



**TACTICS FOR ESB IMPLEMENTATION**  
BEST PRACTICE SERIES

[LEARN MORE](#)



## What Is Cloud Appropriate Data?

Posted by George Crump @ 09:20:AM | Jan,14, 2010



As the use of cloud storage becomes more prevalent this year one of the key challenges is what data should go to cloud storage, when that data should be moved there and how should that data be moved there?

The answer to those challenges in part is going to depend on how you are using the cloud in the first place. If you are using the cloud to utilize software as a service your data essentially is already in the cloud. Your challenge is to get your data out of the cloud, making sure that you have local copies of the information that is being kept out there. This will be the subject of an upcoming entry 'Getting Data Out of The Cloud'.

For many, especially early deployments, cloud storage is going to be used for either backup data or archived data. For small businesses and individuals their entire data set may be backed up to the cloud. Using technologies from companies like Axcient and [DS3 DataVaulting](#) an organization can keep the most recent copy of data local and then replicate that data off-site to a cloud storage facility in case of a disaster. Look for most major backup vendors to add cloud capabilities in the near future if they haven't already.

When it comes to archive the need to recover is more expected. This generally is not because something has gone wrong but more because old data is needed for new research or to meet a legal disaster recovery request. Here is where some intelligence has to be applied to what data is transferred to the cloud. Most cloud providers like Iron Mountain and Nirvanix will provide a NAS gateway to help you move data in and out of the cloud. This provides you with a CIFS or NFS mount view of the cloud locally and then does a more internet friendly protocol translation when moving the data to the provider. Once it is

moved however it is moved and is probably not available locally, so search and retrieval times need to be adjusted.

Through software analysis provided by companies like Tek-Tools or APTARE you can dissect data by size, access time and other parameters. File virtualization systems like those from AutoVirt and F5 can provide much of the same analytics and move the data to the gateway for you, setting up a transparent link. Knowing more than just an access date is critical. Size of the file for example, may affect how aggressive you are with the transfer to the cloud. Small files you may want to archive right away since there is limited bandwidth needed to transfer them. Large files you may want to wait a little extra time to move them, to make sure they are good and cold first.

Then there is the issue of finding the data. With these providers building vast storage repositories how are you going to find the data you need in a timely manor? That's the subject of our next entry: 'Indexing the Cloud'.

*George Crump is lead analyst of Storage Switzerland, an IT analyst firm focused on the storage and virtualization segments. Find Storage Switzerland's disclosure statement [here](#).*