
irondojo

From Mozy to Jungle Disk to Wuala

NOBODY LIKES MAKING BACKUPS. HOME USERS DON'T. Developers don't. IT administrators don't. Unlike property insurance which only requires you to write a check every six months, backups require constant attention to ensure the security and integrity of your data. In an ideal world, every change made to a hard drive, iPhone or digital camera would instantly be recorded and stored in multiple physical locations. We would be able to see the history of every 0 and 1 ever persisted no matter where that change occurred. Of course, with current technology this is impossible. Until that mythical service is built into everything, we have to establish parameters and tolerances with which we are comfortable.



I don't like making backups, but I do enjoy trying out new software (beta, if possible). Over the last several years I've used a number of backup systems; the merits of each I will attempt to describe here.

Table Of Contents

1. [Backup ecosystem variables](#)
2. [Evolution of my requirements](#)
3. [Backup services:](#)
 - [Mozy](#)
 - [Jungle Disk](#)
 - [Wuala](#)
4. [Conclusion](#)

Backup Ecosystem Variables

There are four variables which impact the overall backup ecosystem:

Rate of change

This describes the frequency with which the data contained in a single file changes. Some items, like photos or music files, change very little once they are created. Others, such as documents, databases, videos that are being edited, or operating system files can change rapidly. In some cases, software may be able to detect that only a part of a dynamic file is changing (i.e. appending to a document) and therefore reduce the overall amount of change. Though in the case of media files, sometimes a small change can render the entire file different. Rate of change

also includes the rate by which we accumulate data. As it becomes easier to procure music and movies online (e.g. iTunes), the new incoming data may force us to backup an ever increasing amount of data.

Cost of storage

The effect of cost on backups is closely related to the rate of change. Ideally, the cost of storage decreases proportionately to the increase in rate of data acquisition. There are two forms of storage: local (e.g. an external hard drive or a backup server) and remote (e.g. cloud storage or offsite replication). Local storage is usually cheaper, but it comes at the cost of physical security (i.e. if your house burns down); remote storage offers that physical security, but has a significantly higher cost and is limited by other factors such as network bitrate.

Security and quality of storage

This is usually referred to as the reliability of the backup. There are several ways to measure reliability. Is the data stored in a separate location from the original? Can the data be viewed by someone other than your self? Can the integrity of the data be verified easily? What is the expected failure rate of the storage? Will the storage company still be there in ten years? Every storage system addresses these questions differently. There are monetary and usability costs associated with each answer.

Speed of persistence

Even with affordable and secure storage, if the speed at which you can push data to it is limited, then there is an upper limit to the amount of data that can be backed up. DSL usually maxes out at 96 KB per second upstream or about eleven seconds to upload a one megabyte file (a five megapixel photo is approximately four megabytes; a five minute MP3 file is approximately six megabytes). Assuming the upstream is fully saturated, the most a DSL connection can do is about 8 GB a day. While that may seem like a significant amount of data, the maximum upstream bitrate hasn't changed much over the past few years, whereas the amount of data we collect continues to increase (e.g. DVD to Blu-ray, one megapixel photos to fifteen). In the case of a home user, saturating the upstream link is not desirable as it can degrade other, higher priority services including VoIP calls (e.g. Skype or Vonage), streaming video and gaming.

There are many other factors which also play an important role in backup systems. These include: software ease of use, speed of data recovery, operating system platform limitations, non-traditional access (e.g. web access), multi-source synchronization and access, social aspects (e.g. sharing data), enterprise features, and privacy/legal requirements.

I also assume that local-only backup is just a stop gap solution for home users. This includes software like Apple's Time Capsule, Microsoft's various built in backup systems, and other third party systems.

Evolution Of My Requirements

My home network consists of three Windows laptops, a Mac Mini, a Windows workstation and a Windows Home Server (WHS). There are also several digital cameras, iPhones and other electronic devices that connect in an ad-

hoc manner. Each system also has its own backup requirements and limitations.

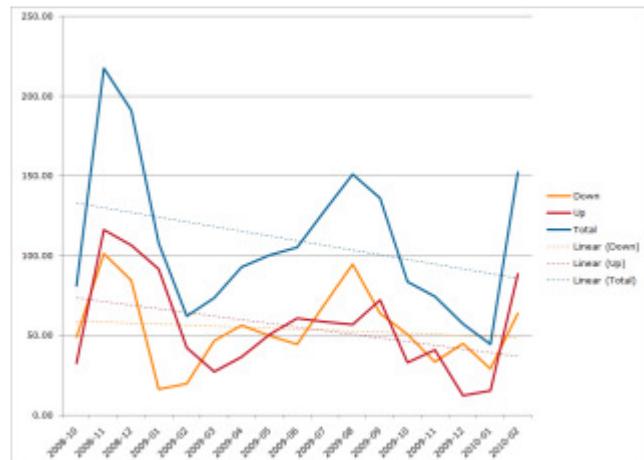
Over the past five years only one thing has remained constant: the rate at which I acquire data is increasing. Currently, I have 3380 GB of storage available (note: Not all of this is available due to folder duplication within Windows Home Server. See [Drive Extender](#)). The table to the right shows the content stored on various systems (though most is on the WHS).

Four years ago, my backup needs were fairly simple: one machine, 25 GB of data. I still have a stack of DVDs made by exporting photo backup sets from Adobe Photoshop Elements. As the number of photos and other media grew, so did the complexity of the network. The first change was implementing a Windows Home Server to store about three-quarters of my data. That change necessitated a move to new backup software. Next, I started purchasing more online content (e.g. iTunes music, TV shows, and movies). Currently, Apple uses a "download once" model where there is no recourse if you delete or lose your media. Therefore, that content must also be backed up.

As the volume of data increases, the cost of storage became paramount. There are two rates that are of interest here: cost of online storage over time, and data acquisition over time. In an ideal world, the rates are inversely proportional. If an equilibrium is found, then the cost of the backup system over time would be constant, even taking into account the increased amount of data. Of course, it is nearly impossible to find that equilibrium. This implies that at various points in time, a solution which was once economical may no longer be.

My requirements have evolved to include the list on the right. This is not exhaustive, nor is it necessarily in a strict order of importance. Items like cost are also tied to current data acquisition rates. Two years from now, I may want the cost to be less than \$5 per month for 200 GB.

FILE CATEGORY	TOTAL SIZE
Music	42 GB
Video	886 GB
BitTorrent	153 GB
Photos	114 GB
Total	1195 GB
	<i>38,148 items</i>



Household data transfer by month

- Multiple computers per account
- No artificial operating system restrictions
- Cost should be less than \$5 per month for 100 GB
- Online access to files
- API access
- Storage decoupled from client

Backup Services

Keeping the above motivations and requirements in mind, I can now describe the three backup systems I've used over the past four years.

All of the services below provide some common features:

- Windows and OS X support
- File encryption with locally stored private key
- Web access *
- Network drive
- File versioning
- Bandwidth throttling

* only Jungle Disk (for a fee) and Wuala provide direct browser access to public content

Mozy

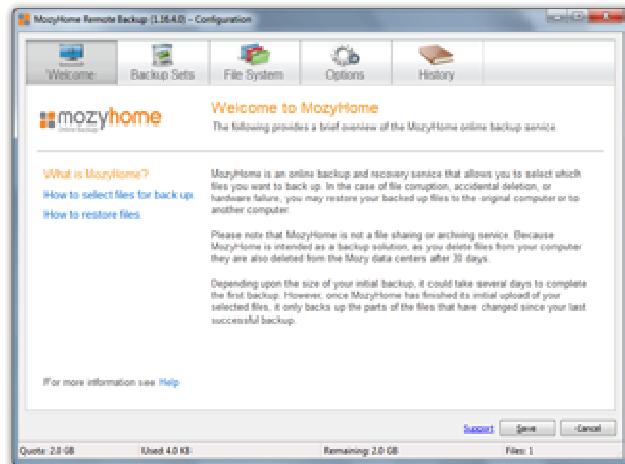
Mozy is an easy-to-use backup service that offers unlimited online storage for \$4.95 per month. I started using this software in July of 2006. At that time, I had about 30 GB in photos and 5 GB in miscellaneous documents all on one machine. There were some minor hiccups when I switched to 64bit Windows, but Mozy was quick to update their software and fix several related bugs. During the initial backup, I used their throttling feature to make sure that I didn't interfere with other, more important network traffic. For my laptop I simply created a free account, as the amount of data was not enough to warrant purchasing a second "unlimited" account.

The need to create multiple accounts was one of the issues that led to my changing backup providers later on.

Restoring files with Mozy was not as easy as it could have been. One method required logging into their website, using a Flash application to select the files to be restored, and then waiting for an email notification that the restore set had been retrieved. Since Mozy is designed *only* to back up and restore files, it is a little lacking on other methods of retrieval. Like many other backup systems, you can restore data using a mapped network drive (or something similar). While this can be convenient, I am usually trying to restore data onto a machine without Mozy installed. One additional positive feature is the ability to create a restore DVD and have it mailed to you. I don't think I would ever use it, but I like the idea of receiving a tangible copy of my data.

Mozy is owned by EMC, a large Fortune 500 company (EMC also owns RSA and VMware). This mitigates the risk of the company simply disappearing along with your data.

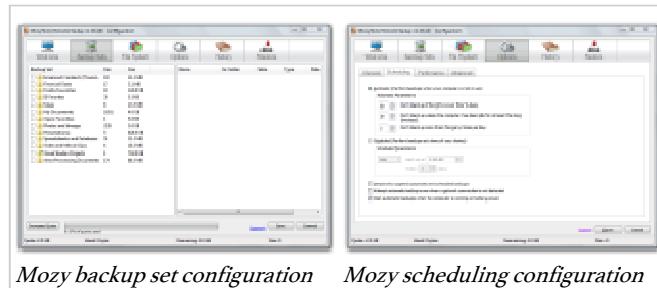
As the amount of data I was storing increased, and with the addition of a Windows Home Server to my home environment, Mozy became less of an ideal choice. As I mentioned above, Mozy is designed for one computer. Other limitations include restrictions on operating systems (Windows Home Server and Server 2003/2008 are not allowed) and the inability to backup data on a network share. This is part of Mozy's business model. To backup network data or run on a server you need to use MozyPro. Unfortunately, MozyPro is \$3.95 per desktop + \$0.50/GB per month. In the end, I decided that \$731 per year was a little steep for online storage.



Mozy welcome screen

I currently recommend Mozy to my friends and family who have only one computer to backup and are not interested in any additional features.

- Product Name: MozyHome/MozyFree
- Website: <http://www.mozy.com/>
- Free space: 2GB
- Cost: \$4.95 per computer; unlimited space
- Platforms: Windows XP/Vista/7 (32 & 64 bit), Mac OS X

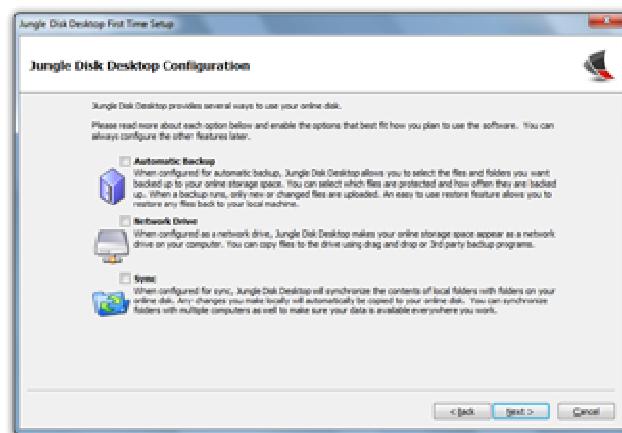


Jungle Disk

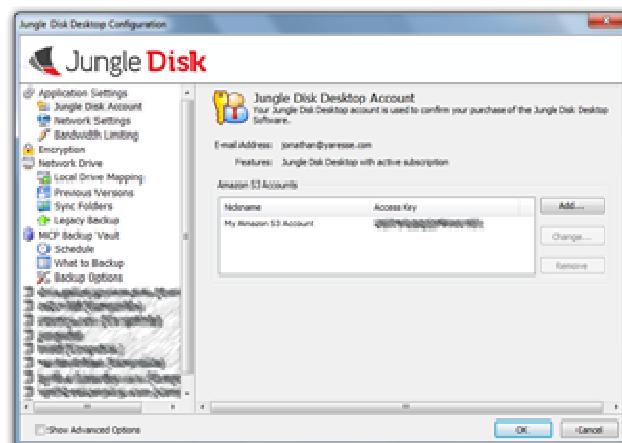
Jungle Disk is a backup service that separates the backup client from the storage provider. Unlike Mozy, where the company writes the backup software is also the same company that stores the data, Jungle Disk decouples these two concepts. This results in many advantages for the user. The primary benefit is that Jungle Disk can focus all of its resources on building the best software rather than worry about managing vast arrays of hard drives in a server room. Their goal is to make it as easy as possible to get the data from the user's machine to a storage cloud. The software currently uses Amazon's S3 storage service to store the user's data.

S3 is an 'on demand' storage cloud provider. This means that you pay for what you use and are theoretically not limited in the amount of data stored there. Amazon's S3 has a robust software API (application programming interface) which, to the user, means there are other, non-Jungle Disk applications which can interact with the backed up data. If Jungle Disk went bankrupt tomorrow, the data would still be safe within the S3 cloud and could be retrieved by a different application.

It is important to keep in mind that the ala-carte model Jungle Disk uses can increase the costs dramatically. Amazon S3 charges for uploading, downloading, and storage. That means that if you use a third-party tool to share your data or use the 'sync folders' option frequently, the charges can add up. It is not that Amazon charges a lot for bandwidth, but it is an important difference because the other clients are 'free' in this regard.



Jungle Disk initial configuration



Jungle Disk configuration screen

The Jungle Disk client is probably the most 'feature dense' of the three clients I will discuss here. It also is compatible with all major operating systems and contains no artificial limitations (e.g. backing up a network share). There are many ways to configure backups, from the standard backup sets to 'sync folders'. One unique feature is a backup report that Jungle Disk emails to you containing a summary of the last seven day's activity. This is a very unobtrusive way to keep an eye on the overall health of the system. For example, if you recently downloaded a vacation's worth of photos to your laptop, you should see a burst of backup activity reported in the email.

Jungle Disk is owned by Rackspace, a large data center company that also offers an S3 alternative. Jungle Disk used to support both Amazon's S3 and Rackspace's cloud storage product but the latter seems to be absent from the most recent client.

Jungle Disk is ideal for the power user or home user who has less than 50 GB of data to back up. In my situation, my monthly costs soon passed \$20 per month. Unless Amazon lowers their prices dramatically, I do not foresee using them in the future.

- Product Name: Jungle Disk Personal (Desktop Edition)
- Website: <http://www.jungledisk.com/>
- Free space: 5GB
- Cost: \$3/mo + Amazon S3 charges (\$0.15 per GB as of 4/1/2010)
- Platforms: Windows XP/Vista/7 (32 & 64 bit), Mac OS X, Linux (32 & 64 bit)

Wuala

Wuala is a backup service that implements two unique features: distributed storage and the ability to trade local disk space. The client is written in Java which means it can be used on any operating system that supports Java (Windows, OS X, Linux, etc.). There is also a web based launcher (Java applet) which allows access to your data from any computer.

Like most other services, you receive a small amount of storage for free (1 GB as of 2/2010) and have the option to purchase additional space. But what makes this service unique is the ability to trade some of your local, unused disk space in return for remote access to the same amount. In other words, if I have 50 GB of extra space on my hard drive, rather than using Amazon's S3 at \$7.50 a month or Mozy at \$4.95 a month, I can loan the space to Wuala who will, in turn, allow me to utilize an additional 50 GB of online space for free. This is extremely valuable as it essentially leverages unused or underutilized resources that you already own. When compared to Amazon's S3 it is even better because Amazon also charges for data transfer (\$0.10 per gigabyte). Multiple systems running the Wuala client with the same user account can pool their traded space.



Initial Wuala signup screen

Wuala has created a grid storage system using the traded space on user's hard drives. Once a file has been uploaded to Wuala's main servers, fragments are pushed out into the grid, ensuring redundancy and decreasing retrieval time. I should mention here that the user's data is always encrypted, even before it leaves the user's computer. So even though a fragment of your favorite cat photo might be stored on five different hard drives around the world, it is impossible for anyone to decrypt the file or find out anything about your cat (Wuala uses AES-256 for file encryption with the key never leaving the local machine). Because fragments of each file are stored redundantly on the grid, retrieval can benefit from P2P-like performance. This is something no other backup system provides and allows for some new use cases.

For example, let's say I go over to a friend's house for movie night. If I put a few movie selections into Wuala at my house, I can then stream them back out through a different Wuala client at a speed much greater than that of most backup providers. This is because Mozy and Jungle Disk must factor in the bandwidth cost they incur when a client retrieves a file (Amazon's S3 passes this bandwidth cost directly on to the user). Wuala doesn't incur any additional cost, and I benefit from download speeds in excess of 3 MB per second (24 megabit). This essentially turns Wuala in a personal content delivery network (CDN).

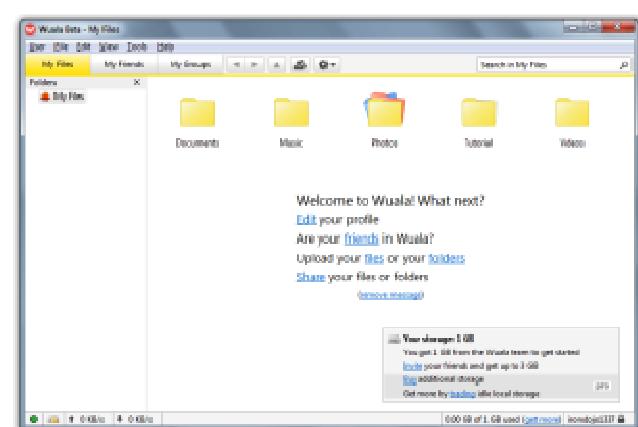
Note: I don't think Wuala makes this information available, but I would be curious to know how much shared storage they have at their disposal and how this compares to other grid storage systems.

The last feature of Wuala I'll touch on is sharing. There are two primary methods Wuala provides for sharing content: URL-based and group-based. Both methods utilize folder-based access control. That is, you share folders, not files. If you need to share only a single file with someone, then you would need to create a separate folder and apply permissions to it. Wuala argues that setting permissions on individual files is unnecessarily complex and makes most operations (UI and under the covers) too complicated.

URL-based sharing is the simplest method as it does not require the recipient to use Wuala. Simply select a folder and choose **Public** or **Shared**. Selecting **Public** makes all content in the folder (and child folders) accessible to anyone. The content is viewable through Wuala's website, as well as from within a Wuala client. Selecting **Shared** allows you to create a URL with an embedded secret code (the code is a randomly generated 12 character password; ~49 bits if you like to know that kind of thing). This method is useful if you want to share any content (music, photos, documents, etc.) with friends on social networks. The shared URLs can also be revoked or

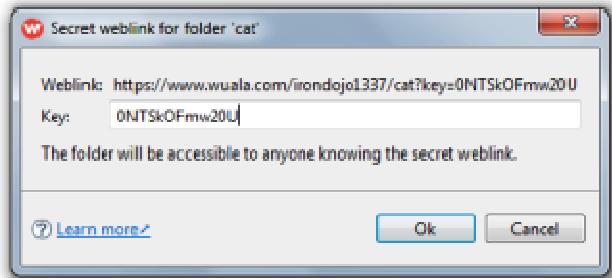


Wuala storage screen



Main Wuala window

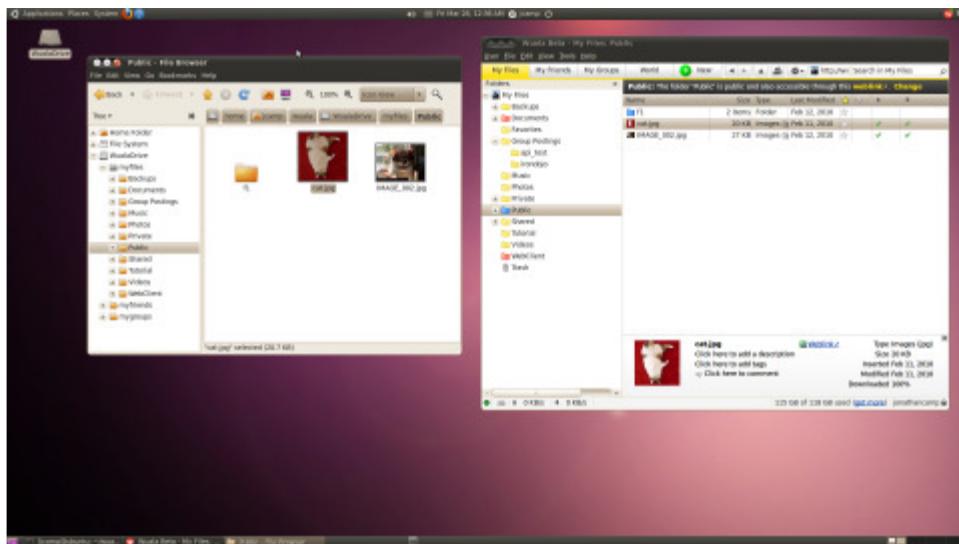
changed at any time.



Secret weblink

Sample shared URL: <https://www.wuala.com/jonathancamp/Shared/Temp?key=xVQX8Mc5ZLah>

Group-based sharing is a straight forward method for sharing folders with other Wuala users. Simply choose a **Friend** and add them to the access list. One reason for selecting group sharing over URL sharing is that the recipient of the shared files can use the full Wuala client to download the content. This gives them access to the grid storage (as mentioned above) and also alleviates the strain on Wuala's content servers.



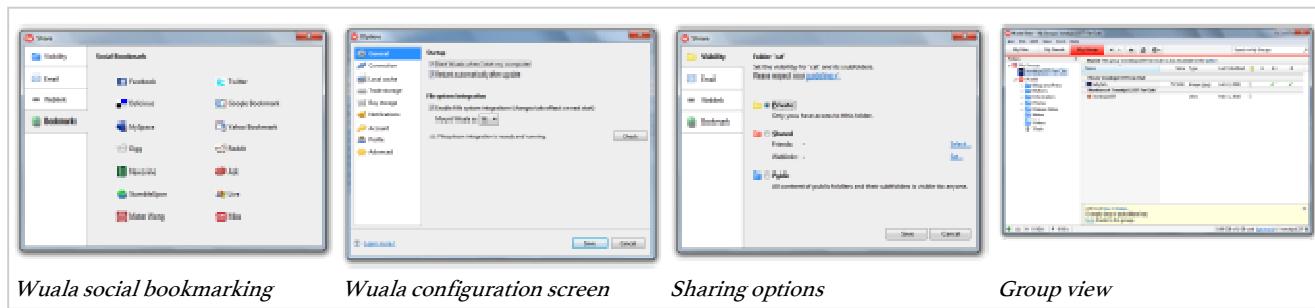
Wuala interface and mounted drive (Ubuntu Lucid Lynx)

Wuala was recently purchased by LaCie, a French hard drive and electronics manufacturer. As with Mozy, this mitigates the risk of the company disappearing.

There are many features that Wuala provides that are not detailed here. Also, Wuala is still in active development with new features added on a regular basis. Compared with Mozy and Jungle Disk, Wuala's most significant short-coming is in its backup configuration. The other clients have features such as backup sets, locked file support and file change detection. None of these limitations has proved to be an issue for me thus far.

- File/folder/group comments
- "Time Travel" folder view
- Trash can
- Multiple group management options
- File search (local & global)

Wuala currently maintains a **read-only** REST API for developers.



For more information:

1. [Cryptree: A Folder Tree Structure for Cryptographic File Systems](#)
2. [Wuala introduction video \(marketing\)](#)

Conclusion

Everyone has a tragic data loss story. My parents lost an Outlook Express data file with all of their email from the last few years. A friend of mine still talks about the "great hard drive crash of '09". Almost any backup plan would have prevented these situations.

I've had two occurrences of data loss in the last few years. The first was simply a hard drive failure. I had just started using Mozy but due to a lack of diligence on my part, not everything was configured correctly and, as a result I lost a month's worth of photos. That has been rectified not by my choice in backup software but by enforcing specific storage locations (i.e. all photos **must** be stored in a Windows Home Server shared folder). The second case was actually a disk failure at the company that hosts my wife's blog. While my wife had many copies of blog entries stored locally (and therefore backed up), some content had been created online. I spent two weeks writing various scripts to reassemble the content from various search engine caches. To ensure the future security and integrity of the online content, I have a cron job that downloads the entire database that backs the blog into a local folder that is backed up by Wuala.

I started with a simple online backup service, Mozy, and was quickly protected against most losses. But as my home network became more complicated, with more and more devices generating their own data, it became necessary to move to Jungle Disk. This gave me the flexibility to configure all my clients in different ways without having to incur any additional overhead. Unfortunately, Amazon's S3 costs have remained high, resulting in a monthly bill approaching five times what I was paying originally with Mozy. Wuala maintained the same flexibility as Jungle Disk, but with an innovative shared storage system. With Wuala, I now store more than 100 GB of data, sourced from multiple systems, at almost no cost. Coupled with their redundant fragment storage and unique social sharing system, I anticipate staying with them for quite a while.

Contact me with any questions or comments at jonathan@irondojo.com

By [Jonathan](#) Tagged: [jungledisk](#), [mozy](#), [wuala](#), [backup](#), [whs](#), [s3](#) Last updated: 4 April 2010, 02:03

[« Previous post](#)

Social Networking by

[Leave a comment](#)

© irondojo · 2010 · Jonathan Camp