Chapter 4:
Type 1 Cloud:
Storage Cloud

How Storage Clouds Work

Type 1 Clouds consist of huge amounts of memory accessible by authorized remote users. As described in the previous chapter, large-scale storage clouds are usually large buildings with lots of electrical power, huge numbers of commodity hard drives for memory, and high-speed access to the cloud.

How does the Storage Cloud provide protection against data loss? Most storage clouds automatically generate back-ups of whatever is stored in them. Since the Storage Cloud service provider (CSP) is using commodity hard drives to store all this data, there is a chance that any hard drive might crash and lose the data that was stored on it. To prevent this, the CSP automatically makes a copy of the data on a second hard drive, and usually makes a copy on a third hard drive on a physically different rack.

The result of this triple redundancy is that the chance of anyone's data being lost when stored in a Type 1 Cloud is very, very low. Since this is much better redundancy than is available in most homes or even offices, there is a greater assurance that the data stored in a Storage Cloud won't be lost by accident. The bigger the cloud of available storage, the greater the chance that the data you store in a Storage Cloud will still exist even if one, or even two, hard drives crash in a short period of time.

A few examples of the many companies that provide cloud storage services include:

- Comcast's online back-up service

- eBay's online back-up service
- Amazon's Simple Storage System (S3)

Benefits of Storage Clouds

Users can place their photographs, music, videos, documents, electronic books, scanned papers, and any other data into the Storage Cloud with the intent of retrieving it later. This not only allows for protection against loss of data when a home computer crashes or is lost in a house fire, it allows users to access these remotely stored items from wherever they can access the Internet. User access to data stored in a Type 1 Cloud can occur from home, school, a friend's house, while on travel, from Mars (well, not yet Mars), or wherever the user has Internet access.

Companies and other organizations can also benefit from cloud storage. Rather than having to retain hard copies of records in a warehouse or storage shed, a company can have the documents scanned and store them in a Type 1 Cloud. This will reduce storage costs, and increase the chance of the survival of the information in case of disaster. For example, a company could include a Storage Cloud as part of their disaster recovery plan, since the data could be accessed by authorized personnel from a new location after a fire, flood, tornado or earthquake.

If you are a PM on a project that has critical information that needs to be retained, or if you are running a distributed project where geographically remote personnel need to share information, a Storage Cloud can be useful as well. As long as access is properly controlled, this approach to sharing information can be very useful to a project manager.

The CSPs also benefit from the use of redundancy within a Storage Cloud. For example, let's say that 10,000 users are all storing the same popular song. The specific recording of that song will have a unique electronic signature (called a hash) that uniquely identifies it from all other songs or data files. Rather than storing 10,000 copies of this song (or 30,000 copies when including triple redundancy), the cloud storage provider need only store a dozen or so copies of the song. As long as all of the users who stored that song in the Storage Cloud can retrieve that song on demand, it doesn't matter whether it is one of a dozen or one of 10,000. Therefore, the

economies of scale increase for a CSP as long as what is being stored by one person is identical to what is being stored by another person.

Note that while using the hash to retain fewer redundant copies of the same file works well for songs, there is little overlap between things like individual photographs. One person's "happy snap" is unlikely to be identical to (or have the same hash as) another person's photograph. So unless the same photo is being stored by many people, there is unlikely to be any comparable economies of scale for the CSP. However, since their cost model is based on storing every user's items uniquely, the CSPs will still make a profit.

Limitations and Issues of Storage Clouds

While a Storage Cloud is excellent for data storage and retrieval, that's all it does. One cannot place data into a Storage Cloud and then run programs or applications on that data while it is still in a cloud. No computational services, such as word processing or spreadsheet software, are available on a storage cloud. There are no opportunities to perform Big Data analysis within a Storage Cloud. A Type 1 Cloud is simply designed to perform receipt, redundant storage, and on-demand retrieval of what is stored within it.

So if you are a project manager and your boss wants you to run Big Data analysis on the company's archives within the Storage Cloud, you can (kindly) inform your boss that such a function is not available in that type of cloud. The company's archives would first have to be transferred to a Type 3 Data-focused Cloud in order to perform that analysis.

Security of the data stored in the cloud is an issue for all three types of clouds. This issue is addressed more thoroughly in Chapter 7. For now, we simply state that if the data you plan to store in a Storage Cloud is sensitive (such as personally identifiable information or PII), then you probably want to take the following steps:

1. Ensure an access control mechanism with strong authentication is in place so that only the personnel you want access to the data are on the access list
2. Ensure that the data being sent to storage or retrieved is adequately encrypted in transit so no one can read it coming or going

3. Ensure that the data is adequately encrypted while being stored.

(Adequate encryption means that the size of the encryption key is sufficiently large, the encryption algorithms strong and well-defined, and that the underlying key management system is not easily compromised. A project's or organization's supporting security personnel should be familiar with, and be able to provide, these encryption features.)

Regarding encryption "at rest" (that is, while it is in storage), it makes sense to encrypt PII or company proprietary information. Other information that is publically available, such as company flyers, logos, or maps to facilities, do not need to be encrypted (unless it is a master copy that needs to be retained and unchangeable).

Another issue to consider is what happens when there is a power outage in the Storage Cloud's region and not where the users are located. In the past, if a company had a power outage and no generator backup was available, the company just made do until power was back up. Now, with the company's data potentially stored at a geographically remote location at the CSP, what does the company do when it has power and is ready to do business, but the access to its stored files is not available? There is no easy answer to this problem, which we will discuss further in Chapter 8 on arranging to use a cloud and service level agreements (SLAs).

Storage Clouds and the PMBOK® Guide

The PMBOK® Guide describes the process groups and the knowledge areas performed by the project manager. This section will discuss where Type 1 Clouds might be involved in the five process groups and the ten knowledge areas of project management. Note that if the PM's organization already regularly uses a Storage Cloud, the organization will probably already have standard operating procedures (SOPs) covering many of these topics.

Process Groups:

Initiating Process Group:

During project initiation, the PM needs to determine whether a Storage Cloud will be used to support the project. The use of a

Storage Cloud may be mandated by the organization, or the use may be selected by the PM. Whatever the cause, a succinct description of why a Storage Cloud will be used and how it will be used is essential. The PM should also ensure that if a Type 1 Cloud is selected to support a project, that it is the most appropriate of the three types of clouds to support that project.

Planning Process Group:

If a Storage Cloud has been selected for use in this project, then a storage CSP will need to be selected during planning (assuming the organization does not already have a standing agreement with a storage CSP). If the use of a Storage Cloud to support a project is an innovation, then how the Storage Cloud will be used to support the existing or new business model will need to be described. If the use of a Storage Cloud is common at the PM's organization, then simply refer to or modify the existing SOP.

For example, if the Storage Cloud will be used as a source of data as part of a disaster recovery program, then how the data will be accessed by whom and from where should all be included in the project plan.

Further planning will be required for determining whether and how which types of the data will be encrypted in transit and/or at rest on the storage cloud. Creating the access control list for who from each location has access to which data will also be important.

Risk planning will include assessing the risk of a data breach, and what to do if the data is temporarily unavailable from the CSP (such as due to a power failure at the CSP site).

Executing Process Group:

During project execution, the PM or his designee will need to set up the access control and determine who on the team has which access to the Storage Cloud and why. The purpose behind the use of the Storage Cloud will determine who gets access, can send data to storage, and retrieve or delete data from storage.

Monitoring and Controlling Process Group:

The PM needs to monitor (or have someone monitor) the access control list to the data stored in the cloud to ensure it is current. For example, if someone leaves the company, you may have

retrieved the key to the building and removed the former employee's access to the local network, but don't forget to remove that person's access to what is stored in the cloud. Other features to monitor are whether the use of the cloud for storage are meeting the expected benefits, or whether there are problems with the efficiency of the system. Making sure the CSP meets the criteria for availability as described in the service level agreement will also be important. It is also good business practice to have your organization's penetration testers periodically check that the data being transmitted and at rest are actually encrypted.

Closing Process Group:

If the use of the Storage Cloud will continue beyond the end of the project, then the PM needs to ensure the access control is correctly closed out. If the Storage Cloud will no longer be used at the end of the project, then the final disposition of the data in the Storage Cloud will need to be addressed.

Knowledge Areas:

Stakeholder Management Knowledge Area:

The PM will need to determine which team members and other stakeholders will need access to the data stored in the cloud. How often and how quickly they will need to access the data will define the requirements for responsiveness and availability. Business needs and constraints on who can access sensitive information will determine which stakeholders get access to which data.

Communications Knowledge Area:

Communications to stakeholders should include how the Storage Cloud will be used to support the project. In addition, the PM needs to communicate the plan for how secure communications of data to and from the cloud will be ensured and monitored during the project. Communications with the CSP management will also be essential, such as when issues occur that need to be resolved.

Risk Knowledge Area:

Before selecting a Storage Cloud, list and evaluate the risks of using a Storage Cloud. Identify the risks, perform risk analysis, and plan risk responses. After selecting a Storage Cloud, monitor the communications and data stored in the cloud for any changes in the risk factors.

Procurement Knowledge Area:

Procurement of a Storage Cloud service level agreement (SLA) is a prerequisite to using a Storage Cloud. If the PM's organization already has a standing arrangement with a Storage Cloud CSP, this step is straightforward.

Cost Knowledge Area:

The use of a Storage Cloud will entail a cost that must be planned and managed. While storing data in a Storage Cloud is relatively inexpensive compared to physically storing hard copies of data, there are monthly costs that need to be estimated, budgeted, and monitored throughout the project life cycle.

Integration Knowledge Area:

The use of a Storage Cloud must be integrated into the organization's business model. How will it be used? What will it be used for? Who will have access to the data and why? How long will the data be stored before being aged off? If the Storage Cloud will be used as part of a continuity of operations plan, how will it be used and how will you train for it?

Scope Knowledge Area:

How will the Storage Cloud be used within your business model? How much data will be stored in the Storage Cloud each month? If the scope of use changes over time, how will that scope creep be managed?

Time Knowledge Area:

When will operations be transitioned to start using a Storage Cloud? Or for a particular project, when will the project start using the Storage Cloud as part of its operations? Is the cloud service provider on-board with the transition time, magnitude of data stored and retrieved each month?

Quality Knowledge Area:

What is the data retrieval time, and does this match the service level agreement (SLA)? How is the quality of service measured according to the SLA? Who will monitor the performance of the Storage Cloud provider and how?

Human Resources Knowledge Area:

If the Storage Cloud is storing personally identifiable information (PII), then the access control for PII stored in the cloud needs to be properly instituted and managed.