Backup vs. Archive
A Practical Approach

Kelly J. Lipp

Storage Solutions Specialists, Inc.

February 24, 2005

Agenda

STORAGE SOLUTIONS SPECIALISTS

- Definitions
- Backup as an Archive Tool
- Why Archive
- A Practical Definition
- Steps to Establishing Archive as an IT function

Definitions - Archive

Pronunciation: 'är-"klv

Function: noun

Etymology: French & Latin; French, from Latin archivum, from Greek archeion government house (in plural, official documents), from archE rule, government -- more at ARCH.

Date: 1603

: a place in which public records or historical documents are preserved; also : the material preserved -- often used in plural, ex: archives.

Definitions - Backup

Pronunciation: -"&p

Function: noun

Usage: often attributive

Date: 1951

1 a : one that serves as a substitute or support <a backup plan> b : musical accompaniment

2 : an accumulation caused by a stoppage in the flow <traffic backup>

3 : the act or an instance of backing up a computer's hard disk

Definitions

- From WWWebster Dictionary Copyright© 1996 by Merriam-Webster, Incorporated
- If Merriam-Webster had meant for these two things to be the same, she wouldn't have created two words!

Archive

STORAGE SOLUTIONS SPECIALISTS

- To remove from the on-line system those objects no longer in day-to-day use, and place them into long term, retrievable, storage.
- Generally, not all objects.
- Why? Provide the ability to retrieve specific information from the past.

Backup

STORAGE SOLUTIONS SPECIALISTS

- To make a copy of an object in the event the original becomes lost or damaged.
- Generally, all objects.
- Why? Provide the ability to restore any information to the most recent state possible.

Backup

STORAGE SOLUTIONS SPECIALISTS

- Contemplates two instances:
  - catastrophic loss of data, due to hardware failure
  - the inadvertent loss of data, e.g., a user accidentally deleting a file.
- In both cases, the most recently backed up version of the data is desired.
- Rarely is it required to restore data from more than a couple of days in the past.

In fact…

- What we really want is the ability to restore data to a point right before we screwed it up!
- Is it possible to backup data more often than once per day?
- Might it be desirable to provide this as a service?
- Hmmm…
- But we digress…

Archive

STORAGE SOLUTIONS SPECIALISTS

- Contemplates a single scenario:
  - The retrieval of specific information for historical purposes.
- Example:
  - The state labor department is interested in seeing payroll records for a specific period of time.
  - The labor department isn't interested in recent backup data, but rather older data in archive form.
  - The labor department is probably not interested in the electronic mail or other files surrounding the payroll data.

Definition

STORAGE SOLUTIONS SPECIALISTS

Backup is Short Term,
Archive is Long Term

Why is this hard?

- The tools we've had to work with only do one thing: Backup.
- Difficult to be discrete with our data:
  - Treat some as backup.
  - Treat some as archive.

So, What do we do?

STORAGE SOLUTIONS SPECIALISTS

Save periodic full backups as archives!

Is this a good idea?

STORAGE SOLUTIONS SPECIALISTS

- **NO!**
  - **Too much data**
  - **Too much of the wrong data: operating system files, applications files, etc.**
  - **Too much of the kind of data that can get you in real trouble:**
    see United States government vs. Microsoft.

Besides…

- Were we told to save all data?
  - In some cases yes. But this is out of ignorance.
- Generally, the regulations specify which data are to be kept and for how long.

So, Why do we do it?

STORAGE SOLUTIONS SPECIALISTS

- It's easy
- Save everything. Then you're bound to have what you need.
- How in world will you find it.
- See needle in the haystack.

Ideally…

- We only save the data we are required to save.
- It's not mixed with all of our backup data.
- It's in Archive Form so it's easy to use.

What is Archive Form?

STORAGE SOLUTIONS SPECIALISTS

- Binary representation of the data?
- The data plus the application?
- The data plus the application plus the operating system?
- The data plus the application plus the operating system plus the hardware it originally ran on?
- None of the above?

Archive Form

- Textual representation of the data.
- Schematic of the data.
- Easily readable by many applications.

Question

STORAGE SOLUTIONS SPECIALISTS

Does it make sense to keep a tape around for ten years?

Archive: A Practical Definition

STORAGE SOLUTIONS SPECIALISTS

- Verb: To place specific data into an easily readable, easily restorable long term storage area.
- Noun: Specific data maintained for legal or business reasons for a long period of time in an easily readable format.

How much data is archived?

- Varies by business.
- Usually much less than the total.
- Only that which is required for legal or other specific business reasons.

So It's Easy, Right!

- Not!
- Hard to find and isolate the data.
- Hard to determine how long to keep which data.
- Hard to agree on a form to store the data.
- Requires cooperation among business and IT folks.

Why do it?

STORAGE SOLUTIONS SPECIALISTS

- Cost: Tapes, Libraries, Server.
- Time: Not enough time to perform a “full” backup.
- Legal reasons:
  - You will have data you are required to have
  - You won’t have data that you don’t want to have

How to Establish Archive in your Business

Step 1:

STORAGE SOLUTIONS SPECIALISTS

- Pick a tool that offers archival services as a separate function from backup.
- TSM comes to mind!

Step 2:

STORAGE SOLUTIONS SPECIALISTS

- Determine which data actually requires archival. Use your legal department along with your business users of the data to hone in on this data.
- Use an IT sponsored workshop to gather business and IT folks to determine your archive requirements.

Step 3:

- Determine the archival retention periods for the various types of data in your environment. Not all data has the same requirements.
- Get your legal people involved.

Step 4:

- Determine the format of the archived data. Remember, systems, applications and tools change, so pick a format that will work with many tools now and in the future.
- Consider simple textual forms.

Step 5:

- Design a solution, using the archival tool, which meets your business needs.
- Using TSM, implement various Archive Copy Groups to meet your business needs.

Step 6:

STORAGE SOLUTIONS SPECIALISTS

- Archive your business critical data using the tool.
- Create TSM schedules or OS scripts to archive the data.

Step 7:

- Perform a test restore of information into a second application to verify the efficacy of your solution.
- Can Microsoft Excel read your data?

Summary

STORAGE SOLUTIONS SPECIALISTS

- Archive and Backup are two distinctly different things.
- An effective Archive strategy can save time and money.
- Data requiring archive services is much smaller than all of your data.
- Archiving is fun!