BACKUP AND RESTORE

BACKUP AND RESTORE

- Being able to restore lost data is a critical part of any environment
- Data also gets corrupted by mistake, on purpose , and by gamma rays from space.
- Backups are like insurance: You pay for it even though you hope to never need it. In reality, you need it.
- Backup and restore service is part of any data storage system.

Full backup- a complete backup of files on a partition

Unix users call this" **level 0 backup**"

Incremental backup- refers to copying of all files that have changed since the full backup.

Unix users call this "**level 1 backup**" or simply as incrementals

INCREMENTAL BACKUP

- Incremental backups grow over time.

  Example:

  A full backup is performed on Sunday and an incremental backup each day of the week follows, the amount of data being backed up should grow each day because Tuesday's incremental backup includes all the files from Monday's backup, as well as what changed since then. Friday's incremental backup should include all the files that were part of Monday's, Tuesday's, Wednesday's, and Thursday's backups, in addition to what changed since Thursday's backup.

- Some systems perform an incremental backup that collects all files changed since a particular incremental backup rather than the last full backup.

BACKUP AND RESTORE SYSTEM

- Engineering your backup and restore system should begin by determining the desired end result and working backward from there.
- Restores are requested for various reasons, and the reasons that apply to your environment affect further decisions, such as development of a policy and a schedule.
- We start by defining **corporate guidelines**, which drive your **SLA** for restores based on your site's needs, which becomes your backup **policy**, which dictates your backup **schedule**.

CONT. . .

- The **corporate guidelines** define terminology and dictate minimums and requirements for data-recovery systems.
- The **SLA** defines the requirements for a particular site or application and is guided by the corporate guidelines.
- The **policy** documents the implementation of the SLA in general terms, written in English.
- The **procedure** outlines how the policy is to be implemented.
- The detailed **schedule** shows which disk will be backed up when. This may be static or dynamic. Such a schedule usually consists of the policy translated from English into the backup software's configuration.

CONT. . .

- Beyond policies and schedules are operational issues:
  - Consumables can be expensive and should be included in the budget.
  - Time and capacity planning are required to ensure that we meet our SLA during both restores and backups.
  - The backup and restore policies and procedures should be documented from both the customer and the SA perspectives.

- Modern backup systems have **three key components**:
  - Automation
  - Centralization
  - Inventory management

REASONS FOR RESTORES

- **Accidental file deletion**
  - ❖ A customer has accidentally erased one or more files and needs to have them restored.
- **Disk failure**
  - ❖ A hard drive has failed, and all data needs to be restored.
- **Archival**
  - ❖ For business reasons, a snapshot of the entire "world" needs to be made on a regular basis for disaster-recovery, legal, or fiduciary reasons.

ACCIDENTAL FILE DELETION

- When the data is on a SAN or NAS with the snapshot feature, it is usually possible to perform a self-service restore with reasonable time granularity.
- The customer may lose up to an hour of changes, but that is better than a day's worth, and the data can be recovered quickly, without losing additional time waiting for an SA to perform the restore.
- Without SAN or NAS, you can typically expect to be able to restore a file to what it looked like at anyone-day granularity and to have it take three to five hours to have the restore completed.

SNAPSHOTS

- To an SA, the value of **snapshots** is that they reduce workload, because the most common type of request becomes self-service.
- To customers, the value of **snapshots** is that they give them new options for managing their work better. Customers' work habits change as they learn they can rely on snapshots.
- Snapshots also increase customer productivity by reducing the amount of lost data that must be manually reconstructed.
- Customers are less likely to attempt to manually reconstruct lost data.

DISK FAILURE

- A disk failure causes two problems:
  - ❖loss of service
  - ❖ loss of data
- On critical systems, such as e-commerce and financial systems, RAID should be deployed so that disk failures do not affect service, with the possible exception of a loss in performance.
- This kind of restore often takes a long time to complete. Restore speed is slow because large volumes of data are being restored, and the entire volume of data is unavailable until the last byte is written.
- To make matters worse , a two-step process is involved : First the most recent full backup must be read, and then the most recent incremental(s) are read.

ARCHIVAL PURPOSES

- Corporate policies may require you to be able to reproduce the entire environment with a granularity of a quarter, half, or full year in case of disasters or lawsuits.
- The work that needs to be done to create an archive is similar to the full backups required for other purposes, with the following differences:
  - ✓Archives are full backups. In environments that usually mix full and incremental backups on the same tapes, archive tapes should not be so mixed.
  - ✓Some sites require archive tapes to be separate from the other backups
  - ✓Archives are usually stored off-site.

CONT. . .

- ✓Archive tapes age more than other tapes. They may be written on media that will become obsolete and eventually unavailable. You might consider storing a compatible tape drive or two with your archives, as well as appropriate software for reading the tapes.
- ✓ If the archives are part of a disaster-recovery plan, special policies or laws may apply.

- When making archival backups,
  - ➢Do not forget to include the tools that go with the data.
  - ➢ Make sure that the tools required to restore the archive and the required documentation are stored with the archive.

PERFORM FIRE DRILLS

- The only time you know the quality of your backup media is when you are doing a restore.
- This is generally the worst time to learn that you have problems.
- You can better assess your backup system if you do an occasional fire drill.
- Pick a random file and restore it from tape to verify that your process is working.
- When doing these fire drills, it is important to time them and monitor such things as disk,tape, and network utilization.

CORPORATE GUIDELINES

- Organizations need a corporate-wide document that defines terminology and dictates requirements for data-recovery systems.
- The guideline should begin by defining why backups are required, what constitutes a backup, and which kind of data should be backed up.
- A set of retention guidelines should be clearly spelled out. There should be different SLAs for each type of data: finance, mission critical, project, general home directory data, email, experimental, and so on.

- For example, the guidelines should require sites to carefully plan when backups are done, not simply do them at the default "midnight until they complete" time frame. It wouldn't be appropriate to dictate the same window for all systems. Backups usually have a performance impact, so they should be done during off-peak times. E-commerce sites with a global customer base will have very different backup windows than offices with normal business schedules.

A DATA-RECOVERY SLA AND POLICY

- The next step is to determine the service level that's right for your particular site.
- An SLA is a written document that specifies what kind of service and performance that service providers commit to providing.
- To establish an SLA, list the three types of restores, along with the desired time to restoration, the granularity and retention period for such backups (that is, how often the backups should be performed and how long the tapes should be retained), and the window of time during which the backups may be performed (for example, midnight to 8 AM).

- The example SLA as follows:

Customers should be able to get back any file with a granularity of one business day for the past six months and with a granularity of one month for the last three years. Disk failures should be restored in four hours, with no more than two business days of lost data. Archives should be full backups on separate tapes generated quarterly and kept forever. Critical data will be stored on a system that retains user-accessible snapshots made every hour from 7 AM until 7 PM, with midnight snapshots held for one week.

Databases and financial systems should have higher requirements that should be determined by the application's requirements and therefore are not within the scope of this example policy. The policy based on this SLA would indicate that there will be daily backups and that the tapes will be retained as specified. The policy can determine how often full versus incremental backups will be performed.

BACKUP SCHEDULE

- Backups should be performed every business day.
- Even if the company experiences a nonredundant disk failure and the last day's backups failed, we will not lose more than two days' worth of data. Since full backups take significantly longer than incremental backups, we schedule them for Friday night and let them run all weekend. On Sunday through Thursday nights, incremental backups are performed.

TIME AND CAPACITY PLANNING

- Restores and backups are constrained by time.
- Restores need to happen within the time permitted by the SLA of the service.
- Most systems slow down considerably when backups are being performed.
- Some services must be shut down entirely during backups.

- **Backup Speed -** if the server cannot provide data quickly enough, backup speed is dramatically reduced.
- **Restore Speed -** Restoring an entire disk is extremely slow, too. The main issue affecting the restore speed is not the drive read speed, but rather the file system write speed.
- **High-Availability Databases** –Some applications, such as databases, have specific requirements for ensuring that a backup is successful. A database manages its own storage space and optimizes it for particular kinds of access to its complex set of data tables. it is not acceptable to shut it down each night for backups. However, the risks associated with not doing a backup or performing the backup while the database is live are also unacceptable.

SUMMARY

- There are three kinds of restore requests: **accidental file deletion**, **recovery from disk failure**, and **archival**.
- The backup and restore policy is set based on these parameters.
- The policy should also state that the validity of backups must be tested with fire drills.
- Once a backup and restore policy is in place , all decisions flow easily. From the policy, you can develop a backup schedule that is a specific list of which systems are backed up and when.
- One of the most difficult parts of determining this schedule is deciding how many days of incremental backups should be done before the next full backup.

- Current software does these calculations and can create a highly dynamic schedule.
- The policy helps you plan time, capacity, consumables, and other issues.
- Communicating the policy to the customers helps them understand the safety of their data, which systems are not backed up, and the procedure they should use if they need data restored. Making customers aware of which systems are not backed up is important as well.

- A modern backup system must be automated to minimize human labor, human thought, human decisions, and human mistakes.
- Modern backup systems are centralized. Doing backups over the network to a central, large backup device saves labor.
- Backup technology is changing all the time. As disk capacity grows, SA's must upgrade their ability to maintain backups.
- Restores are one of the most important services you provide to your customers.
- The inability to restore critical data can bankrupt your company. The flawless execution of a restore can make you a hero to all.