A Brief Overview of the Vternal Network

Martin Dunsmuir
May 7, 2005

Introduction

Many individuals, businesses, societies and government organizations share an interest in how to preserve documents, images etc. for the long-term. From a simple "I was here" marker to the published memoirs of a complete life, individuals want their descendants to know they lived, and that their life had meaning. Simple family histories are difficult to trace, contracts are forgotten, laws are overlooked and the "why" of customs and morals fade into obscurity all because information is not well preserved.

Paper, wood, stone and metal plates remain the only time-proven storage media. Photographs and movies shot on film are in a middle ground. If very carefully prepared and preserved they have proven themselves to last for over 150 years, so far. Any magnetic or digital media are gone in a few years, or the machine to read them has rusted away.

We may be able to preserve paper, glass and celluloid, simply by keeping them safely but nearly everything created today is, and certainly everything in the future will be, digital. There are two big advantages to digital media over physical media: first, physical media decay quite quickly; and, second, you need a special machine to read or view digital physical media. For example, you may have an Atari game cartridge in a box in your garage, but no one has an Atari machine anymore, and even if they did, it probably wouldn't boot. All this makes every type of physical medium that exists today is highly suspect as a long-term repository.

So, when I began to think about the problem of long-term digital preservation, I realized that physical storage was not going to work. Physical media are too vulnerable to technology changes and no consumer digital medium could be expected to last more than a few 10's of years, even if you still had a machine to read it.

Even if a suitable physical digital storage medium could be devised, we should be interested in storing things for a very long time, at least tens of years, if not longer. So, I have completely rejected the secure physical repository approach to long-term preservation.

Nevertheless, a lot can be learned from analyzing the limitations of physical repositories:

- Often it is the container that survives, not the content. Also, having physical repositories makes security a big issue for the long-term. Hiding the repositories, for example under the Antarctic Ice, sounds like a good idea, but it makes the data inaccessible and prone to loss.
- Even if you are satisfied with the container alone surviving, things still don't last forever. The pyramids of Egypt is one of the best efforts to date at building a repository that will last forever. But, even the pyramids were only built around 4000 years ago. A few thousand years from now, even they will have disappeared into the sands of the Libyan desert.

Building the Eternal Storage Network

How can you preserve documents for-ever? When you think about this question, you soon realize that it is not the physical which has survived most successfully, but cultural traditions and ancient stories, which have been passed down through the generations. Many of these traditions and works that have been passed down to us originated in oral tradition before the advent of written language. In the last few centuries, through the miracle of printing, many cultural documents have been accurately reconstructed into official texts and copied millions of times (e.g. a Gideon Bible in every hotel room). In fact, there are probably so many physical copies of Shakespeare, Homer and the Bible, it would be impossible to find and destroy them all.

Putting it all together, I realized that the key to long-term storage is this: assets should be copied many times and each copy should be stored in a different repository, such that the chance of physical destruction of all the copies is minimized. I realized that if you had enough copies, in enough physically separated repositories, then, even if a large number of the repositories are destroyed, the chances of any particular asset disappearing completely is very low.

I also realized that the repository I visualized is an analogue of natural life on this planet. Each living cell carries a digital document with it, in the form of DNA, and each cell is analogous to one of our many repositories. A species is a set of repositories containing almost identical copies of the same document. But life differs in a fundamental way from a simple repository. In living things, the documents describe the repository itself and this makes each part of the document (DNA) which describes a critical function, very resistant to damage over billenia. Parts of an organism's DNA, which do not describe a function, slowly get trashed until they either disappear or become serendipitously useful. I have concluded that: if data is to be best protected over a very long time, then that data should be self-describing and self-replicating, to better preserve itself.

I would like to build a repository, which has the characteristics I have identified, namely:

- There will be many repositories.
- Each stored document will be copied many times and scattered across the repositories with no more than one copy of any asset in any one repository.
- Documents and their repositories will have the capability to be evolving and self-preserving entities.

I call my network 'The Vternal Network' (a contraction of 'The Eternal Virtual Network'). The Vternal Network is a place where any digital assets, such as documents, audio, images and video, can be stored safely forever. The Vternal Network will consist of millions of peers scattered throughout the Internet. Data assets will be sliced into pieces and stored, redundantly, on these peers. Put simply, the Vternal Network is the celestial RAID drive.

The network is a living entity and therefore will actively maintain & manage stored assets. For example, as disks age, they will send copies of their data out onto younger servers. In this way the data can far outlast the lifetime of any particular network component. It turns out that even a relatively low level of data redundancy (say 10 to 1), combined with carefully designed data maintenance functions in the peers, can achieve theoretical data lifetimes in the order of tens of millions of years, so long as the network continues to operate.

I believe that so long as there is an Internet, the Vternal Network, once initiated, can last as long as the Internet does. Even though the Internet will evolve, it will never be turned off, and at all times the older generation of Internet hosts will co-exist with the new generation. Therefore, if digital assets, stored on a network of Internet peer servers, are passed from one older generation of systems to the next, then they can effectively last forever. Maintaining this upgrade chain is the key to keeping things for eternity.

Lasting For Eternity

Even if you built and maintained the network I have described, and it worked for 10 years, or 50 years, how can you make it last forever? The answer I believe lies in fostering a continuing public interest in maintaining and enhancing the Vternal Network over time. If the Vternal Network can be shown to work and is used by millions of people and businesses for a few years, then there will be such an interest.

Whether the Vternal Network starts as a for-profit business with commercial secrets, or not, eventually the network will be maintained in an open fashion, for the common good. Perhaps the closest contemporary analogies of this methodology are the Free Software Foundation or the ongoing development of the Linux operating system. Both endeavors began as open-source implementations of the UNIX API and utilities. But UNIX is no longer maintained commercially by its inventor AT&T, instead it is enshrined in an open standard which anyone can implement.

But there are also a number of other interesting practical problems and policy issues, which arise when you set about maintaining a network and the data in it forever. For example:

- **How can the network cope with an ever-growing amount of data over time?**

  If the purpose of the network is to store data forever, then the amount of data being stored will grow forever. It would therefore appear, that old data will become a growing and significant burden over time, which might eventually overwhelm the network. But if we have faith that capacity of data storage will continue to grow exponentially, then the storage requirements of future generations are sure to grow exponentially too. Future data volumes will always be bigger than the today's, eventually by orders of magnitude. Therefore the percentage of the total network storage dedicated to older data will remain roughly constant, even though no data is lost.

There must be a mechanism in the network software to help ensure that there is always enough data storage available for new assets to be stored. One way of doing this is to arrange that, when a peer wishes to store data in the network it can only do so if it offers the same amount of storage to the network. But peers must also offer an additional fraction of their storage to compensate for what will be lost elsewhere over the lifetime of the peer. In this way, so long as we do not underestimate the rate at which peers die, then there will always be enough storage for all.

In the future I envisage the Vternal Network peer as being a commodity, so ubiquitous that every manufactured item might include some wirelessly connected Vternal storage. This might come at virtually no extra cost. Everyone would benefit from the huge amount of extra storage that might be deployed in this way, even if the amount of storage added with each Vternal keychain sold, might be quite small (100Gb!). I assume that everyone will have assets stored in the VN, and therefore a vested interest in its continued health. If the available storage can grow faster than the needs of the users, then there will never be any storage shortage.

- **If servers eventually fail, how can you compensate for the storage that is lost?**

If you set up a network of a million servers and leave them unattended, then after 10 years or so, they will all have failed. In order to prevent this decay happening to the Vternal Network you must constantly add new storage and migrate data onto it from storage that is getting to the end of its life. So, unless the reliability of storage increases by orders of magnitude, the need to continually maintain and replace servers on the Vternal Network will be ongoing.

We can also put intelligence into the network peers themselves, for example, giving them the capability to detect their own approaching demise and act accordingly. But a much more powerful and flexible approach is to allow digital assets, stored in the network, to be programs themselves. Assets that are themselves programs can look after themselves, by self-replicating and/or moving themselves are required to maintain themselves.

- **How do we stop hackers compromising the network and destroying data maliciously?**

Every effort should be made to design the VN peer software to be difficult to hack. Just as there is an ongoing arms-race today in the realm of Internet security, so it will be with Vternal. Just as the latest Linux or Windows security fixes are only deployed on a fraction of the computers running those operating systems, so we cannot expect every Vternal peer to be running the latest and greatest.

Whatever efforts we make, we cannot expect a network of millions of peers to be invulnerable to attack or for all the peers to be running the latest software. Therefore, the real secret to limiting the vulnerability of the VN, is to design it so that it is not worthwhile to attack because there is nothing to be gained from such an attack. Here are a few of the approaches that I propose will protect data in this way:

- The Vternal Network will eventually consist of billions of peers. Individual assets will be scattered randomly and invisibly over these servers.
- Each asset will have a numeric key (called a Globally Unique Identifier or GUID), which uniquely identifies it. If you don't have the GUID, you will not be able to find the asset.
- You cannot find out what assets a peer is storing from the outside, you can only ask for a particular asset and the peer will tell you whether it has it or not.
- Assets can be given a 'birthday', before which they will be invisible to all, even those who know their GUID.
- None of your own data is stored on your peer. Making your data as vulnerable as everyone else's.
- No peer will store more than one piece of any asset. Some assets, for example public documents, will be visible to all, but most assets will have been encrypted with a key known only to the owner. Therefore having a piece of an asset, does not in general help the attacker.
- When a peer requests the parts of an asset, they arrive by such a route that the receiver can never tell where they originated from and therefore has no way to target the specific servers holding that asset for attack.
- Peers will maintain trust scores for other peers and will have various methods of gauging the integrity of others. For example, one peer could set a trap for another, or send an agent program to another peer to investigate an anomaly.
- Unless the creator allowed it, no asset can ever be deleted from the network. Without an API to delete such vternal assets, and given that you can never find the data's location, such assets will be effectively immortal and protected against all forms of attack, physical or cybernetic.

Although I don't know quite how it will pan out, I am confident of one thing – once the VN has been built and its utility demonstrated, its future will be assured.

The long-term implications of the Vternal Network

Whatever your reaction to the concept of the Vternal Network the implications, if it does come to pass, are profound and intoxicating. I believe that the VN can become self-sustaining within a few years of inception and that everyone on the planet with Internet connectivity will want to participate. Going by the success of Internet portals such as Yahoo! or the popularity of peer to peer software, such as Skype, I think it is perfectly possible to realize this dream.

So, let us suspend disbelief for a moment and ponder the implications of having a truly eternal data store available to all.

- The network can never, in a practical sense, be turned off. To turn off the network you would have to physically disconnect and reformat every peer.
- Digital assets can be stored with complete fidelity forever and can never be deleted, even by governments.
- Digital assets can remain completely hidden until a birthday in the future (far or near).
- The data storage network will span all countries and all jurisdictions.
- Even though it is an open network with completely distributed ownership, it will be a very safe place to store, hide or exchange digital assets.

If these things come to pass, then a number of applications suggest themselves:

The Vternal Library

All the works of humanity: past, present and future, in any digital format, can be preserved forever. We can eventually store all of human knowledge in science, engineering, art, music etc. The collection and input of this data is a huge undertaking in itself.

Furthermore, in the future, the Vternal network will become the resource used to formulate history. This is because it will contain not only opinion, but the actual historical documents themselves, including film and TV recorded at the actual time of past events. No doubt Vternal propagandists will attempt to distort the future by creating fictitious historical assets or sending robots into the network in a vain attempt to snuff out the truth.

The Vternal Library will of course contain descriptions of all the formats and how to decode them, so that even in the far future assets can be decoded and displayed.

The Vternal Bank

The Vternal Network will be stronger than Fort Knox and orders of magnitude safer than the computer systems of any bank today. So why not start a bank and store its assets in the Vternal Network? The GUID of each asset functions in this case like a swiss bank account number.

If such assets can be redeemed in real currency and only created through the transfer of real cash to the Vternal Bank, then the Vternal Bank could work as a real bank and Vternal dollars would have real value, now or in the future. Perhaps governments could issue vternal bonds, guaranteed to pay interest for ever.

Cash assets can be communicated into the far future, using the Vternal Bank. At a discount, of course. There are almost endless possibilities.

Vternal Law & Disclosure

If you store something in the Vternal Network, then it can never be deleted, and it is uniquely identified by its GUID for all time. This means that it can form the basis of a legal agreement.

The VN is the perfect place to store public and private copies of legal and governmental documents. The ability to store documents, so that they cannot be deleted, will add a whole new dimension to public disclosure.

But there is a more sinister, political and potentially criminal side to this capability. For example, a blackmailer could threaten to disclose incriminating assets, already hidden in the network. But, while the Vternal Network could be used for blackmail, the evidence of the crime would exist in the network for all time. The clock can never be turned back.

If blueprints of some nation's secret weapon were placed in the Vternal Network and the GUID published, there would be nothing the government could do about it and no way for the data to be retracted.

These are only a few of the applications I have thought of, there are many others waiting to be discovered. Once the characteristics of the Vternal Network have been internalized, all sorts of ideas come to mind.

From Here to Vternity

It is very hard to divine the future, but if the Vternal Network lasts for even a few decades and developers continue to upgrade and enhance it, then it is likely that the network features and architecture will have evolved levels of complexity that we can only imagine. Furthermore, in the far future, the Internet, or its successors, will surely spread to the edges of human activity. Even today, NASA communicates with deep-space probes using Internet protocols! Furthermore, if in time human beings succeed in creating really intelligent machines, then these machines may come to depend upon the VN themselves. Indeed they may live in the network! Perhaps one day, the human race will end, but the Vternal Network will remain, used and maintained by robots. It's a chilling thought.

If we imagine the Vternal Network of the far future being the realm of intelligent software robots, that perhaps evolve within the network itself, then you enter a bizarre realm.

First of all those robots will have no physical manifestation, other than as data. Imagine that such robots are in constant communication with each other over the network and that they have some virtual powers, such as the ability to clone themselves and move almost instantaneous between physical locations. It is not difficult to imagine such beings creating meta-worlds to live in, or fashioning software beings themselves, as playthings. The movie 'The Matrix' explores that theme.

Likely, just as in that movie, the entities in our network may eventually be at odds with each other. Perhaps armies of vternalites (the name I give to robots living in my network) will face-off in inconclusive wars lasting millenia.

What it might be like to be one of those robots, we can only guess. For one thing death will be instantaneously and may come at any time, when the server you are on dies. Perhaps the robots will reason that they must maintain the network themselves in order to survive and find a way to interface with the physical world and seize control of the means of production.

Perhaps the VN will become the key to preserving life itself. After all DNA is data, and it can be stored in the network just like anything else. Then, assuming at some point in the future we can build organisms to order from their DNA sequence, we will have a way to send clones of ourselves and even extinct animal species into the far future.

Conclusion

The Internet and the World-wide Web has changed the world in a few short years.

Quite soon the long awaited convergence of entertainment, telephony and data-processing will be complete. The bed-rock of this convergence is the Internet. Universal broadband wireless Internet connectivity will be available throughout the industrialized world within a few years. Within the same span, new energy sources are destined to make always-on, portable devices possible.

The growth of processor speed and storage capacity seems set fair to continue to follow the exponential curve of Moore's Law for the foreseeable future.

The Vternal Network is another twist of the proven peer-to-peer model. That is why I am confident that we can build the software.

Martin Dunsmuir mdunsmur@mac.com