

1.0 Personal Information Space

1.1. Basic Concept

This document addresses the features of the *dataSentinel* Distributed Storage and Web Services products from the perspective of the end user. It looks at how these technologies can enhance the daily experience for consumers of different levels of computer familiarity.

1.2. Supporting Documents

The following documents pertain to the descriptions in this document.

1. dataSentinel-10266-2-1 Distributed Storage Specification.doc
2. dataSentinel-10262-1-1 Web Service Specification.doc

1.3. Definitions

- *Application*: A computer program that allows a user to view and manipulate data.
- *Hard Drive*: A mechanical device that stores large amounts of digital data.
- *Virus*: A malicious computer program that attempts to impair the user experience.

1.4. The Personal Computer Paradigm

1.4.1. Background

The user's perception of computers has gone through a number of transformations as the industry has evolved. In the days of ENIAC¹, the computer was a magical mystery thought to be capable of replacing the human brain that was tended by an army of wizards in white lab coats. During the heyday of IBM (International Business Machines) the computer symbolized the brute calculating force of modern business. It was big, powerful and never wrong.

The average person became a little more comfortable with the technology when the personal computer arrived with the mass-produced Apple II. Oddly, the personal computer quickly developed a negative reputation for being complicated and unfriendly as novices had to come to terms with the early manifestation of the technology. To use the new IBM PC it was necessary to memorize arcane commands and difficult procedures. Despite this, the arrival the SpreadSheet 'killer app' brought the PC to the corporate office and back home in large numbers.

¹ Electronic Numerical Integrator And Computer. The first computer ever built.

A conceptual leap forward was made in the 1980's with the arrival of the Apple Macintosh and its GUI (graphical user interface). The iconic representation of files and folders coupled with WYSIWYG (What You See Is What You Get) views helped the novice user navigate the operational steps of producing documents. A new wave of users came on stream to primarily perform Desktop Publishing. Microsoft cemented its dominance in the Operating System market by releasing its first acceptable GUI in the form of Windows 3.1, and continued to enhance its product with new releases in the decades to follow.

The next breakthrough was the Internet Browser. This application allowed users to easily wander from page to hyperlinked page of views from HTML servers around the world. The computer user was catapulted from creating and managing personal documents and images into a world where it was possible to find and view publicly available data on every conceivable subject. The browser became the default tool for most users and the various search engines became the most popular web sites. Email became commonplace.

1.4.2. The State of the Art

Today's user is now faced with a deluge of data. Digital images created by cameras and phones and music collections account for many gigabytes of data that accumulate on the home PC every year along. Large collections of bookmarks are created from daily browsing. The average user creates (and forgets) dozens of passwords as they subscribe to different Internet services.

All of this data resides on the hard drive of the user's computer and, unless that user has an affinity for technology, is rarely saved to a back-up device. Should a computer fail or become unusable by computer viruses, the typical user loses all of the data and must start from scratch. Sadly, even a smooth transition from an older, functional computer to a newer, faster model results in the loss of data. How many people know how to transfer their Internet bookmarks from one machine to another?

1.5. A New Storage Paradigm

Ultimately, all computer data will be stored ‘on the net’. Just as we all pull our television channels from a plug in the wall and use the telephone network for a monthly subscriber fee, so too will our private and public data reside on massive public infrastructure. Technology will exploit the vastness of mathematical number spaces to protect personal data protection from theft in a way that is superior to simply locking our equipment away in our homes. At the same it will prevent us from ever losing our data.

Today's computer users should be able to work with their own data securely from any computer. They must be freed from the drudgery of managing the storage of this data. *dataSentinel*'s technology allows all users to comfortably interact with their complete, lifetime collection of private data through their own customized interface. They won't have to lug a laptop around and they will never lose data should a computer hard-drive fail.

1.5.1. Your Personal Space: Safe and Secure

dataSentinel has developed a new paradigm for Personal Data Space. Users will all have their own private place within a massive data space created by a large bank of storage computers on the Internet. *dataSentinel*'s technology will guarantee that a user's Personal Space is protected from eavesdropping by any other user -- even from the *dataSentinel* computer administrators, by next-generation encryption technology. *dataSentinel* will not lose your data. Your home computer may crash, but the next computer you buy will see your data as soon as you switch it on.

1.5.2. Your Personal Space travels with you

dataSentinel will give you your own electronic key. It is a small piece of plastic that fits on your key-chain and plugs into any computer (on the USB port) through a small cable. All of the encryption steps are performed by the key when you access your Personal Space. When you disconnect your key, it is impossible for the computer to continue to visit your Personal Space. Now you can go to any computer in the world that is on the Internet and work with your data with the absolute assurance that your data remains safe from others. You can also move from the office to home without worrying about the security of your sensitive business data. Your Personal Data Space is safe, secure and available exclusively to you.

1.5.3. Share your data, but only with those you know

There are times when you want to share. An electronic key differs from a conventional key in that it is actually a container of many keys. If you plug your key into the same computer with another electronic key, you can create a secure shared space between you and the holder of the other key. This space is independent of your Personal Space. Once the shared space is created on the two electronic keys, either of you can access that space independently. With just your own electronic key plugged into any computer, you can place data in that shared space. Any time afterwards, the other holder of that key can plug their electronic key into another computer, view the same shared space and retrieve that data. Like your own private space, this shared space is exclusively available to the two of you; it is hidden from all other users. You can, of course, delete a shared key at any time.

1.5.4. Your Public Space is available to everyone

You also have the option of publishing data for everyone to see. Your Public Space is a sub-directory of an Internet wide file system available to everyone. You exclusively write the files there from the convenience of your own computer. They cannot be modified by anyone else. If you place web documents there, they can be visited from a web browser.

1.5.5. Rent your Applications

Applications such as word-processing and spreadsheets do not need to occupy space on every user's hard drive. Each application can exist within its own Application Space on the *dataSentinel* system. This means you do not have to install software. It also means that you don't have to pay for the full cost of a software product when you only expect to use it a few times. *dataSentinel* will allow you to lease the use of any commercial application for a period of time at a fraction of its cost. If you find that you do use an application frequently, you can increase the length of the lease to save money. As a result, you will not have to install applications, nor will you have to upgrade them. Every time you launch an application, you will be using the latest version.

1.5.6. Controlling Computer Viruses

All computer viruses are spread by the distribution of executable files (e.g. exe files in windows). Either the unwitting user installs a rogue application on their computer and launches it, or the virus is hidden in the executable content (e.g. a macro) of an attachment. *dataSentinel* will store all major applications in read-only application spaces on the *dataSentinel* system where they will be protected from malicious modification. *dataSentinel* will also offer to place the smaller applications on its system once the author has proven that the executable is free of virus code. If you only use the *dataSentinel* application spaces to execute programs, then you are safe from viruses. If you do experiment with unknown programs and your computer becomes infected, all the software in *dataSentinel* application spaces is protected and will remain clean. You can continue using them. If necessary, you can reinstall your Operating System to remove the viruses and immediately resume using the protected applications with the data in your Personal Space.

1.5.7. The new role of the Personal Computer

dataSentinel's technology redefines the role of the computer for the user. It is no longer the holder of personal data, nor does it contain copies of application programs. It now becomes just a computing commodity; an engine to crunch the data from your Personal Space based on the programs executing from *dataSentinel's* application space. The computer is no longer the critical piece. It can be replaced, upgraded or repaired with little interruption for the user. Without the need to install applications, transfer configuration data or perform backups, there is no need to manage the computer experience. It is just a matter of going to any computer at any time to begin working.

1.5.8. Next Generation User Interfaces

The *dataSentinel* system allows thousands of users to have independent Personal Spaces. Portions of each user's Personal Space can be structured in similar ways. This allows custom user interfaces to be designed and made available to all users. For example, should you be a novice user who is not comfortable with hierarchical file systems,

dataSentinel can provide you with a 'house analogy.' You take one digital picture of every wall in your own home, store it in your Personal Space and we will let you navigate through a virtual copy of your home. Where do you store all those pictures of the kids? Just navigate to the child's virtual bedroom and place them on the wall above the bed. Where are last year's tax returns? You placed them in drawer of your virtual desk in the spare room. You now have a familiar context to help you remember where data is within your Personal Data Space.

1.6. How Does It Work?

dataSentinel adds many orders of magnitude of security to existing storage systems by hiding both the content and the location of your data. Its managed redundancy technique ensures your data is never lost.

1.6.1. The Conventional Way to Store Data

Conventional systems encrypt the channel over which your data flows or encrypt your file before it is sent. They have complex authentication processes that prove that you are entitled to store or obtain data from their equipment. In all cases, each file is treated as a single entity and resides on a final server or storage device as one piece. Your data is protected from theft through the physical security of the site or by the encryption of its content. It is protected from loss by a regular backup procedure to off-line media such as magnetic tape.

1.6.2. Using Mathematics to Hide Location

The *dataSentinel* Mass Storage System uses mathematics to hide your data. In preparation for writing, each of your files is broken into small blocks that are individually encrypted. A mathematical algorithm called an 'Address Transform' is executed which uses exactly two pieces of information: the fully-qualified name of your file (which includes the names of the sub-directories it which it is located) and your Personal Encryption Code. This algorithm creates a series of instructions that specify on which of one hundred or more computers each block should be stored and where on the specific computer it should be placed. These instructions are executed to transfer each of the blocks to their final positions. When it is time to read back your file, the Address Transform uses the same information (the name of your file and your privately held Personal Encryption Code) to recreate the storage instructions and get each block back. These blocks are then decrypted and used to reassemble your file.

Think of a jigsaw puzzle. Imagine that each piece of the puzzle has its own 'Peer Index' and 'Block ID' numbers. The Peer Index tells you in which of hundreds of piles of puzzle pieces the piece will be stored. The Block ID allows you to quickly locate the piece in the pile of 100 million pieces. If you know the sequence of Peer Indices and Block IDs that belong to the puzzle in the order in which the pieces fit, then solving the puzzle is not too hard. Now try solving the puzzle without this information when the puzzle pieces are mixed in with many separate piles of 100 million other similar-looking puzzles.

It is not possible to recreate the sequence of Peer Indices and Block Ids calculated by the Address Transform without knowledge of the Personal Encryption Code. This is how *datSentinel* hides the 'location' of your data. An attacker would have to make $2^{64} = 18$

billion, billion guesses on each of hundreds of servers to locate each block in the sequence that forms your file. Each guess requires time to transmit to a *dataSentinel* server computer, be processed and return an answer. Even then, the attacker would have to decrypt each block and reproduce the correct order. This adds up to eons of time even if the attacker has a large number of computers working on a fast network to solve the problem. Your data is safely hidden by the huge numbers involved. At any time, you can decide to recreate your Personal Encryption Code and restore your data. This forces the attacker to start again from scratch.

Mathematics is also used to protect your data from loss. The *dataSentinel* system actually stores a minimum of three copies of each block on different servers. Should any of these computers fail, 'data regeneration' is automatically scheduled within the hour which replaces the third copy of the block on a different computer. This technique of 'Managed Redundancy' ensures that your data is always safe.