

Inside Risks

Risks and Myths of Cloud Computing and Cloud Storage

Considering existing and new types of risks inherent in cloud services.

CLOUD COMPUTING AND STORAGE are often seen as general blessings, if not financial salvations. There are good reasons behind this claim. Cloud services are indeed usually much cheaper than their dedicated counterparts. Administration and management oversight are simpler under a single, central authority. Small businesses and startups have taken advantage, using low-cost cloud services during their first few years. Cloud platforms are critical avenues to getting started for many companies, giving them access to many customers at low cost. Many business leaders see the cloud as an engine for small businesses and job creation.

Cloud storage services are also a boon for individual users, most of whom do not back up their computers and mobile devices regularly or at all. Cheap, automatic backup to cloud storage protects their valuable data from loss.

Despite all these bounties, cloud services also present new kinds of risks, which are considered here. Prospective cloud users should evaluate these risks before making their decisions about how to use clouds. The main issue is that expectations of trustworthiness may be unrealistic. Confidentiality, system integrity, data integrity, reliability, robustness, resilience may be questionable. Protection against surveillance, and

denials of service are essential, as are perpetual access and long-term compatibility of stored data. The integrity, accountability, and trustworthiness of potentially untrustworthy third parties and even unknown *n*th parties must also be considered. Those parties may have business models that are radically incompatible with user needs; furthermore, they might go out of business—with users holding the bag. Moreover,

insider misuse may create additional risks. All these risks are relevant to many different types of applications. As one example, from users' perspectives, having unencrypted email maintained by a cloud provider may be particularly risky.

The basic concept of cloud computing and cloud storage has a lineage spanning two generations, with significant experience in designing and administering these systems. Time-

sharing systems with common computing resources and possibilities for collaborative data access have been around since the 1960s (CTSS, Multics, Tymshare), with varying types of sharing. Since the 1980s, Project Athena at MIT has employed the Sun Network File System, and later the Andrew File System, to provide three services that would be identified with today's cloud services: remote storage of application programs, remote storage of personal files, and remote backup of personal files. However, time-sharing and Athena's three services have been under single operational administration, thus minimizing the number of entities that users must trust (while at the same time providing a single point of failure). We know from experience how to offset some of the risks when cloud services are the responsibility of a single administration. With respect to both local and remote servers, some of the risks can be reduced. For example, private systems and intranetworks under local control or more likely the control your own employers (with respect to hardware, software, certificate authorities, and pooled system and network administration) are likely to have greater trustworthiness.

What is new—and the source of new risks—is the scale and distributivity of some of the clouds. They are large distributed systems with few centralized controls. Clouds that provide access to vast amounts of information (such as Google and Amazon) are extremely valuable resources. However, other clouds that store your own data (along with everyone else's) can present serious problems relating to trusting potentially untrustworthy entities.

Giving the “cloud” name to the old concept of large, shared, distributed systems is misleading. It creates a new buzzword, and hides the problems of risks that designers and admins have otherwise grappled with for years. Some of the cloud providers have ignored many of the old risks and are evidently largely oblivious to newer risks as well. Clearly, *cloud computing* is simply *remote computing*, which was one of the primary reasons for the creation of the ARPANET—to allow people in one coastal time zone to benefit from unused resources in other time zones at

Some of the cloud providers have ignored many of the old risks and are evidently largely oblivious to newer risks as well.

certain hours of the day. This has clearly been an even greater benefit in the Internet, with its worldwide coverage. Similarly, *cloud storage* is simply *remote storage*, which in early days became common as off-site backup for obvious reasons of fault tolerance, emergency preparedness, and other reasons.

One risk in identifying remote storage for offsite backup as “cloud storage” is that this term masks the existence of in-house alternatives, such as the common practice of periodically recording file-system snapshots on small detachable media, and keeping them in a safe place. This can be particularly important after nasty penetration attacks that may have compromised a system with the insertion of malware, sniffers, and so on. Furthermore, remote archiving—especially if widely distributed among different repositories—leaves users unsure of whether their information is still retrievable in its original form (unless they have actually retrieved it).

Ron Rivest has been quoted as saying, “Cloud computing sounds so sweet and wonderful and safe ... we should just be aware of the terminology; if we [are] calling it swamp computing. I think you might have the right mind-set.”²

To paraphrase a quote often attributed to Roger Needham, Butler Lampson, or Jim Morris, if you think cloud computing and cloud storage are the answer to your problems, you do not understand those would-be solutions, and you do not understand your problems.

A few examples of relevant recent risky exploits are worth noting here. (Further background on the first nine items and other examples of cloud

compromises can be found in the ACM Risks Forum: <http://www.risks.org>.)

- ▶ Dropbox's sharing services were hacked, resulting from a security hole in its link-sharing scheme. The exploits were also disseminated by the perpetrators.

- ▶ No-IP had 22 of its most frequently used domains taken down by Microsoft, under a seemingly overreaching federal court order.

- ▶ Amazon Web services have gone down (briefly) several times. Code Spaces (a valued source-code repository built using Amazon's AWS facilities) was effectively destroyed by an attacker demanding ransom.

- ▶ Cisco Systems had a private crypto key embedded in their VoIP manager that allowed unauthorized control of sensitive messaging gear.

- ▶ Cryptolocker and other ransomware programs have forcibly encrypted stored information, and demanded payment to decrypt it (although in some cases have never done so even after receiving the ransom!). Although most of these attacks have been on individual users, the opportunity for attacks on remote storage repositories is clearly a risk. (Recently, an antidote website has reportedly been created.)

- ▶ TrueCrypt (full disk encryption) was discontinued as source-available software by its pseudonymous authors, “as it may contain unfixed security issues.” (Uncertainty remains as to the severity and impact of those possible issues, and how they found their way into the codebase.)

- ▶ Similar things happened to Lavabit, which provided privacy and security features in email services to over 400,000 customers, but was then withdrawn after prolonged legal harassment that attempted to coerce the installation of surveillance equipment.

- ▶ Megaupload.com was taken down by authorities, blocking both illicit and legitimate users.

- ▶ Nirvanix went belly-up financially, giving its users two weeks to exit.

- ▶ Various talks at Black Hat and DEF CON in August were rather disenchanting. In short, essentially every device seems to be compromisable, often with a fixed master password embedded in the system, but with many more subtle vulnerabilities as well. This is old news to Inside Risks readers, but

could be shocking to everyone else.

Among old risks that are still pervasive, even in-house use of local storage can result in hardware outages and database software failures. Redundant copies might actually all wind up in a single vulnerable cloud repository. Furthermore, older data formats may no longer be supported. Local storage still requires attention to backup that can be successfully retrieved—in some cases many years later. Furthermore, if the original information is encrypted, the ability to manage and recover old keys becomes critical.

Recently, increasingly efficient cryptographic schemes are emerging in research communities for proof of data possession and proof of data retrievability. Unfortunately, simple and inexpensive techniques along these lines have not yet found their way from theory to practice. Perhaps more useful are the efforts cloud providers make to ensure their own data storage is recoverable. It may well be that, on average, cloud providers' systems are better administered than the information technology groups of many organizations and agencies. At least, cloud users certainly hope so! Nevertheless, various risks remain.

Another old problem that has been exacerbated involves the ability to delete information ubiquitously. The existence of pervasive copies and different versions has clearly exploded as a result of copies that have been replicated for resilience. Internet mirrors have proliferated far beyond anyone's ability to keep track of unsearchable versions. With storage in some unaccountable remote repository, pervasive deletion will always seem to be questionable. Besides, approaches that may succeed in pervasive deletion may also be victimized by accidental or malicious deletion. In this case, some sort of time machine would be desirable.

Many socially relevant risks also need to be considered, such as different versions of unauthentic data; the presence of misinformation in not quite identical searchable versions of what purports to be the same information; and situations in which people or organizations desire that certain information disappear completely.

Of course, international laws and regulations also present numerous problems—first by their imprecision


or overextension, and second by the uncertainty surrounding the origins and destinations of data and other resource requests. For example, if a nation insists that all information belonging to its citizens must be stored within systems under its own legal jurisdiction, how can that be assured when it is so easy to subvert, and when ownership is itself murky? In addition, we must be cognizant of the risks of ubiquitous surveillance in unaccountable and in some cases unknown remote resources.

As noted in many past Inside Risks columns (this is the 234th in the series), almost every computer or human entity is potentially untrustworthy, with respect to accidents, intentional misuse, and attacks. As an example that remains problematic, the outsourcing of elections with regard to dependence on proprietary systems and software, computing resources, registration databases, networks (whether open or private), and—above all—dependence on potentially untrustworthy people, aptly illustrates the end-to-end nature of the risks from the very beginning of the election cycle to the disputes that result from sources of error, fraud, and confusion—with concomitant fear, uncertainty, and doubt. Insider misuse is serious in all shared resources, but particularly in elections (numerous cases have been noted in the Risks Forum and elsewhere). In this example, outsourcing to unaccountable entities is problematic.

Despite the risks discussed here, there are some hopes for constructive alternatives. Research communities have various approaches to pieces of this puzzle, but rarely to systems as a whole. As a result, many of the previous columns in this series are relevant to the use or misuse of remote resources—even if they focused on problems that were previously considered as local. For example, cryptography that is managed solely by end users for information stored remotely in encrypted forms is often touted as a solution to the problem of having to trust an untrustworthy remote storage provider. Homomorphic cryptography has the potential to allow computations on encrypted information, without the need for that information to be decrypted. These approaches can improve the confidentiality of the information,

and also provide a means for sharing the information through out-of-band shared cryptographic keys. However, they remain vulnerable to other compromises such as accidental or malicious deletion, lapse of contracts with remote providers, loss of cryptographic keys, unavailability of servers, invasive usage monitoring, and so on. As is true in general, key management becomes a fundamental risk in itself. Furthermore, convenient schemes for recovery of lost keys (for example, backdoors) are always vulnerable to misuse—as are any backdoors that can be misused by insiders or external attacks.

Virgil Gligor¹ has considered some of the risks inherent in virtualization in a context very similar to what is examined in this column. Virtualization has certain aspects that are common to the abstractions provided by remote execution and remote access, in the sense that there are well-defined interfaces for dealing with both cases—whether they are virtually remote or physically remote. There are also questions of the trustworthiness of the underlying mechanisms for enforcing the virtualization abstractions—for example, encapsulating, avoiding, or otherwise masking lower-layer vulnerabilities. Gligor's article implicitly addresses some of the topics noted here, and deserves a careful reading for those readers who would like further background than that included here.

I emphasize that clouds can offer real and significant benefits. They also bring many risks, which can be masked by the simplicity of the cloud abstraction. You should weigh these risks when designing, selecting, and configuring your cloud services. 

References

1. Gligor, V. Security limitations of virtualization and how to overcome them. Security Protocols Workshop, SPW 2010, Cambridge, U.K., 2010.
2. McMillan, R. Cloud computing a security nightmare, says Cisco CEO. *Computerworld*; http://www.computerworld.com/s/article/9131998/Cloud_computing_a_security_nightmare_says_Cisco_CEO.

Peter G. Neumann (neumann@csl.sri.com) is Senior Principal Scientist in the Computer Science Lab at SRI International, and moderator of the ACM Risks Forum.

The author is enormously grateful to the members of the ACM Committee on Computers and Public Policy for their continued wisdom and counsel in acting as an advisory group for risks-related activities. This column gained significantly from their feedback.

Copyright held by author.