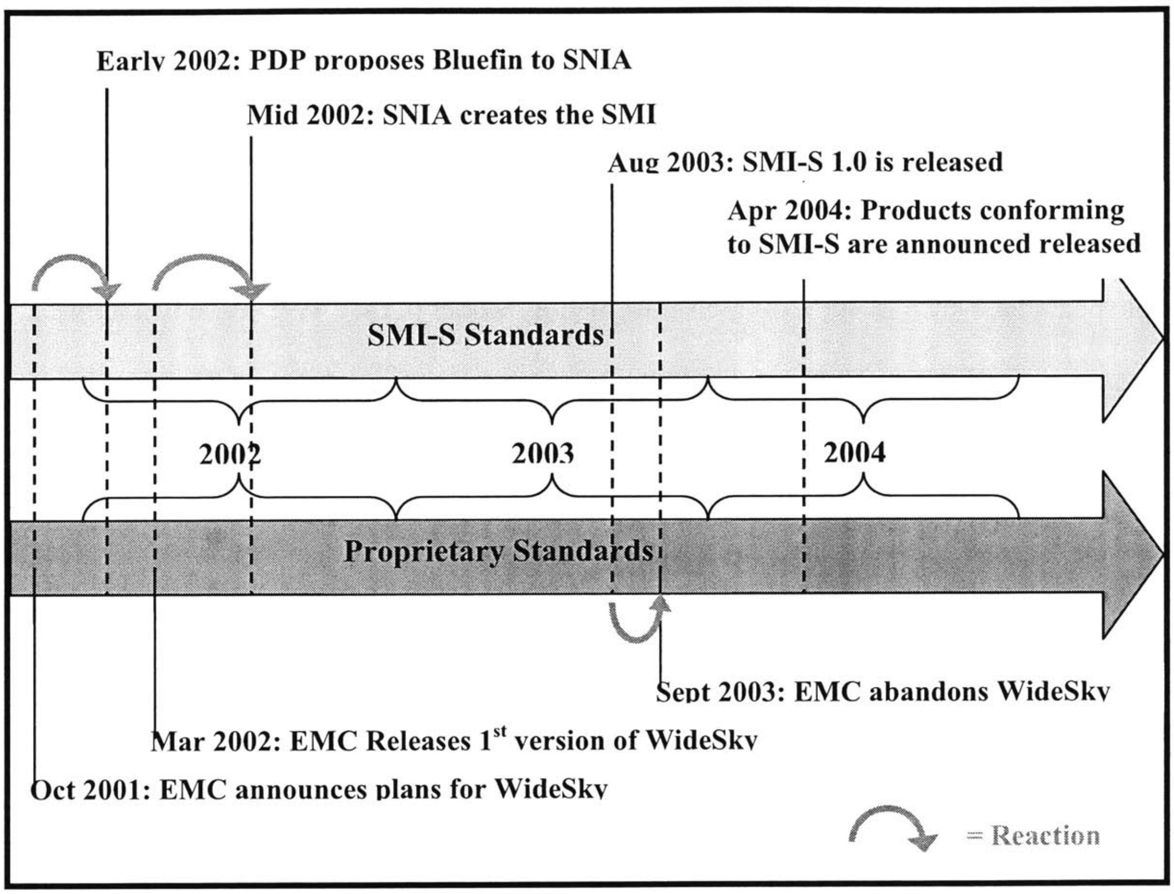


Figure 19. Timeline comparison of the evolution of storage standards.

5 The SMI-S Technology

This section aims at providing the readers with enough technical information about SMI-S and its underlying dependencies so as to allow them to understand the capabilities of the technology and the extent of its usefulness. The target reader is not someone who wants to use this technology, but rather someone who wants to understand the extent of its enabling capabilities and apply this knowledge in a business context.

The SMI-S was not built from the ground up. It has dependencies on other standards and implementations such as the Common Information Model (CIM), Web-Based Enterprise Management (WBEM), XML, and HTTP. In order to better understand SMI-S, a brief introduction of these dependencies is therefore necessary.

5.1 Hyper Text Transfer Protocol, HTTP

HTTP, the Hypertext Transfer Protocol, is a W3C specification which has become the de facto standard for exchanging messages over the internet. It is a stateless protocol, in which a requester sends a request for information using the *get* method targeted at a particular address identified by a URL (Uniform Resource Locator). The messages exchanged consist of meta-data provided in headers, the request/response line, and an optional body that contains the application data. Once a client sends a *get* request, the server at the location responds with the data that is requested via a Get response message, or with the relevant error message if applicable. Other methods that HTTP provides include *post* and *head*. Post is used to send data to the server that is expected to update its databases, such as when a form is filled out. Head requests only the headers. An example usage of "head" is for deciding whether to do a full *get* by first finding out whether the server has a newer version of available data.

HTTP is stateless; the connection lasts only for the length of a single transaction. It is often necessary, however, to maintain state over a set of requests and responses. While HTTP itself does not provide this capability, multiple mechanisms exist in which state may be transmitted with the exchanged messages. Examples include the use of cookies and opaque tokens.

5.2 Extensible Markup Language, XML

XML is a cross-platform, text-based language for specifying structured information. Simplistically, an XML document is tagged data. It consists of nested "elements". Each element has a name, and may have attributes and/or content. The content itself may be other elements. Taking a data-centric view, XML separates style from content. We provide an example of an XML element below:

```
<item itemID="234345435">
    <receivedOn>12/14/2004</receivedOn>
    <priority>high</priority>
</item>
```

XML is extensible in that it allows one to define other languages using it. It is therefore often referred to as a "meta-language". This is its main strength. An XML based language must at least specify the names of its allowed elements, their attributes, and the cardinality and nesting relations of these elements. DTDs or XML Schema can be used for this purpose. Regardless of the language it is in, an XML document can be read using any XML parser without requiring the DTD or Schema. However, having that information enables the parser to automatically validate the document.

5.3 The Common Information Model, CIM

The Common Information Model is being developed and maintained by the Distributed Management Task Force, or DMTF. The DMTF is a body comprised of multiple members of industry. Its main objective is to develop internet and enterprise management standards. There are various levels of membership in the DMTF; the most important ones being Board membership and Leadership membership (see Table 5).

Board Members	Leadership Members
Cisco	Adaptec Inc.
Dell	Advanced Micro Devices
EMC	AppIQ
Hewlett-Packard Company	Argon Technology Corporation
IBM	BMC Software
Intel Corporation	Brocade Communications Systems
Microsoft Corporation	Computer Associates International
NEC Corporation	ETRI
Novell	Hitachi, Ltd.
Oracle	Motorola
Sun Microsystems, Inc.	Newisys, Inc.
Symantec Corporation	OSA Technologies, an Avocent Company
VERITAS Software	Peppercon AG
WBEM Solutions	RLX Technologies
	SAP AG
	VIEO

Table 5: DMTF Board and Leadership members.

CIM is an object-oriented model for the representation and management of information and is based on the Unified Modeling Language (UML). CIM provides for the representation of data and of associations in a commonly understandable format that is applicable to a wide array of information types and fields. In addition to the specification, a CIM schema is also defined. This schema defines the data model and allows for the establishing of a common framework. The schema is broken down into three models[29]:

The *Core schema*: describes information applicable to all areas.

The *Common schema*: extends the Core schema and applies to specific management areas, such as Systems, Applications, Networks, and Devices[30]. This model remains pretty stable, and changes or additions can only be made by the DMTF.

The *Extensions schema*: extends both the Core schema and the Common schema, and allows for technology specific extensions. These extensions are usually industry specific or company specific, and are at the complete control of whoever chooses to create them.

5.4 Web-Based Enterprise Management, WBEM

WBEM is an initiative led by the DMTF. Its aim is to develop a common way to manage computing systems and heterogeneous computing environments. It utilizes existing standards such as HTTP for communication, XML for data format representation, and CIM for information representation and management. Figure 20 is a simplified view of the WBEM architecture.

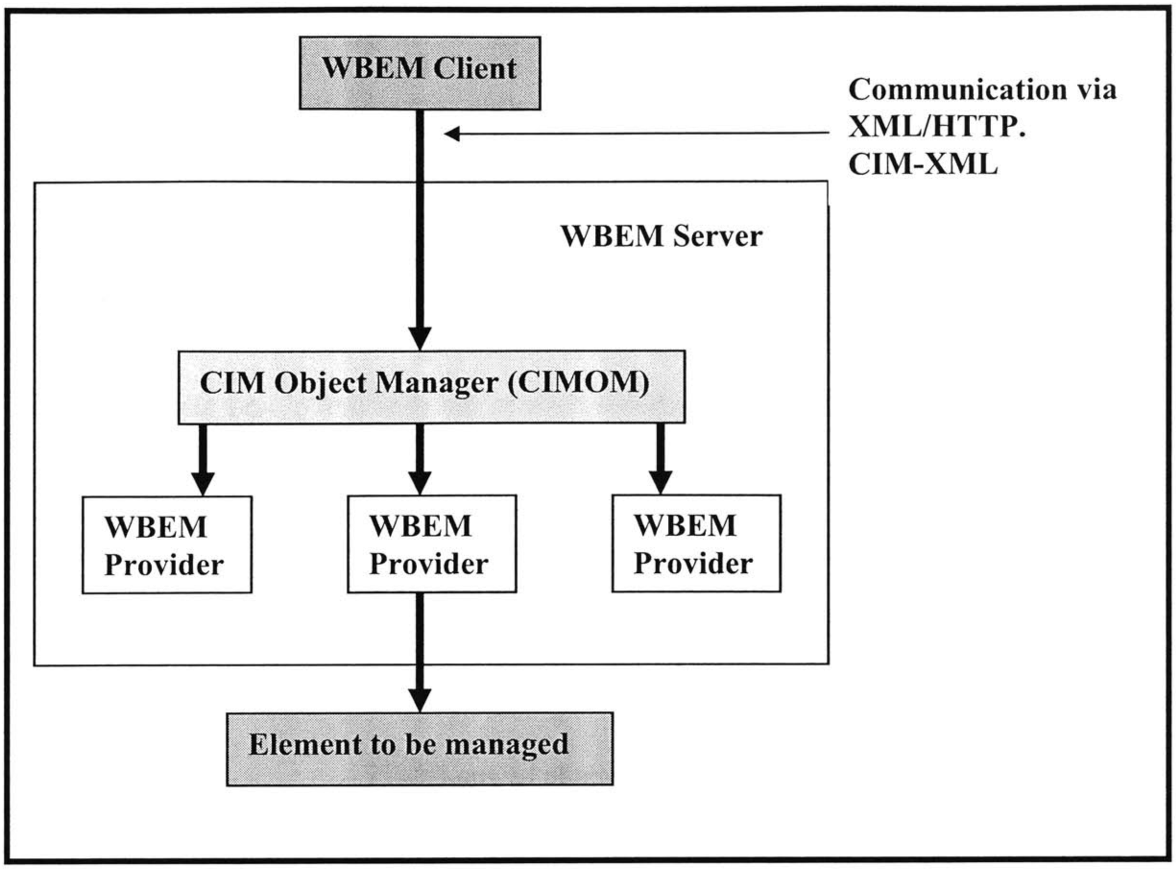


Figure 20: Basic architecture of a WBEM implementation (adapted from Sandlund[31])

- The *WBEM Client* can be a GUI or a third party application that wants to manage a WBEM resource. The front end of the client is application specific tailored towards the specific desired functionality; while the back end has to conform to the CIM-XML communication protocol.

- The *WBEM Server* is comprised of three major components:
  - CIM-XML is the communication protocol used by the WBEM server to interface with a WBEM client. CIM-XML allows the client to do WBEM

operations. All data transfer is done over HTTP, and the format of the data is xmlCIM. xmlCIM refers to the CIM data in XML format.

- CIMOM is the CIM Object Manager. It is the main component of the server that handles the management of the CIM objects.
- WBEM Provider is the part of the server that is specific to the elements that are to be managed. It is usually developed by whoever wants to expose and manage their interface with WBEM. It interfaces with the CIMOM on one end and with the actual elements to be managed on the other.

- The *Element to be managed* is the actual hardware and software application that a provider wishes to manage through WBEM. Examples of such elements are switches, servers, or storage arrays.

5.5 The Storage Management Industry Specification, SMI-S

The SMI specification is based on WBEM and therefore it is dependent on CIM, XML, and HTTP. SMIS specifies the information that a WBEM server must support and provide to its clients; with the focus being on networked data storage. Part of the specification is also a naming convention that reconciles the various vendor specific nomenclatures into a single uniform vocabulary. In practice, SMI-S compliant products are shipped with a WBEM provider extension that adheres to the specification. Firms don't spend resources on creating their own server; they normally use a generic server and extend it with their own specific SMIS compliant needs. There are several well known and widely used open source and proprietary WBEM server implementations such as Open Pegasus, Open WBEM, and WBEM Services. In addition, for functionality that is not yet part of the specification or that is vendor specific and is not generic enough to make it into the specification, the vendor is able to offer it through extensions; and the SMI specification provides for the creation of such extensions. In order to be considered SMIS compliant, vendors need to pass the Conformance Testing Program (CTP) that is administered by SNIA at its facilities in Colorado Springs.

The current release of SMI-S is version 1.0.2. The high level functionality that SMIS supports is called a 'profile'; and each profile is broken down into several subprofiles that further define detailed functionalities of the parent. The current capability of SMIS can therefore be summarized by the list of profiles in Table 6 [32]:

SMIS Supported Profiles
Fabric
Switch
Router
FC HBA
Host Discovered Resources
Array
In-Band Virtualization
Storage Library
Server

Table 6: SMIS Supported Profiles in version 1.0.2

5.6 Additional Aspects and Plans for SNIA and SMI-S

This section highlights some additional aspects of SNIA and SMI-S. The focus is on the current and future releases of the specification, the Conformance Testing Program (CTP), and the international expansion of SNIA.

5.6.1 Releases and Features

SNIA is continuing support for SMI-S through the Storage Management Forum (SMF). The current available release is version 1.0.2. At the same time, there are two releases in the works. Version 1.1.x is currently in the final stages or ratification and approvals. It is slated for availability in the middle of 2005. At the same time, version 1.2.x is in the specification phase. From a scheduling point of view, the aim is to have one release every year.

5.6.2 Conformance Testing

The Conformance Testing Program (CTP) was devised by SNIA to certify vendors' adherence to the standards. CTP consists of a set of test suites that are geared to run against a vendor's product. It tests the compliance of the product with one or more profiles of the specification. Once a vendor has received the certification, they get the right use the CTP mark on their product. Currently, there are 17 CTP certified vendors[33]: Brocade Communication Systems, Hitachi Limited, Cisco Systems, IBM, CNT, McDATA Corporation, Dell Computer, Network Appliance, EMC Corporation, QLogic, Emulex, Silicon Graphics, Engenio Information Technologies, StorageTek, Hewlett-Packard Company, Sun Microsystems, and Hitachi Data Systems. Figure 21 shows the percentage of SMI-S products that are CTP certified. Immediately after the concerted April 2004 announcements (previously mentioned in section 4.2.2), the percentage of certified products significantly jumped from 0% to about 70%, and kept continuously creeping up. Currently, it stands at 80%.

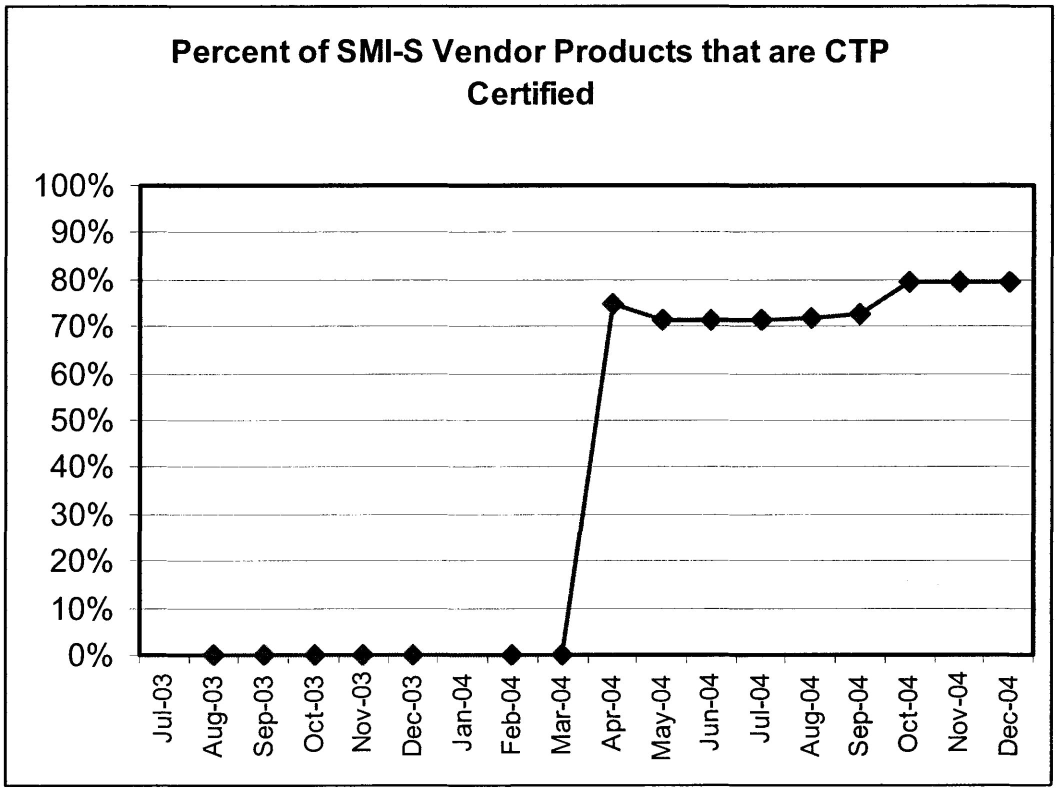


Figure 21: SNIA CTP certification trends of SMI-S supported vendor products[34]

In 2005, SNIA plans to introduce CTP for SMI client software. This partially highlights the increase of attention to client products, which are necessary to extract the value from the SMIS vendor compliance and deliver it to the end user. Thus far, vendor adoption has been occurring at a faster rate than client availability, so the success of the entire standards endeavor currently hinges on the success of the latter. A more detailed discussion of these dynamics will be presented in a later section.

5.6.3 Expansion of SNIA

Beyond its plans of expansion through membership, vendor support, and new specifications, SNIA is also going after a geographical expansion. Outside of the Unites States, there are currently six international SNIA chapters, some of which are country based, while others are regional:

- Australia and New Zealand
- China
- Europe
- India
- Japan
- South-Asia

6 Growth and Sustainability of the Standards

In the previous chapters of this thesis, a complete overview of the data storage industry was presented, both from a technology point of view as well as from a business point of view. This was followed by an analysis of the failed attempts at establishing platform leadership through dominant designs that are based on proprietary standards; and the resulting emergence of an open standards movement that is promoting SMI-S to be the industry wide standards for storage management. This chapter further extends the research effort and provides an analysis of the dynamics governing the evolution of the SMI-S standards. Given the current state of affairs, three years after the standardization effort was initiated and less than two years after the first version SMI-S was released, what are the factors (direct and self-reinforcing) that are aiding or hampering the adoption of these standards? Has SMI-S reached its tipping point for continued sustainability, and if not, when and under what conditions will it be expected to reach it in the future? In addition to answering these questions, this chapter also offers an analysis of the SMI-S standards as a disruptive technology, and uses this analysis to yield further insight into its predicted course of evolution.

6.1 Direct Factors Promoting the Growth of SMI-S

In the data storage industry, there are various factors at play that are fueling the sustainability of the SMI-S standards. An in depth discussion of the most important ones is necessary in order to better predict the standardization direction of the industry.

6.1.1 Fear of Vendor Lock-in

This push for standardization is indirectly reinforced by IT customers as a reaction to the fear of lock-in. The best way to illustrate such risk of lock-in is through an example:

> Suppose that company XYZ makes a $10 million purchase of HDS storage arrays to embed in a SAN solution. In addition, XYZ utilizes HDS HiCommand to manage the storage arrays. Two years later, XYZ plans to expand the deployed SAN with additional storage capacity. Although at this time IBM or EMC storage arrays might be superior in performance and functionality, XYZ yet might find itself forced to buy the HDS product.

The problem is that each storage array vendor exposes the functionality of its arrays through its management API and its own applications. Other than being able to discover third party storage systems, applications are optimized for the management of the arrays that are supplied by their own vendors. This means that in order to mix heterogeneous storage products and efficiently manage those, customers need to use the control applications provided by each vendor, and modify their scripts to interface with the APIs provided by each vendor. Furthermore, they have to develop in-house expertise in these applications and APIs, and this also comes at an additional time and monetary cost. The result is incurred expenses that could make the TCO of an optimal heterogeneous solution (multiple suppliers) higher than that of a heterogeneous sub-optimal one (single supplier). Naturally, IT customers want avoid this potential lock-in; and this is what a standard such as SMI-S would provide. The software management level standardization would remove the lock-in and allow price and performance based optimization at the hardware level. It is important to note however, that the customers are not directly asking for SMI-S standards. Their request is for better management tools that would allow interoperability between heterogeneous products. Thus far however, vendors in the data storage industry have been unable to fulfill this requirement mainly because of the technical difficulties involved in such an endeavor if undertaken in the absence of standards. But now that the SMI-S effort is underway, IT customers might start seeing their interoperability requirements fulfilled.

6.1.2 Fear of Technological Lockout

Another factor contributing to the support of standards is the vendors' fear of technological lockout. Merriam-Webster defines *lockout* as follows[35]:

> *Pronunciation:* 'läk-"aut
>
> *Function: noun*
>
> *the withholding of employment by an employer and the whole or partial closing of his business establishment in order to gain concessions from or resist demands of employees*

Although this definition expresses a theme similar to the one I am interested in, namely that of *someone* being unable to do *something* by being prevented from entering *somewhere*; it does not accurately express the desired meaning. For our purposes, I refer to the definition of technological lockout that was presented by M. Schilling[36]:

"...a situation in which a firm finds itself unable to develop or competitively sell products to a particular market because of technology standards."

Schilling further divides technological lockout into two types. Type I lockout occurs when in the absence of a dominant design, a firm adopts a proprietary technological path; but then reaches an obstacle in market penetration as the market starts converging towards the adoption of the dominant design. Type II lockout occurs in the presence of a dominant design that establishes a standard, when a firm is unable to enter the market with products that conform to the standard. It would be interesting to utilize the Type I lockout concept in analyzing EMC's dual strategy of developing WideSky while simultaneously backing SMI-S, at a time when there was no dominant design in the horizon. However, given that SMI-S is well underway and has the potential of becoming the dominant design in data storage management, it is clear that what the vendors fear is the Type II lockout.

Schilling further elaborates on the Type II lockout by presenting two factors that would increase the likelihood of its occurrence. The first factor is one where the firm is restricted out of the market by the competitors' patent protection; while the second factor is the firm's "lack of core capabilities and absorptive capacity". Given that SMI-S is an open standard being developed under the umbrella of the SNIA body; the first factor therefore becomes of no issue, and the focus shifts to the second one. As the dominant design starts emerging, the firm might fail to recognize it and consequently fail to invest in learning it and in integrating it within its core competencies. If and once this dominant design is established, in this case SMI-S becomes the universal interface through which storage is managed; the firm will find itself lacking the products, the technology to develop those products, and the knowledge necessary to develop the technology. It will be at a disadvantage with respect to the competition and late in the delivery of standards-conforming products to the market, while in the meantime, its existing technology and products are becoming obsolete. In a fast paced technology arena such as the data storage industry for example, the firm might never be able to catch-up. The results could prove to be disastrous, either to the division of the firm that is competing in this particular market, or to the overall firm if it is not well diversified. In order to exemplify this, a story comes to mind; it is that of Remington, Underwood, IBM, and the electric typewriter, which was recounted by Utterback[37] in his book, *Mastering the Dynamics of Innovation*. In 1920, Underwood was the dominant firm in the typewriter market. Previously, that role had been held by Remington. But in 1933, both companies had passed up buying Electrostatic Typewrites Inc. and its electric typewriter technology, which was bought by IBM. As the standard in the industry starting shifting towards electric typewriters, IBM had the advantage, while the other two were locked out, at least temporarily, until they had developed the core competence to enter the electric market. But it was too late, because by 1967 IBM had 60 percent of the market share of electric typewriters and 74 percent of the market of the high-end electric typewriters.

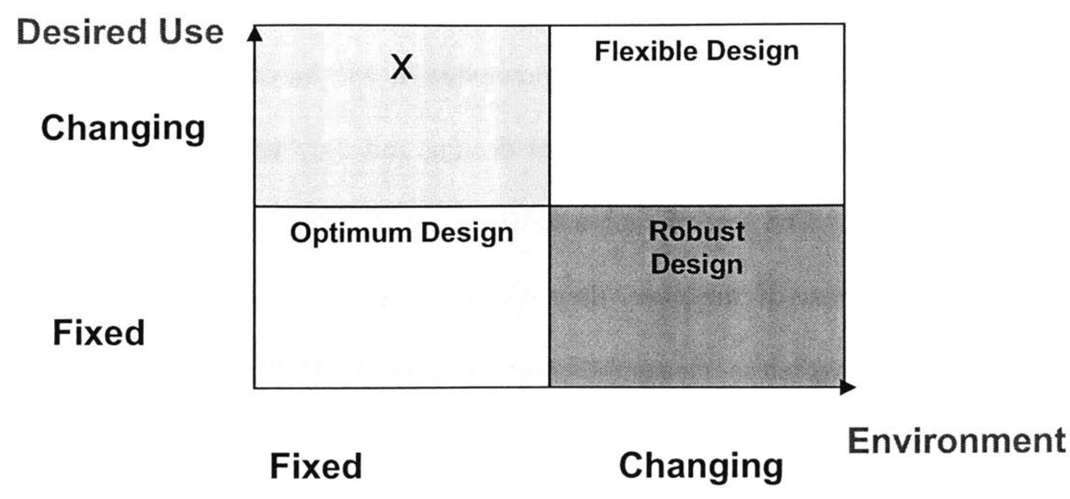
Data storage firms fear succumbing to similar lockout fate; and although the technology is different, the underlying dynamics are identical. If SMI-S becomes the adopted standard for storage management and SMI-S based software becomes the dominant design; firms that did not invest in learning and core competency development will be at a disadvantage. Without the technology and knowledge, it will be difficult for them to develop products that conform to this dominant design, and they will find themselves locked out of the market. The fear of such a fate is a big driver behind the adoption of SMI-S. Although in most of the case, they are not reaping any benefits from the compliance, firms are nonetheless releasing their products with SMI-S compliance.

6.1.3 Investing In Flexibility

System flexibility can be understood as the ability of a system to perform new functions at later stages of its lifecycle, in response to changes in its environment. Saleh more formally defines it as:

> *"...the property of a system that allows it to respond to changes in its initial objectives and requirements –both in terms of capabilities and attributes- occurring after the system has been fielded"*[38]

Figure 22 graphically illustrates this concept, and further explains flexibility by differentiating it from robustness. While a robust design is one that is able to handle changes in its environment to continue fulfilling its intended usage requirements sometime in the future; a flexible design can also handle changes in the environment, but is also able to fulfill temporally evolving usage requirements.



Source: adapted from previous work by Banerjee, Kommandur, and Saghbini, who adapted it from Saleh

Figure 22: Graphical illustration of flexibility.

In general, flexibility's attractiveness comes when there is uncertainty in the future requirements of the system and its operating environment. The system in this case is the actual data storage product that the firm is producing, or even the firm's organizational structure. By embedding SMI-S support in its products and investing in SMI-S learning and SMI-S technology development in its organization; the firm is in fact investing in flexibility.

Design Flexibility

Through its SMI-S compliant products, the firm would then be able to deal with changes in its environment, if SMI-S becomes the widespread standard of the industry; as well as changes in the requirements, if interoperability between heterogeneous systems becomes a must. It would be ready and able to support SMI-S functionality and competitively release SMI-S conforming products should the specification prevail as the standard of the industry, all with minimal additional investment. Flexibility however comes with an associated cost, and it comes at the expense of other system characteristics. In order to make a decision as to whether or not to incur this cost, the firm must be able to assess the *value of embedding flexibility*. Although the study of methodologies to compute this value goes beyond the scope of interest for this research, I will mention one such methodology as an example. The work is done by Banerjee[39] and builds upon earlier work on real options, by de Neufville[40] and others, in order to propose a method for computing the value of the option of embedding flexibility in a system using the real options approach. Utilizing such a methodology to compute the value of embedding the flexibility of SMI-S provider support in a storage system - given that the use and need of such functionality at the time the product is fielded is still very limited - would make an interesting research topic. On the other hand, one could argue that the firm's decision path for supporting SMI-S is not explicitly analogous to the logic presented in this section. A firm might not be calling it flexibility, and might not be trying to compute its value using the real options approach. However, the end result is the same, because the reasoning behind it is intuitively the same: "let us provide support for SMI-S because it seems to be on the uptake towards becoming the standard; and we want our products to be equal to those of our competitors' in terms of support". For the purposes of this analysis, we are mainly interested in the effects of the decision to embed flexibility; to summarize it, investing in flexibility (SMI-S support) yields contributions to the growth of SMIS.

Organizational Flexibility

In addition to designing flexibility in its products, the firm can also introduce this flexibility into its organization. This is done in the form of learning. The aim is to develop a knowledgeable R&D organization that will enable future inexpensive and rapid development of SMI-S based technology, should this technology become a widely adopted standard. From the firm's point of view, this early-on investment in learning is analogous to an investment in flexibility, along with all the previously discussed costs and benefits that come with it. On the other hand, from the industry's point of view (which is the point of interest here), these investments undertaken by the firm, result in a direct impact on the uptake of the standards. They yield an increase in the overall number of users and contributors to the standards.

6.1.4 Protection against Proprietary Standards

By backing the SMI-S open standards effort, firms are aiming to fend off any attempts at proprietary standards where a single firm is successful at establishing platform leadership through a dominant design based on proprietary standards. This is certainly true at least for the majority of firms who are not in the hunt for establishing their own standards for the management of heterogeneous systems. As a matter of fact, the previous discussions of the dynamics of evolution of EMC's WideSky and SNIA's SMI-S attest to that. The industry rallied behind SMI-S partially in order to fend off the imminent threat that EMC was posing through its solo introduction of WideSky and its vision of making WideSky the de facto middleware through which all other storage systems can be managed. The argument is therefore that the apprehension from seeing potentially one single firm dominate the market is a driver for the support of the open standards movement that is embodied by SNIA. For this to be true, we must get in terms with the fact that if an industry is highly standardized, the likelihood is highly reduced for a firm to have an established dominant design by means of its own proprietary standards. This does not mean that no dominant design can exist in the industry, or that once a standard is established, it cannot be overthrown by a proprietary standard backed by a single firm. What it means is that the dominant design in the industry cannot be *based* on a different standard that the established one. The following helps in explaining and exemplifying this.

In battles for standards and battles for standards based platform leadership, rare are the times where multiple standards can coexist and have an equal foothold in the market. It is not impossible, but historical evidence suggests that such a scenario is rare. Once a standard is widely adopted and a dominant design emerges based on this standard, the market dynamics become such that the majority of firms compete within the framework of this dominant design, until it is eventually overthrown. This holds true for formalized standards developed through planned consensus, as well as de facto standards that are a result of the continued and growing dominance of a proprietary implementation. Note that when referring to *dominant design* in this section, the actual reference is to a subset of dominant designs; those that are highly based on standards. Many examples can be cited in support. The battle for VCR format dominance between Betamax and VHS is one such case[15, 41]. The struggle occurred between the Betamax format introduced by Sony in 1975 and the VHS format introduced a year later by JVC. For details on the dynamics of the competition, I urge the reader to refer to the referenced papers, especially the one by Cusumano, Mylonadis, and Rosenbloom. The outcome however was that although being second to market, VHS dominated and Betamax all but disappeared and was confined to a small share in niche markets. Figure 23 shows how these two standards could not coexist, and how the VHS based dominant design for VCRs got established.

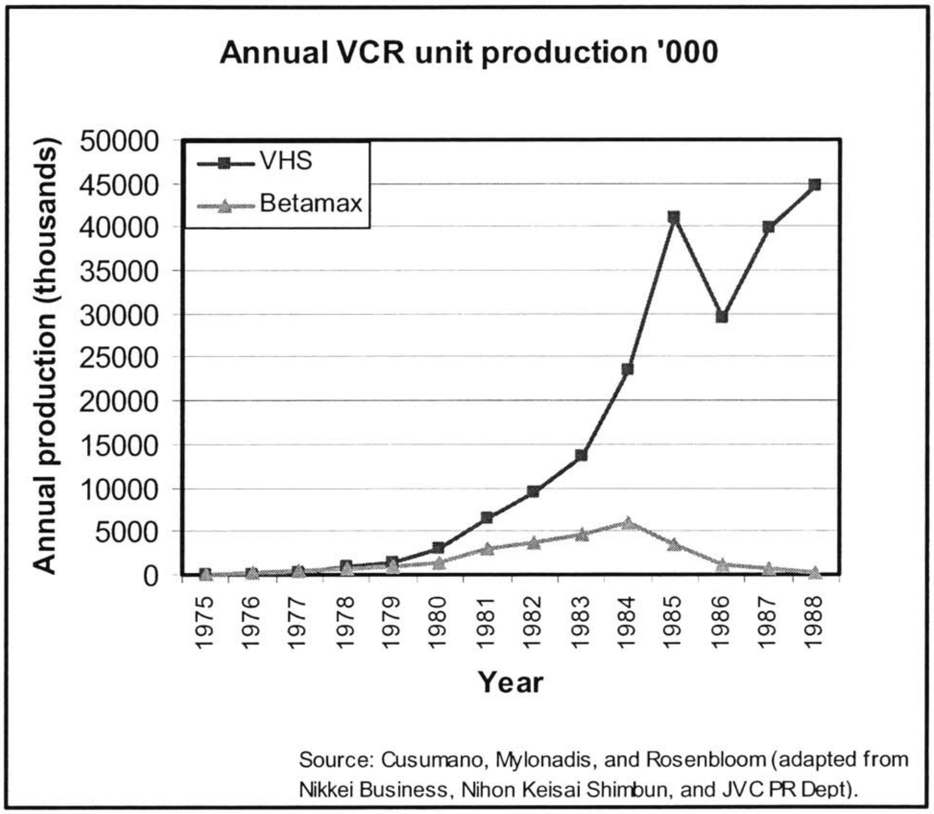


Figure 23: Comparison between VHS and Betamax annual sales.

The VCR case is one example of a single standard dominating the market; but many other industries are filled with such examples. Figure 24 shows somewhat similar dynamics for the SMIS-WideSky interaction. In this case however, the situation is a bit simpler because although EMC had higher aspirations for WideSky, it had a maximum of *one* supporting vendor at any point in time, itself. Moreover, the data is not as complete as the VCR data because the timescale covers only 16 months of the life of SMI-S, while the VCR timescale covered 13 years. The main point is to illustrate how one standard, loses as it battles with another.

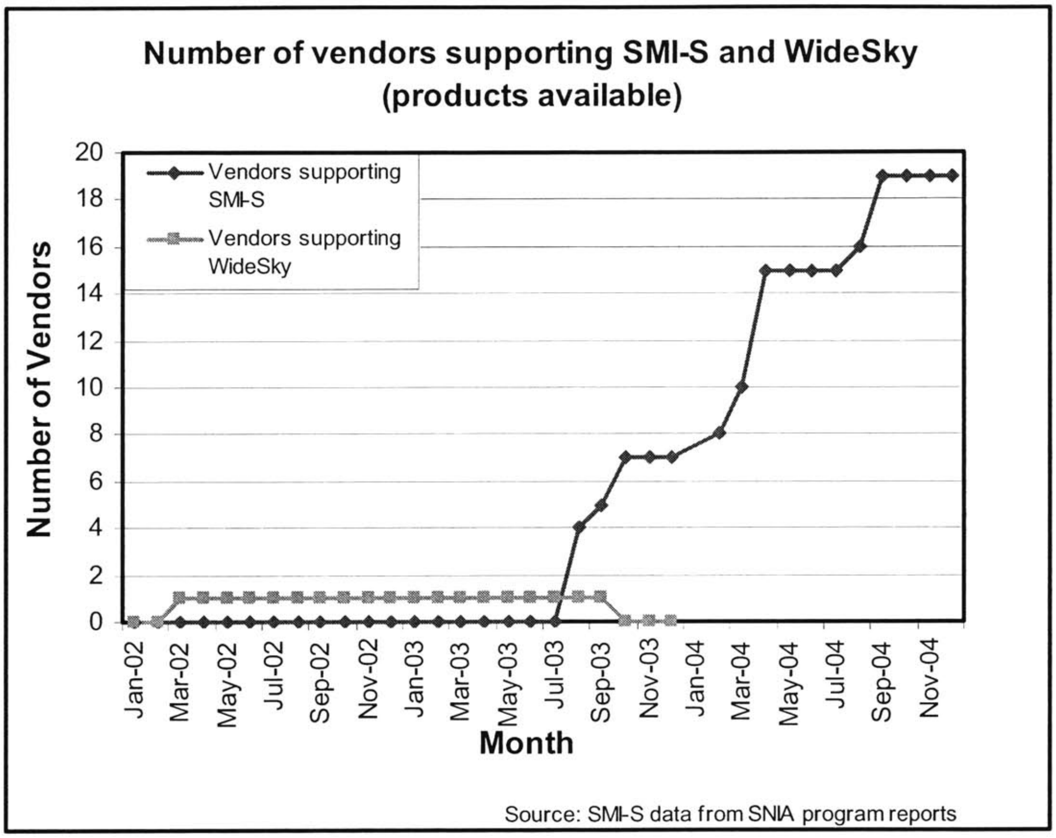


Figure 24: Number of vendors supporting WideSky and SMI-S (products available).

In summary, industries are generally dominated by a single standard which prevents dominant designs from being established based on other standards. This fact is a great incentive for firms in the data storage industry to actively back SMI-S. They have much interest in its success, because it will provide them with protection against a single player dominating the market based on a proprietary standard.

6.2 Self-Reinforcing Mechanisms Promoting the Growth of SMIS

In addition to the direct factors promoting the sustainability of SMI-S, there are also self-reinforcing mechanisms that have similar effects. These mechanisms are based on what is referred to in systems dynamics as the "success to the successful" archetype[42]. Their behavior follows one of the most basic reference modes in systems dynamics: the exponential growth[43]. Two such mechanisms can the identified: the *learning effects* mechanism and the *network externalities* mechanism[44]; both of which were extensively discussed by Schilling (2004).

6.2.1 Learning Effects

The learning effects mechanism can be described as follows. The more experience and knowledge are accumulated around a technology, the more effective it becomes. This increase in effectiveness generates value (revenues) which allows further investments in the technology. Moreover, the increase in effectiveness also increases the appeal of the technology and draws investments towards it. These investments in the technology in turn increase the experience and knowledge. This mechanism applies in a very straightforward manner to the SMI-S. As storage firm develop towards the specification, the cumulative knowledge and experience of their organizations increases. This allows them to enhance the specification, which then becomes more attractive, and entices them to allocate even more resources to it.

Another aspect of the learning effects is the *absorptive capacity* of a firm[45]. Cohen and Lenvinthal (1990) define it as *"the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends"*. The absorptive capacity of a firm is highly dependant on the firm's prior knowledge in the field. For example, the more a storage vendor gains knowledge and experience in SMI-S implementations, the faster they are able to advance their SMIS knowledge.

Figure 25 is a representation of the dynamics. The top two positive feedback loops correspond to the learning effects mechanism, one through experience (left) and the other through active investment in learning (right). The bottom positive loop is the absorptive capacity loop. All three loops exemplify the "success to the successful" archetype and have a cumulative effect of increased SMI-S adoption.

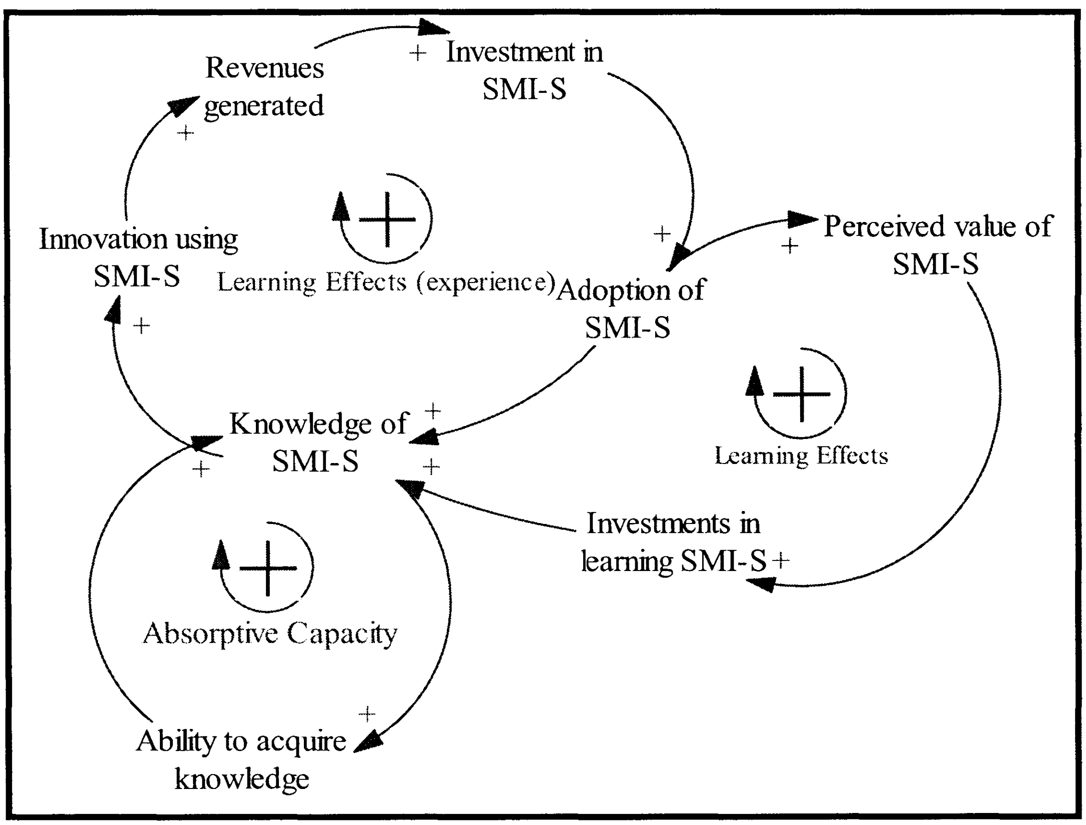


Figure 25: Learning effects and absorptive capacity dynamics

6.2.2 Network Externality Effects

Network externalities are another set of self-reinforcing mechanisms that positively influence the evolution of standards. Their fundamental behavior is also an exemplification of the "success to the successful" archetype.

Explaining Network Externalities

The main premise of network externalities is that the benefit of a technology increases with the increase in the number of users of this technology. The function that relates this benefit to the number of users is known as the *'network benefit function'*. Various studies have been made on the effects of network externalities on the adoption of technology[46]. There is also a multitude of historical examples of these dynamics. The predominance of these has been in industries characterized by the existence of physical networks, such as railroads and the telephone[47]. There are also additional studies on some famous industries, such as the VCR industry with the success of VHS over Betamax, the PC industry with the success of the IBM-compatible platform over Apple, and the OS industry with the success of Windows. They all drew on network externality dynamics during their growth stages.

In addition to promoting the growth of a standard, network externalities can also have the opposite effect; that of preventing the adoption of a standard. In fact, this is expected in any industry where more than one standard or technologies are in direct competition; the benefit of some comes at the expense of the others. Postrel (1990) presents the case of the quadraphonic failure at replacing stereo circa 1976 as an example of network externalities as barrier to entry[48].

It is important to distinguish between two types of network externalities, '*installed base*' and '*complementary goods*', which are sometimes also referred to as direct network externalities and indirect network externalities respectively[49]. The following explains the difference between the two, and how each applies to influencing the adoption of SMI-S.

Installed base or direct network externalities

This is the situation where the increase in benefits from a technology or a standard is dependent on the increase in availability of implementations of this technology or standard. In the data storage industry, for the standardization case, this amounts to the increase in availability of vendor products that are SMI-S compliant. The literature emphasizes the terminology of 'availability of the same good' and of 'users of the same good'. This is also true in this case, but the definition of 'same good' that I use is broader, because the underlying good is actually a group of products, as opposed to a single self-contained product. So, for our purposes, the availability of the same good refers to the availability of products, albeit heterogeneous, that support the SMI-S interface. These can be storage arrays, switches, volume managers, and pretty much any component of a SAN. The increase in availability of these SMI-S products provides customers with more choices of appliances and software that can be managed through the same interface, SMIS. This yields direct benefits to the customers, who are now able to avoid vendor lock-in while minimizing their expenses; and thus increases the overall value of the technology.

Complementary goods or indirect network externalities

In addition to the benefits reaped from the increase in SMI-S compliant products, there are benefits related to the availability and increase of complementary goods to the products. Examples of such complementary goods are SMI-S client management applications, or third party niche products for specific applications, that utilize SMI-S. The more of these complementary products are available on the market, the more benefits customers will be able to extract from the specification. In turn, this increases the value of the specification and its underlying technology, both to the customers, who will be able to tap into new functionality via the complementary good, as well as to the vendors, who will get additional usage of their products along with the associated revenues. These dynamics, along with the previously discussed dynamics of direct network effects, are represented in the causal loops diagram of Figure 26.

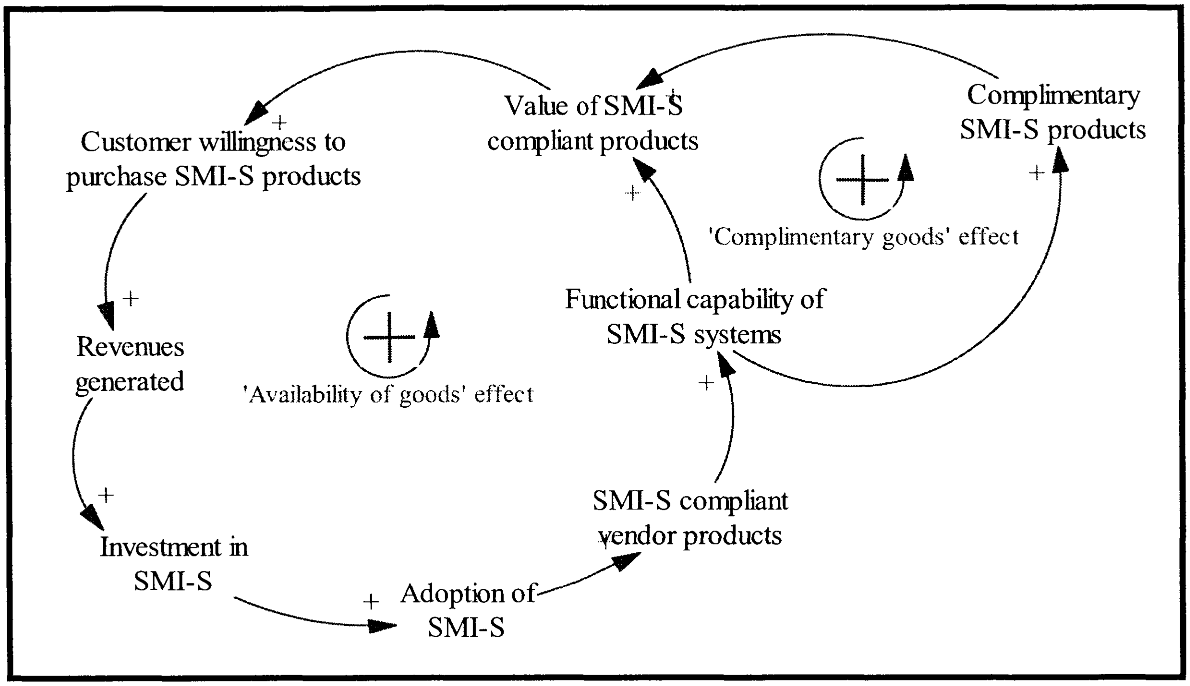


Figure 26: Network effects on the adoption of SMI-S

6.3 Factors Dampening the Growth of SMI-S

In addition to the previously listed major factors that positively contribute to the adoption of SMI-S, there is an equally important set of factors but with exactly the opposite effect. They act in a capacity that undermines and slows down the adoption of SMI-S. These are what I refer to as the dampening factors. The following is a list of the most significant of these factors along with an explanation for each.

6.3.1 Missing Core Functionality

Perhaps what has the greatest impact on slowing down the adoption of SMI-S is the fact that the specification still does not support all the functionality that proprietary interfaces and native APIs support. Moreover, in some instances, this missing functionality is actually core functionality of the products and essential to end users. For instance, let us take storage arrays as an example. Each vendor has implemented and supports a multitude of copy mechanisms for the replication of data from one storage location to another. The SMI-S 1.0.2 specification however states the following:

> *"While copy services functionality is broad in scope and represents vital functionality in any enterprise storage environment and the time of this specification publication, copy services design has principally only been validated for use in support of volume snapshots."*[32]

This is a clear indication that the specification as it stands at version 1.0.2 still lags behind the actual functionality that the vendors support. While this lag is not uncommon when a new specification is introduced to an established market; one must acknowledge that it has considerable delay effects on the adoption of the specification. This is especially true for client developers who need this functionality because of imposed requirements by the current business rules where there products are being used. Looking ahead though, the effects of the lag will be diminished. As the specification matures, and new versions of it are released each year, the expectation is that the functionality *delta* between SMI-S and native APIs will shrink, and become less of a factor, especially given that the specification provides for the implementation of vendor extensions. These extensions can compensate for the functionality shortcomings of the specification, until this functionality is embedded therein.

6.3.2 Complexity of the Implementation

Technically, interfacing with a provider's SMI-S interface is not as straightforward as interfacing with a native API written in C for example. There is a steep learning curve associated with the standard. In addition, one needs to learn WBEM and CIM. By virtue of being a standardized interface, the provided functionality is generalized. This results in yet more difficulties when one tries to use the general to control the specifics. Figure 27 shows how the difficulty of a standards-based implementation changes as the number of integrated systems to be managed increases. The difficulty variable is the ratio of a standards-based client implementation over a native-API based client implementation. Note however that the graph does not have any units and is meant to give only a qualitative representation of the trend. With a few number of systems to be integrated, it is easier for the client developer to use the native interfaces for each. As this number increases, so does the difficulty of coding to each specific API, and therefore the relative difficulty of using a standard such as SMI-S decreases. There is point at which the relative difficulty is 1, which means that it is equally difficult (or easy) to use SMI-S than the use a native API interface. A resistance to the use of SMI-S might therefore come at the early stages of implementation, if the client has a small number of systems to integrate and does not see the scalability as an immediate issue. The ease of interfacing with "straightforward" native APIs might be more alluring than embedding the flexibility in the clients by adopting standards at an early stage, for a long term ROI.

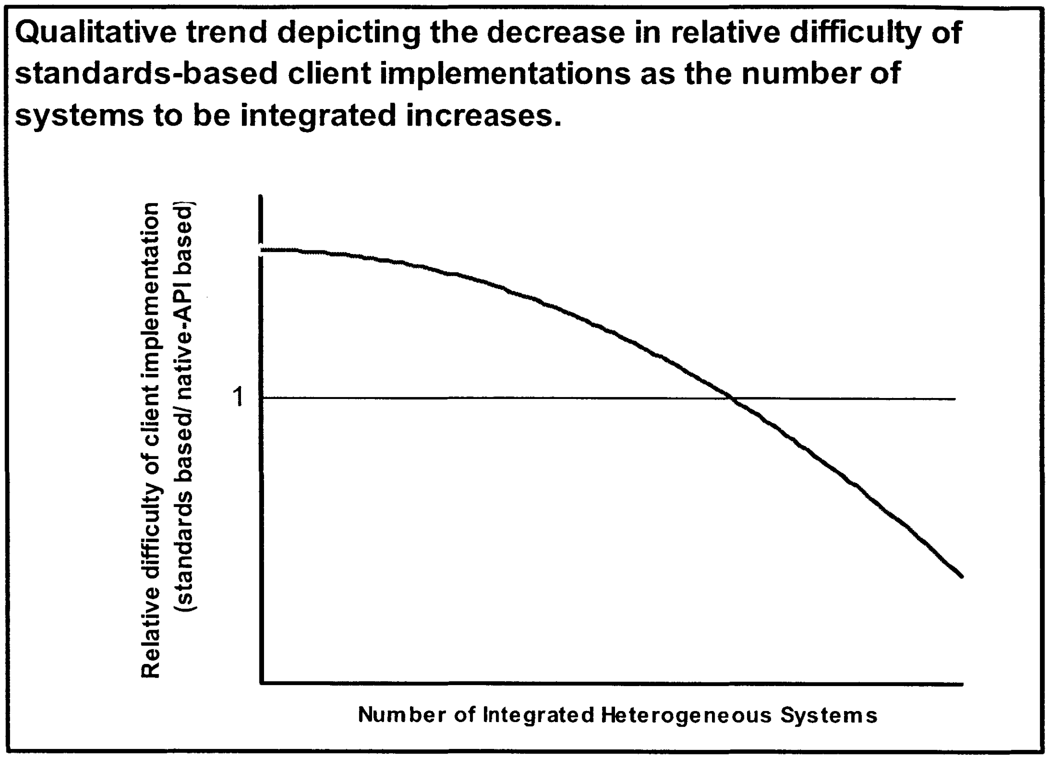


Figure 27: Standards-based client implementation difficulty vs. number of systems

6.4 Determining the Tipping Point and the Tipping Direction

> *...The company announced yesterday its plan to drop support of SMI-S from its future products. In an interview that followed, the CTO cited low adoption rates and various technical concerns as the reason behind this decision. "Technically, SMI-S did not deliver on its promises", he said. "We were not able to generate enough revenues from the technology to justify maintaining it. In simple terms, our customers did not want it."*

The above statement is fictitious and is only presented to emphasize the point that such a scenario could happen; in which case it would be one of many announcements just like it made by various firms throughout the storage industry. On the other hand, the exact opposite could take place; a continuing stream of announcements of vendor product support, new client availability, and success stories of IT customer implementations.

Has the SMI-S standards market tipped yet? When will it tip? In which direction will it tip? These are the questions that will be addressed in this section, with the aim of providing an insightful analysis - not a prediction - as to the expected progress and final outcome of SMI-S.

6.4.1 Explaining the Concept of 'Tipping'

Perhaps the best explanation for what it means for a market to tip can be given using Figure 28. Just as a quick disclaimer, the following explanation that does not claim to cover all possible trends of market behavior, but is rather simplified to emphasize the two most prevalent trends. First, we start with the assumption that the market has experienced some sort of growth early on. Without this assumption and this initial trend, the analysis would not be interesting. After this initial period of slow growth comes the tipping point. From there on, the market will experience one of the two most common evolution patterns. The first is that of accelerated growth, following the S-curve trend that has been extensively discussed in literature. The second is that of decline, in which case the market is unable to 'cross the chasm'[50]. This theoretical and qualitative description of behavior can be related back to the example of the VCR that was discussed in a previous section. The curves of VHS and Betamax of Figure 23 closely resemble the growth and decline curves of Figure 28.

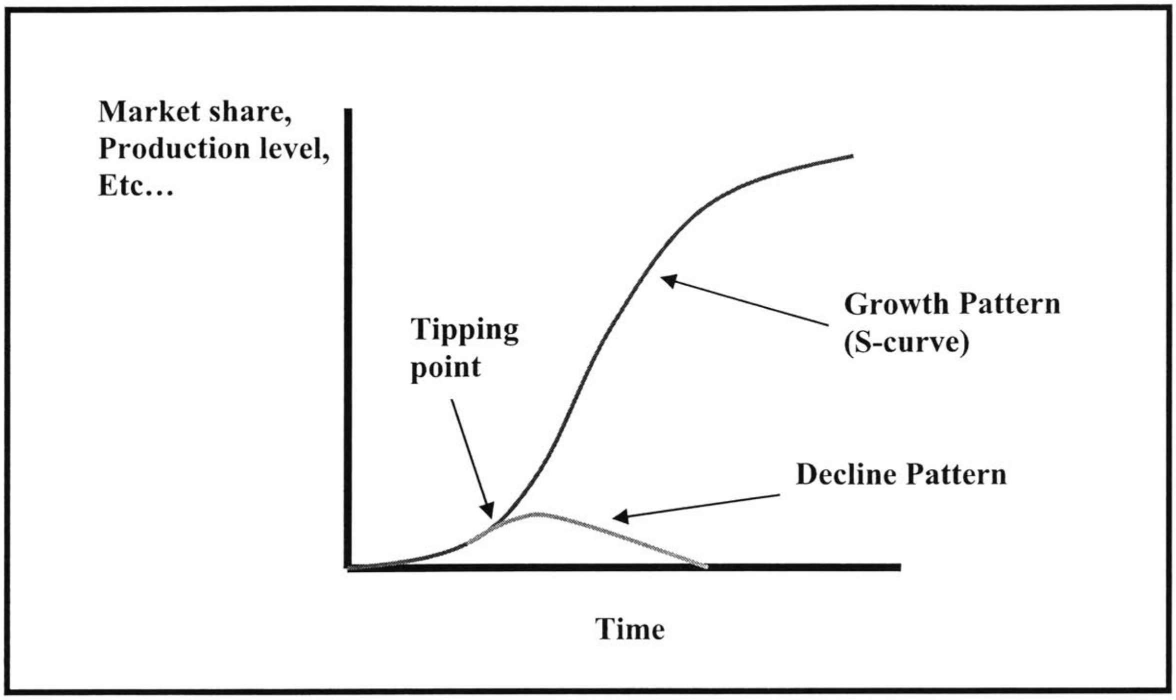


Figure 28: Representation of the tipping behavior of a market.

6.4.2 Adoption Trends of SMI-S Providers and Clients

The factors of interest in this analysis are those that will contribute to the temporal location of the tipping point and to the direction of the tipping. I will start by presenting the value chain around SMI-S (Figure 29). The success of the specification at becoming the standard and generating value to the IT industry is dependant on the first two links of the chain. A critical mass of SMI-S provider products and a critical mass of SMI-S client products need to be reached. I therefore propose to proceed with the analysis by examining two separate growth trends, and their interaction. The first trend is that of SMI-S provider products, and the second one is that of SMI-S client products.

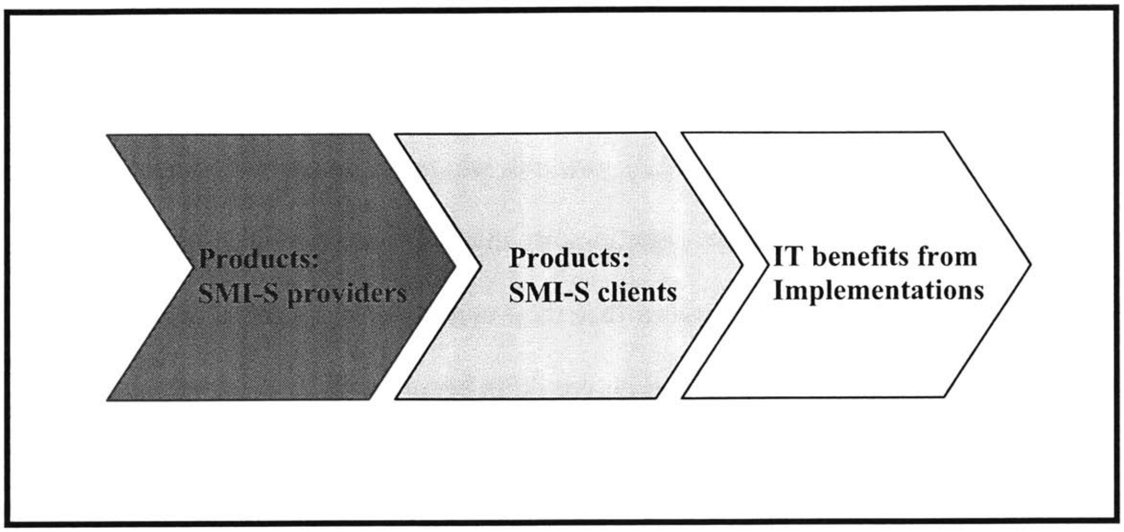


Figure 29: Simplified SMI-S value chain

SMI-S Providers

The current situation is that most major hardware vendors in the data storage industry have already released SMI-S provider products. Most of the firms discussed in the Data Storage Business chapter are therefore also present in Table 7, which is a list of firms with SMI-S provider products.

HP	Veritas
IBM	Engenio
EMC	3Par
Hitachi	Cisco
Sun	NetApp
Emulex	Qlogic
Brocade	StorageTek
McData	

Table 7: List of firms with SMI-S provider products as of December 2004 (source: www.snia.org)

A question therefore arises: has a critical mass been reached? From a number of providers point of view and a number of products point of view, the answer is 'yes'. Most major firms have most of their hardware products CTP certified (Figure 30 and Table 8). However, from a functionality point of view, the answer is 'no', and this is how looking at product count and company support indices alone could be misleading. They are certainly necessary for the adoption of the standard, but not sufficient. The current version of the specification (1.0.2), towards which all the above listed products are certified, is still missing some key functionality that is essential to end users, who are able to access it via the native interfaces that they currently use. The effect of lag in functionality support between SMI-S and native APIs becomes more apparent when one moves up the value chain to examine trends of SMI-S client support.

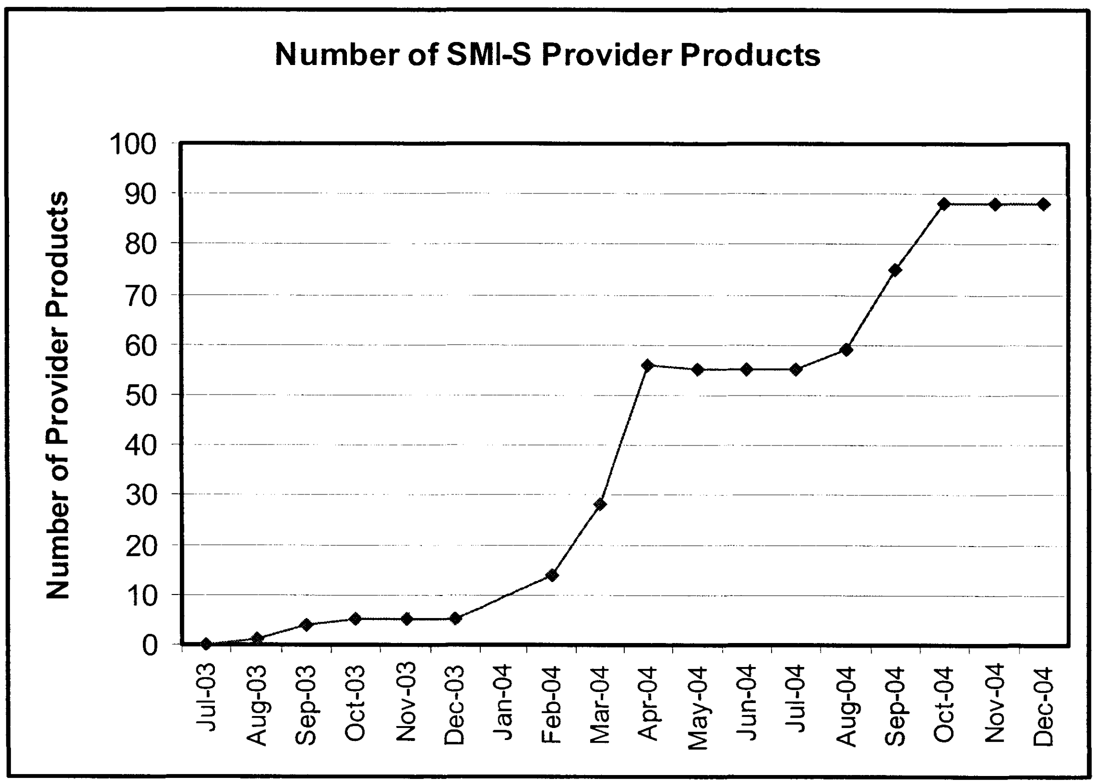


Figure 30: Number of SMI-S provider products over time (source: www.snia.org)

HP StorageWorks Enterprise Virtual Array	HP VA7400
HP StorageWorks Disk Array (XP)	HP StorageWorks Modular Disk Array (MSA)
HP StorageWorks Core Switch 2/64 PowerPak	HP StorageWorks Core Switch 2/64
HP StorageWorks SAN Switch 2/16 EL	HP StorageWorks SAN Switch 2/16
HP Sturestore FC 1GB/2GB Switch 16B	HP StorageWorks SAN Switch 2/16 PowerPak
Engenio Information Technologies E-Series	HP Brocade Silkworm 2800 FC Switch
EMC CLARiiON Storage Arrays	EMC Symmetrix Storage Arrays
Hitachi Data Systems Lightning 9970V	Hitachi Data Systems Lightning 9980V
Hitachi Data Systems Lightning 9960	Hitachi Data Systems Lightning 9910
Hitachi Data Systems Thunder 9580V	Hitachi Data Systems Thunder 9570V
Hitachi Data Systems Thunder 9533V	Hitachi Data Systems Thunder 9532V
Hitachi Data Systems Thunder 9531V	Hitachi Data Systems Thunder 9530V
Hitachi Data Systems Thunder 9200	Hitachi Data Systems Thunder 9585V
Hitachi Data Systems TagmaStore USP	Hitachi Limited SANRISE USP
Hitachi Limited SANRISE 9970V	Hitachi Limited SANRISE 9980V
Hitachi Limited SANRISE 2200	Hitachi Limited SANRISE 2800
Hitachi Limited SANRISE 9585V	Hitachi Limited SANRISE 9580V
Hitachi Limited SANRISE 9570V	Hitachi Limited SANRISE 9534V
Hitachi Limited SANRISE 9533V	Hitachi Limited SANRISE 9532V
Hitachi Limited SANRISE 9531V	Hitachi Limited SANRISE 9530V
Hitachi Limited SANRISE 1200	Hitachi Limited SANRISE 1100
Sun Storedge 3000	Sun Storedge 6000
Sun Storedge 6130	Sun Storedge 6920
McDATA Intrepid 6140	IBM Enterprise Storage Server
McDATA ED-5000	McDATA Intrepid 6064
McDATA Sphereon 4300	McDATA Sphereon 4500
McDATA ES-3232, ES 3216	McDATA ES-3032, ES 3016
Dell | EMC CLARiiON Storage Arrays	CNT FC/9000 Fibre Channel Directors
Cisco MDS 9509 Multilayer Director Switch	Cisco MDS 9506 Multilayer Director Switch
Cisco MDS 9216 Multilayer Fabric Switch	Cisco MDS 9140 Multilayer Fabric Switch
Cisco MDS 9120 Multilayer Fabric Switch	Brocade Silkworm 2010
Brocade Silkworm 2040	Brocade Silkworm 2050
Brocade Silkworm 2210	Brocade Silkworm 2240
Brocade Silkworm 2250	Brocade Silkworm 2400
Brocade Silkworm 2800	Brocade Silkworm 3200
Brocade Silkworm 3250	Brocade Silkworm 3800
Brocade Silkworm 3850	Brocade Silkworm 3900
Brocade Silkworm 12000	Brocade Silkworm 24000
Network Appliance Enterprise Storage System	SGI Infinite Storage
StorageTek D-Series	StorageTek B-Series
IBM Total Storage FAStT Storage Server	Emulex LightPulse FC HBA Family
QLogic SANbox 5200	HP VA7110
HP VA7100	HP VA7410

Table 8: List of SMI-S provider products, December 2004 (source: www.snia.org)

SMI-S Clients

On the client side of the fence, the uptake from a numbers point of view is as show in Table 9. The top hardware manufacturers have already embedded client support in their management products. In addition, there are some new comers such as CreekPath Systems for example, who specialize in management software, and provide a full suite of such software to control and manage third party products.

Although the client applications support SMI-S, they do not exclusively use this specification when interfacing with managed objects. As it was previously stated, the specification is still lagging behind the full functionality of vendor products. Client developers therefore utilize SMI-S where it is supported. In addition to that, and in order to provide their clients with control over the full functionality of the SAN components, they supplement the SMI-S interface with an interface to the native APIs of the managed systems. On their website, CreekPath Systems state the following:

> *"...CreekPath is compliant with the existing Storage Management Initiative (SMI) specification. CreekPath supplements the SMI interface with native APIs to expose the unique features and capabilities of the installed storage elements."*[51]

This means that although utilizing SMI-S, client developers still need to invest in learning and interfacing with native APIs in order to fully support their customers' needs. What value then, does SMI-S being to client products? An inherent but indirect value is that of flexibility and learning, as presented earlier. On the other hand, in order to estimate the direct value, although this goes beyond the scope of this study, one would have to examine the internals of the products to understand what percentage of functionality is accessed via SMI-S and what percentage is accessed via native calls. The aim would be to evaluate the direct benefits that client developers are collecting from their support for the specification.

AppIQ StorageAuthority Suite	IBM SAN Integration Server
Computer Associates BrightStor SAN Manager	IBM SAN Volume Controller
CreekPath Suite 3.1	IBM Tivoli Storage Resource Manager
CreekPath Suite 3.2	SGI InfiniteStorage Resource Manager
EMC ControlCenter Storage Management Software	SGI InfiniteStorage Shared Filesystem CXFS
EMC VisualSAN Storage Management Software	SGI InfiniteStorage Volume Manager XVM
Hitachi Data Systems HiCommand Device Manager	Sun StorEdge Enterprise Storage Manager
Hitachi Data Systems HiCommand Storage Manager	Veritas CommandCentral Storage
HP OpenView Storage Area Manager	Veritas Volume Manager

Table 9: List of SMI-S client products (source: www.snia.org)

6.4.3 When and How Will Tipping Occur?

I can think of no clear argument that would oppose standardization in the data storage industry. Simplification of complex networked storage systems is a prerequisite for continued growth. There is clearly an unmet market need for this simplification, and SMI-S has all the attributes that would make it a promising potential solution for this unmet need. To summarize:

- It is a solid technical specification that solves the interoperability problem
- It is an open standard and it is based on other standards
- It is developed by an open industry body (SNIA) that also manages its PR
- It has the backing and support of all the industry leaders
- It is continuously growing in functionality support
- Most products from the major vendors have embedded provider support for it
- Many client products with SMI-S compliance are available

What is still missing is that the value chain is still not functioning properly: value is still not being transmitted up the chain. IT customer benefits from the providers' SMI-S implementations and subsequent client support for SMI-S, have been minimal. And although there are a lot of success stories of support and interoperability, there are still very few success stories of end-user added value. This means that providers and clients alike are still not reaping the benefits of their investments. The question therefore becomes: Is this a normal trend of new standards? Yes. It takes time for the value chain to be completed and for the players at various stages of this value chain to extract benefits from it. This is certainly true for many other industries. The performance rate starts out being very slow and gradually increases over time, following an S-curve pattern[52]. At the early stages, where the performance of the new technology or standard are still very low compared to the established technology, benefits are still not being extracted from the value chain. *Time* is therefore the key catalyst. If the current trends in SMI-S performance growth, technology development, vendor support and adoptions rates, and marketing hype continue; the specification will become the standard of the industry. This will mark the upward tipping of the SMI-S market. The risk however is that time delays for the performance to reach an acceptable level become too great. This will lead firms to start 'cutting' their losses and dropping support for SMI-S, which will mark the beginning of a decline phase and the downward tipping of the SMI-S market.

This 'performance over time' issue that SMI-S is facing in trying to become an adopted technology standard in an established market is not unique. In fact, such problems have been extensively studied in a *disruptive technologies* context. In the next section, I will take a similar approach and analyze SMI-S as a disruptive technology. The goal is to determine whether or not it is a disruptive technology; and it if is one, determine the implications that this will have on its expected course of evolution.

6.5 SMI-S as a Disruptive Technology

Before addressing SMI-S in particular, I will start with a brief overview of *disruptive technologies* and of the studies done thus far in this area. This will be followed by the actual analysis of SMI-S within that framework.

6.5.1 Brief Overview of Disruptive Technologies

The term 'disruptive technology', in its current contextual meaning, was first used by Christensen (1997) in his famous book *The Innovator's Dilemma.* One of the main points that he presents is how a disruptive technology enters an established market and grows until it is able to attack the established products of this market. He uses mainly examples from the disk drive industry to illustrate the concept. Furthermore, Christensen proposes that a disruptive technology must meet three criteria. The first criterion is lower cost than the established technology. The second one is lower traditional performance, as measured by the established technology performance metrics. And the third one is higher ancillary performance: performance dimensions that are beyond the scope of the established technology. Later work by Utterback and Acee (2003)[53] presented an expansion of Christensen's notion. From the criteria point of view, they made the argument that a disruptive technology does not have to meet the criteria set by Christensen for the three metrics of cost, traditional performance, and ancillary performance. To that effect, they presented a tabular list of examples of disruptive technologies that cover all eight (2^3) possible combinations of the metrics. In addition, Utterback and Acee further advanced the notion that a disruptive technology's impact is not limited to the displacement of existing technologies in a particular market. Rather, they propose that it will also contribute to the overall expansion of the market that it enters.

6.5.2 SMI-S through the Disruptive Technology Lens

Although Utterback and Acee have already proved that Christensen's three criteria do not have to be met all the time, and that there are disruptive technologies that meet any of the eight metric combinations; it is nonetheless important to understand where SMI-S falls with respect to these metrics.

Traditional Performance

Presently, SMI-S technology offers lower traditional performance than non-SMIS technology. Products based on SMI-S therefore cannot meet the same performance criteria that established products meet. In this case, since the technology under examination is interface management technology, then we look at performance as being a measure for the management functionality that is enabled by an interface. When discussing SMI-S, we have addressed the fact that the current specification does not cover all the functionality that is allowed though native interfaces. The example in that case was that of copy services and data replication. To that regard, the new specification is clearly lagging, and if traditional performance and access to functionality was the only measure by which an interface technology gets chosen, SMI-S would be at a clear disadvantage. This however, is not exactly the case, and we turn to other two metrics.

Ancillary Performance

SMI-S based solutions add a functionality dimension that is not available from traditional solutions. This new dimension is that of allowing the control and management of heterogeneous systems through a common interface. To that regard, this functionality *is* the ancillary performance that the specification brings; and it is clearly higher than the one provided by traditional technologies.

Traditional Cost

The traditional cost in this case is the cost of development of a solution using a particular technology. For clarification, it is the cost of developing either a single vendor-provider interface, or a client interface to manage a single vendor product. This cost is clearly higher when it comes to using SMI-S because of three main reasons:

- The technology itself is more complex than a native API, especially when developing towards a single implementation.
- There is a steep learning curve, and investments must be made into this learning before the technology can be implemented.
- Most firms do not have the SMI-S absorptive capacity or a development infrastructure that is similar to the one they have for their native API development

These reasons force the firms to make large investments in order to use the technology, and when these investments are attributed to the cost of SMI-S based product development, this cost becomes higher than that of development based on traditional technologies.

Trends of Cost and Performance

SMI-S implementations therefore have higher cost, lower traditional performance, and higher ancillary performance that non-standards based interfaces. From a cost point of view, the expectation is that it has started falling and will continue to fall. After the initial investments in learning and infrastructure development have been made, the cost of development per product should start dropping. Moreover, the more encompassing the development is, as far as number of heterogeneous products to be managed, the lower the cost will get.

As far as traditional performance is concerned, the expectation is that similar trends will take place. The SMI-S traditional performance will continue to increase as more and more functionality support is added with each new version of the specification. At the same time however, the actual functionality of the underlying elements to be managed is also increasing. The question therefore becomes, which rate of functionality increase is higher, that of SMI-S or that of the actual systems and their native interfaces? Empirical analysis is needed to answer this question with absolute certainty. In the absence of that, I offer the following insight with respect to the trends. As firms develop new products, they go through the process of developing the underlying technology that provides the functionality, as well as the interfaces to access this functionality. On the other hand, development of SMI-S is mainly an interfaces development. Granted that it carries with it a complexity because of conformance and generalization requirements, I would argue that advances in it can be done much faster than advances that require fundamental technological innovation in the underlying products. Qualitatively, the traditional performance curves would be similar to the ones in the following graph.

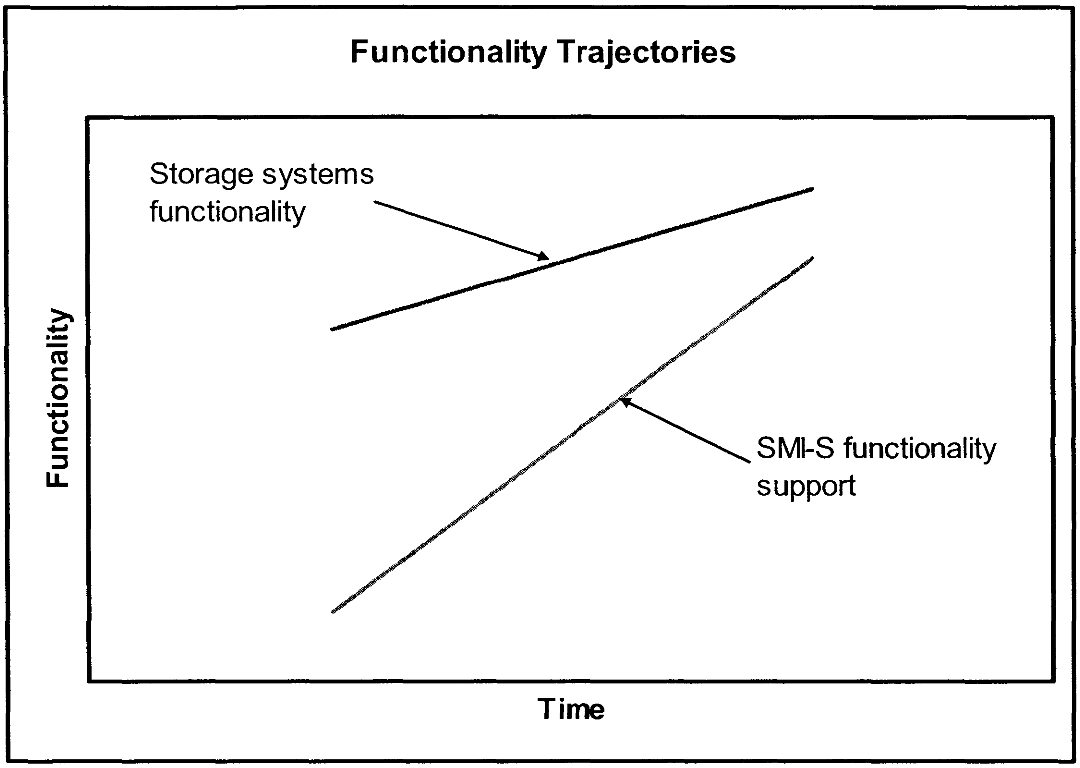


Figure 31: SMI-S and storage systems functionality trajectories

SMI-S interfaces could eventually reach the same level of functionality support that native interfaces provide. More optimistically, it could be that the market demand curve is actually lower than the storage systems functionality curve (oversupply in functionality). It could be argued that this is the case based on the fact that the statements: 'need improved management tools' and 'better interoperability', rank first and second with very high percentages, on the customer needs list of Figure 16. The functionality curve of SMI-S interfaces might therefore intersect with the market needs curve much earlier than it would with the system functionality curve, and this would make the capabilities of the SMI-S technology sufficient from a traditional performance (functionality) point of view. In addition, SMI-S interfaces would also still be providing the ancillary functionality of interoperability and commonality of management interfaces. Such dynamics would lead to the establishing of SMI-S as the industry standard. In essence, SMI-S technology becomes the dominant design for data storage management.

7 Conclusion

This thesis has explored the topic of management software standardization in the data storage industry. In particular, it has analyzed the dynamics affecting the sustainability of these standards in an aim at providing insight into their future evolution towards either widespread adoption or obsolescence. The approach was methodical. It started by establishing the context for the discussion through a technology and business overview of the industry. Next, the evolution of proprietary standards efforts was presented, along with a root cause analysis of its demise and of the subsequent rise of the SMI-S open standards movement. Finally, the last portion of the thesis focused on understanding the dynamics influencing the tipping point of the SMI-S technology adoption, and on predicting in which direction its market will tip.

The following is an executive summary of the effort presented in a point by point bulleted format.

7.1 Executive Summary

7.1.1 The Data Storage Technology and Business

- Storage solutions are complex systems composed of various heterogeneous products, the major ones being: servers, storage arrays, switches, directors, HBAs, and tape libraries.
- The worldwide data capacity storage requirements are continuously increasing. In addition to the traditional demands of capacity and performance, this

proliferation of data is increasing the requirements for complex functionality to manage this data.

- Storage solutions are increasing in complexity, and the relatively simple DAS systems are being replaced by more complex networked solutions such as NAS and SAN.
- Improved management tools and better interoperability top the list of IT customer demands.
- Hardware is becoming commoditized, and traditional hardware firms are looking at software for new sources of revenues. The battle for dominance is shifting to the software arena.

7.1.2 Proprietary Standards

- The complexity and diversity of the components of a storage system (NAS or SAN) has generated a need for standardization in order to allow the management of heterogeneous products, supplied by multiple vendors, in a simple and cost effective manner.
- In 2001 EMC announced WideSky, a proprietary middleware that would be able to control third party APIs and provide the client with a common interface to manage the hardware of third party vendors. The aim was to make WideSky the de facto standard of storage management.
- In March 2002, EMC announced the availability of WideSky. This was met with apprehension from the rest of the industry players.

- WideSky failed at becoming the dominant design due to the lack of industry support. It was finally abandoned in 2003, and EMC pledged its backing for SMI-S.

7.1.3 SMI-S

- The Storage Networking Industry Association is comprised of most of the industry players, including all the major ones. SNIA is the body that develops the Storage Management Initiative Specification (SMI-S)
- SMI-S is built on top of other open standards and specifications such as HTTP, XML, WBEM, and CIM. It is an interface specification for the management of storage systems.
- SMI-S adoption is on the uptake. All major vendors have released SMI-S compliant products, and there are currently around 90 announced provider product with the SMIS CTP certification.
- On the other hand, client applications and end user implementations have not seen similar growth because the functionality of the first version of the specification is still limited in comparison with the functionality that vendor specific native APIs provide.

7.1.4 Sustainability of SMI-S

- Fear of hardware lock-in because of lack of interoperability between vendors is prompting customers to demand management applications that are not vendor specific. This request can only be accomplished through standardization.

- Firms are afraid of technological lock-out in case SMI-S becomes the industry standard before there are ready to release SMI-S compliant products. This is pushing them to invest early in supporting the technology even before it becomes the adopted standard.
- Some firms view the success of SMI-S as a medium with which to prevent the dominance that a single player could exercise through proprietary standards.
- Network externalities (the benefit from a technology increases with the number of users) and learning effects (the more knowledge is accumulated around a technology the more effective it becomes) are two positive feedback mechanism contributing to the growth of SMI-S.
- For a single vendor support, learning SMI-S is more complex than learning a native API. Moreover, the current version of the specification (1.0.2) is missing come core functionality that is available in native APIs. These two factors clearly have a dampening effect on the adoption rate of the specification.
- Although the adoption of the specification is on the uptake, the value chain has not been completed yet, and direct benefits have not been reaped from SMI-S implementations. Time is of the essence. The 'tipping' (positive) of the technology will depend on whether the value chain gets fully established and benefits are generated along all its links, before firms start abandoning support for SMI-S.
- SMI-S is a disruptive technology competing with an established storage management interface paradigm that is based on native APIs. It has higher initial

cost, lower traditional performance (functionality), but higher ancillary performance (commonality of interfaces) than native API technologies.

- The performance gap between SMI-S and native APIs is closing. An SMI-S success at becoming the dominant design will not only displace the existing technology from the market; in addition, it will allow for the expansion of the total market.

7.2 Future Work

By using this thesis as a stepping stone, future research could focus on developing a predictive understanding of the effect of standards on the construct of the storage industry; along with generalizations that apply to other industries. The questions to be answered are the following:

- What value do standards generate, and what must firms do to capture this value?
- How do standards in the industry influence innovation and the architecture?
- What is the effect of standardization on the creation of entrepreneurial opportunities in the industry?
- What is the effect of standardization on the diffusion of knowledge?

In essence, such research will form the bottom cone of Figure 32; the top cone being what was achieved in this thesis.

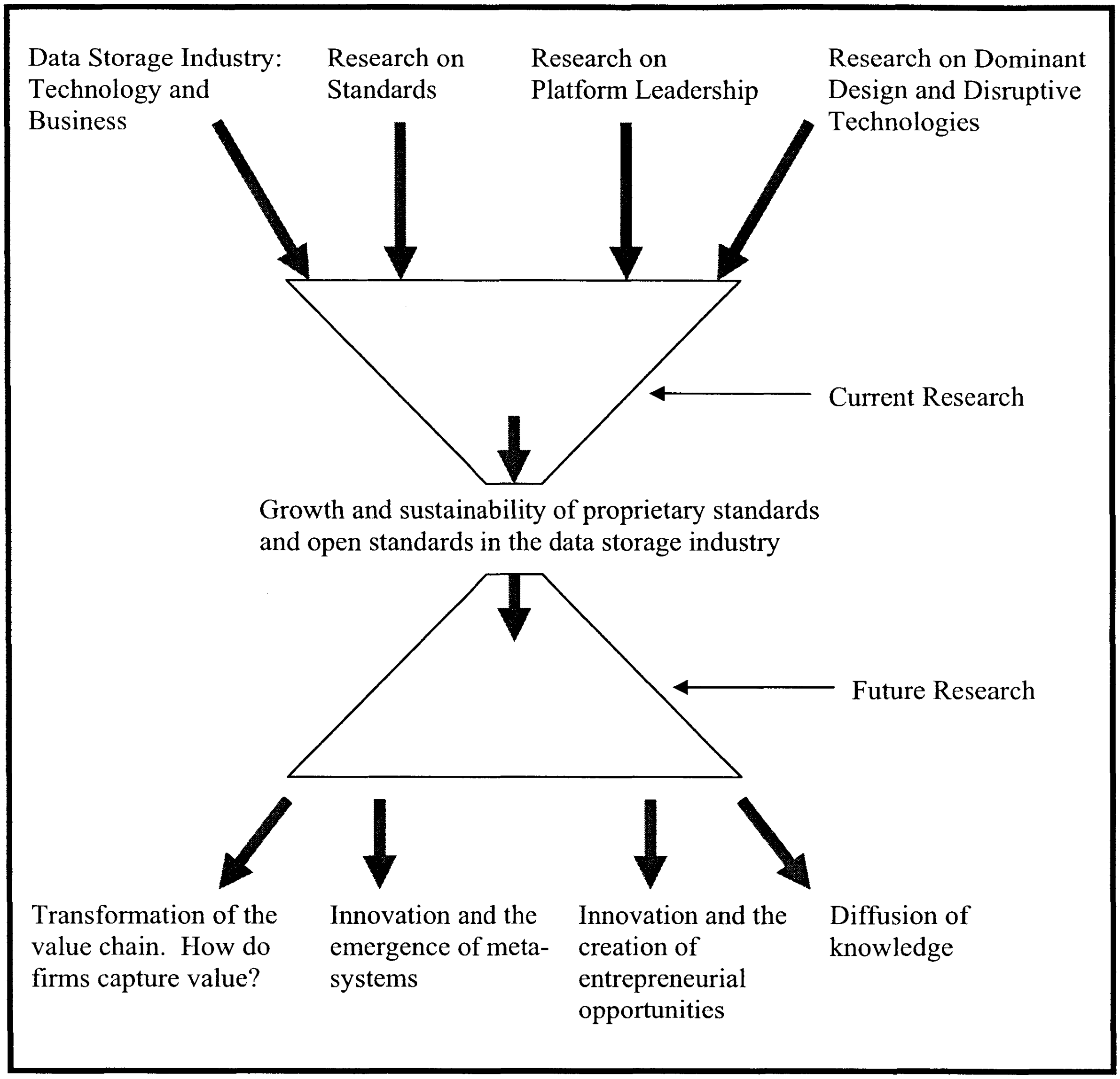


Figure 32: double-cone approach to understanding data storage software standards

Appendix 1: Glossary and Definitions

List of Used Acronyms

API:	Application Programming Interface
ATA:	Advanced Technology Attachment
C:	C programming language
CIM:	Common Information Model
CIMOM:	Common Information Model Object Manager
CLI:	Command Line Interface
CPU:	Central Processing Unit
CTP:	Conformance Testing Program
DAS:	Direct Attached Storage
DB:	Database
DMTF:	Distributed Management Task Force
DTD:	Document Type Definition
ECC:	EMC Control Center
ESCON:	IBM mainframe fiber connectivity
FC:	Fiber Channel
FICON:	Fiber Connectivity
FS:	File System
GUI:	Graphical User Interface
HBA:	Host Bus Adapter
HDD:	Hard Disk Drive
HSM:	Hierarchical Storage Management
HTML:	Hyper Text Markup Language
HTTP:	Hyper Text Transfer Protocol
HW:	Hardware
ILM:	Information Lifecycle Management
I/O:	Input Output

IP:	Internet Protocol
IT:	Information Technology
iSCSI:	Internet SCSI
JBOD:	Just a Bunch of Disks
LUN:	Logical Unit Number
NAS:	Network Attached Storage
OEM:	Original Equipment Manufacturer
OO:	Object Oriented
OS:	Operating System
PC:	Personal Computer
PDP:	Partner Development Process
PR:	Public Relations
RAID:	Redundant Array of Independent Disks
R&D:	Research and Development
ROI:	Return on Investment
SAN:	Storage Area Network
SATA:	Serial Advanced Technology Attachment
SATA II:	Serial Advanced Technology Attachment II
SCSI:	Small Computer System Interface
SMF:	Storage Management Forum
SMI:	Storage Management Initiative
SMI-S:	Storage Management Initiative Specification
SMIS:	Storage Management Initiative Specification
SNIA:	Storage Networking Industry Association
SNMP:	Simple Network Management Protocol
SRM:	Storage Resource Management
SW:	Software
TCO:	Total Cost of Ownership
TCP:	Transmission Control Protocol
UDP:	User Datagram Protocol

UML:	Unified Modeling Language
URL:	Uniform Resource Locator
VCR:	Video Cassette Recorder
VHS:	Video Home System
VM:	Volume Manager
WBEM:	Web-Based Enterprise Management
XML:	Extensible Markup Language

Data Capacity Definitions

	Acronym	How much is that?
Byte	B	8 bits
Kilobyte	KB	1,000 bytes
Megabyte	MB	1,000,000 bytes
Gigabyte	GB	1,000,000,000 bytes
Terabyte	TB	1,000,000,000,000 bytes
Petabyte	PB	1,000,000,000,000,000 bytes
Exabyte	EB	1,000,000,000,000,000,000 bytes

Table 10: Definition of storage capacity units (true values are multiples of 1024)

Appendix 2: References

1 "How Much Information? 2003", School of Management and Information Services, University of California at Berkley, California, 2003.

2 J. Band. *The Storage Outlook: Global Storage Solutions*, Reuters Business Insight, Technology and eCommerce, 2003.

3 "Disk Interface Technology", Intel Corporation, available from www.intel.com

4 "Implementing a FICON Solution for Mission-Critical SAN Environments", Brocade white paper GA-WP-552-02, available from www.brocade.com, 2004

5 R. Paquet, "Determine you Hierarchy of Storage Needs", Gartner research note, January 2005.

6 S. Zaffos, R. W. Cox, J. Krischer, and M. Chuba, "Magic Quadrant for Midrange Enterprise Disk Arrays, 1H04", Markets M-23-1481, Gartner, July 2004.

7 S. Zaffos, R. W. Cox, J. Krischer, and M. Chuba, "Magic Quadrant for High-End Enterprise Disk Arrays, 1H04", Markets M-23-1173, Gartner, July 2004.

8 "Market Share: Network-Attached Storage, Worldwide, 2Q04 (Executive Summary)", available from www.gartner.com, 2004.

9 M. L. Katz and C. Shapiro, "System Competition and Network Effects", Journal of Economic Perspectives, Volume 8, Number 2, pp 93-115, 1994

10 News release from www.emc.com, New York, NY, October 2001.

11 Komiega, Kevin, "Vendors Denounce EMC's WideSky as de facto Standard", www.searchstorage.com, February 2002.

12 News release from www.emc.com, Hopkinton, MA, March 2002.

13 Utterback, J. M. and Abernathy, W. " A Dynamic Model of Process and Product Innovation", Omega, 33, pp639-356, 1975

14 Suarez, F. F. and Utterback, J. M. "Dominant Designs and the Survival of Firms", Strategic Management Journal, Volume 16, Number 6, pp415-430, September 1995.

15 Cusumano, M., Mylonadis, Y., and Rosenbloom, R. "Strategic Maneuvering and Mass-Market Dynamics: The Triumph of VHS over Beta", Business History Review, Volume 66, 1992.

16 Bowes, J. and Elliott, S. "Tradition, Confusion and Multimedia; Developing New Standards from Old Industrial Roots", International Engineering Consortium, Annual Review of Communication, 47, pp548-557, 1993-1994.

17 Hemenway, D. *Industry-wide Voluntary Product Standards*, Ballinger, Cambridge, MA, 1975.

18 "The Storage Outlook: Global Storage Solutions", Reuters Business Insight, Technology and eCommerce, 2002.

19 Gawer, A. and Cusumano, M. *Platform Leadership, How Intel, Microsoft, and Cisco Drive Industry Innovation"*, Harvard Business School Press, Boston, Massachusetts, 2002.

20 Lelii, S. "One-on-One with Gary Bloom", available from VARBusiness ay www.varbusiness.com, July, 2002.

21 R. Henderson, "Competing is Standards Driven Markets", Technology Strategy lecture notes, MIT Sloan, 2004.

22 Sharead, Paul. "EMC Says Good-bye to WideSky, Hello to SMI", www.internetnews.com, September, 2003.

23 "The SNIA's Storage Management Initiative", www.snia.org/smi/home.

24 News release from www.emc.com, April 2004.

25 News release from www.ibm.com, April 2004.

26 News release from www.hds.com, April 2004.

27 News release from www.sun.com, April 2004.

28 News release from www.hp.com, April 2004.

29 "Common Information Model Specification', version 2.2, Published online by the DMTF at www.dmtg.org, June 1999.

30 "CIM/WBEM Technology Backgrounder", Storage Networking Industry Association, 2002.

31 Sandlund, K. "Network Management Using WBEM", Master's Thesis, Lulea Tekniska Universitet, 2001.

32 "SNIA Storage Management Initiative Specification, Version 1.0.2 (FINAL)", available from www.SNIA.org, February 2004.

33 J. Rigger, "SNIA-CTP Update for the SNIA End-User Council", available from www.snia.org, December 2004.

34 SMI Program Reports from July 2003 to December 2004, available from www.snia.org

35 Merriam-Webster Online, available at www.m-w.com

36 M. A. Schilling, "Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technology Success and Failure", Academy of Management Review, Vol 23 No 2 pp 267-284, 1998.

37 J. M. Utterback, Mastering the Dynamics of Innovation, Harvard Business School Press, 1994.

38 J. Saleh, D. Hastings, D. Newman. *Flexibility in system design an implications for aerospace systems*, Acta Astronautica, September 2002.

39 P. Banerjee, "Describing, Assessing and Embedding Flexibility in System Architectures with Application to Wireless Terrestrial Networks and Handset Processors", System Design and Management thesis, Massachusetts Institute of Technology, Cambridge, 2004.

40 R. de Neufville, J. Clark, F. Field. Real Options 2, Engineering Systems Analysis for Design, MIT, Cambridge, Massachusetts.

41 S. Park, "Quantitative Analysis of Network Externalities in Competing Technologies: The VCR Case", Department of Economics, SUNY at Stony Brook, New York, January 2003.

42 J. Hynes, Systems Dynamics lecture notes, MIT Sloan School of Management, 2004.

43 J. D. Sterman, Business Dynamics, Systems Thinking and Modeling for a Complex World, Irwin Mcgraw-Hill, 2000

44 M. A. Schilling, Strategic Management of Technological Innovation, McGraw-Hill Irwin, New York, 2004.

45 W. M. Cohen and D. A. Levinthal, "Absorptive Capacity: A New Perspective on Learning and Innivation", Administrative Science Quarterly, March 1990, pp 128-152.

46 M. L. Katz and C. Shapiro, "Technology Adoption in the Presence of Network Externalities", Journal of Political Economy, Vol 94, No 4, 1986.

47 A. Brian, "Competing Technologies and Lock-in by Historical Small Events: the Dynamics of Allocation under Small Returns", Policy paper, Stanford, California, Stanford University, Center of Economics and Policy Research, January 1985.

48 S. R. Postrel, "Competing Networks and Proprietary Standards: The case of Quadraphonic Sound", The Journal of Industrial Economics, Volume XXXIX, No 2, December 1990

49 S. Park, "Network Benefit Function in the Presence of Indirect Network Externalities", Department of Economics, SUNY at Stony Brook, New York, January 2002.

50 G. A. Moore, Crossing the Chasm, HarperBusiness, New York, 1991

51 Excerpt from the Product Detail section, entitled "CreekPath Suite – Standards Based", available from www.creekpath.com, 2004.

52 C. M. Christensen, The Innovator's Dilemma: How new Technologies Cause Great Firms to Fail, Harvard Business School Press, 1997.

53 J. M. Utterback and H. Acee, "Disruptive Technology", Presented at the Pravitt Conference, Sussex, England, 2003