Database Backup and Recovery Best Practices

Ali Navid Akhtar, OCP, has more than two decades of experience with databases. He works as a lead database administrator at Solo Cup Co.

Jeff Buchholtz has more than 18 years of design, implementation and support of global IT technology solutions. He works in an IT leadership role and is an Oracle database administrator.

Michael Ryan, CIA, CPA, is the director of internal audit for Solo Cup Co., with the primary responsibility of building and executing US Sarbanes-Oxley Act 404 compliance strategies.

Kumar Setty, CISA, has more than 10 years of experience in the areas of data analysis, systems administration, auditing and computer security. He is a manager at PricewaterhouseCoopers LLP.

The ability to restore databases from valid backups is a vital part of ensuring business continuity. Backup integrity and restorations are an important piece of the IT Governance Institute's *IT Control Objectives for Sarbanes-Oxley, 2nd Edition*. In many instances, IT auditors merely confirm whether backups are being performed either to disk or to tape, without considering the integrity or viability of the backup media.

This article covers the topics related to data loss and the types of database backup and recovery available. Best practices that can assist an auditor in assessing the effectiveness of database backup and recovery are also provided. This article focuses on the technologies and capabilities of the Oracle relational database management system (RDBMS) and Microsoft (MS) SQL Server because, together, they cover approximately 40 percent of all database installations. **Figure 1** provides a short comparison of Oracle and MS SQL Server.

One of the key responsibilities of a database administrator (DBA) is to prepare for the possibility of media, hardware and software failure as well as to recover databases during a disaster. Should any of these failures occur, the major objective is to ensure that the database is available to users within an acceptable time period, while ensuring that there is no loss of data. DBAs should evaluate their preparedness to respond effectively to such situations by answering the following questions:

- How confident is the DBA that the data on which the company business depends are backed up successfully and that the data can be recovered from these backups within the permissible time limits, per a service level agreement (SLA) or recovery time objective, as specified in the organization's disaster recovery plan?
- Has the DBA taken measures to draft and test the procedures to protect as well as recover the databases from numerous types of failures?

The following is a checklist for database backup and recovery procedures that are explained throughout this article:

1. Develop a comprehensive backup plan.
2. Perform effective backup management.
3. Perform periodic databases restore testing.
4. Have backup and recovery SLAs drafted and communicated to all stakeholders.
5. Have the disaster recovery plan (DRP) database portion drafted and documented.
6. Keep your knowledge and know-how on database and OS backup and recovery tools up to date.

COMPREHENSIVE BACKUP PLAN

DBAs are responsible for making a comprehensive backup plan for databases for which they are accountable. The backup plan should include all types of RDBMSs within the enterprise and should cover the following areas:

- **Decide what needs to be backed up.** It is imperative that the DBA be aware of database and related OS and application components that need to be backed up, whether via an online backup or an offline cold backup. The following are details of what needs to be backed up:
  - OS software—An event such as a hardware failure will require a complete system restore, starting with the OS, so there is a need to back up the database server OS initially and after any system updates or configuration changes.
  - RDBMS software—The RDBMS software should be backed up initially and after any patches/upgrades.
  - Application software where applicable—This applies especially to Oracle E-Business Suite, Oracle Application Server and Oracle Enterprise Manager (OEM). The application DBA should complete an initial full backup of the applications to disk using an appropriate OS command and, then, schedule future incremental backups, e.g., after any patches/upgrades. These backups should also be transferred to tape.

Figure 1—Comparison of Oracle and MS SQL Server		
Item	**Oracle RDBMS**	**MS SQL Server RDBMS**
General	In Oracle, a database when started refers to the entire Oracle RDBMS environment, including memory structures and background processes called Oracle instance and control files, datafiles, online redo logs and some other files, such as the parameter or server parameter file and the password file.	An instance of SQL Server when executed allocates memory pools, uses background processes, and has multiple databases including system and user databases. The master database is the main system database that contains the system catalog as well as some information about individual databases.
Catalogs	Each Oracle database runs on one centralized system catalog, or data dictionary, which resides in the SYSTEM tablespace.	In SQL Server, the system catalog, which is analogous to the Oracle data dictionary, is broken up among the individual databases, the master database, and the (hidden and read-only) resource database (found in later versions).
Storage structures	The Oracle RDBMS is comprised of logical structures called tablespaces, which, in turn, are comprised of physical datafiles. Tablespaces/datafiles are formatted into internal units, called blocks. An Oracle extent contains a chain of contiguous blocks and varies in size.	SQL Server uses filegroups, which are logical containers of one or more files. Data contained within a filegroup is proportionally filled across all files belonging to the filegroup. SQL Server formats files into internal units called pages, which are organized into extents that are fixed in size.
Logins	Oracle provides logins for authorized users to connect to the database, which are referred to as the user or username, and any operation the user can perform is controlled by the privileges granted to the login.	In SQL Server, the login enables a user to connect to an instance. However, access to other databases within the instance is not automatic and is controlled by additional accounts (called users) that are created in each of the databases to which the login requires access. The privileges at the instance level are assigned to the login, and privileges inside a database are given to the related database user. A database user is mapped back to an instance login.
Authentication	Authentication is the process of verifying that the login ID or username supplied by a user to connect to the database belongs to an authorized user. Oracle allows authentication through the OS or through the database (server).	SQL Server also allows authentication through the OS or through the database (server). In SQL Server, the OS mode is called Windows Authentication, and the database mode is called SQL Server Authentication.
Logging mode	Online redo logs are used by Oracle to record transactional changes made to the database before those changes are committed to the database files. Oracle also uses rollback or undo segments to capture an image of data before they are changed to facilitate transaction rollback, recovery and read consistency.	In SQL Server, the redo logs are called transaction logs. A transaction log combines the functionality of Oracle redo logs and the rollback or undo segments. Each database in SQL Server has one or more transaction log files.
Automatic recovery	Oracle performs automatic recovery each time it is started. It verifies that the contents of the datafiles are coordinated with the contents of the online redo log files. If they are not, Oracle applies the contents of the online redo log files to the datafiles, and then removes any uncommitted transactions that are found in the rollback or undo segments. If Oracle cannot obtain the information it requires from the online redo log files, it consults the archived redo log files.	SQL Server also performs automatic data recovery by checking each database in the system each time it is started. It first checks the master database, and then launches threads to recover all of the other databases in the system. For each SQL Server database, the automatic recovery mechanism checks the transaction log for any committed and uncommitted transactions and applies these to the database. Each database has its own transaction log, which records all changes to the database.
Backup and recovery	In Oracle, backup methods can be categorized as physical and logical. There are two ways to perform Oracle physical backup and recovery: Recovery Manager (RMAN) and user-managed backup and recovery. Oracle segments its backups by consistent and inconsistent states. These can also be viewed as cold or hot backups.	SQL Server offers full, differential, partial and transaction log backups, which aid in complete recovery of databases during disk, server or instance failure. There are a variety of hot and cold backups available in SQL Server to suit any business environment. SQL Server databases can also be quickly detached and the files copied, and then they can be attached to another instance.
Logical backups	The goal of a logical backup is to be able to recover at the individual schema object level. In Oracle, logical backups are mainly performed using the Export or Data Pump utility. This utility exports the schema objects into a binary file, which can be read only by the Import or Data Pump utility, and imports the schema objects into a database.	In SQL Server, individual schema objects can be backed up to flat files in any of the supported file formats. Then flat files can be restored using tools such as the bulk copy program (bcp) utility, the Import and Export Wizard, or the SQL Server Integration Services tools.

– Passwords—All superuser passwords that may be required during recovery should be preserved. It is a good idea to ensure that the default passwords that came with the initial installation of the RDBMS are changed.
– All components of Oracle databases:
  · Database parameter file—A parameter file or server parameter file (SPFILE) defines persistent initialization parameters of a database, including information about database control files.
  · Database control file(s)—The control file stores the status of physical structure of the database. If it becomes unavailable, the database cannot operate. It is imperative that these files be backed up while backing up other components of the database. In later versions of Oracle (9i onward), the DBA can configure automatic backup of the parameter file as well as the control file to ensure that these get backed up after each backup and after any structural changes in the database.
  · Database data files—These should be backed up during cold backup as well as during online backup, using Oracle's Recovery Manager (RMAN) or, in Oracle Database versions in which RMAN was not introduced, by putting tablespaces in backup mode. The DBA should try to run all production databases in Archive log mode so that recovery to the point of failure is possible.
  · Redo log files and archived redo logs—While making a cold backup, the DBA needs to backup redo logs. When the database is running in archive log mode and doing and online backup, the DBA needs to archive redo logs manually or automatically and then back up all archive redo logs.
  · Oracle network files—It is important to back up all Oracle network files initially and after any change.
  · Password files—Password files when used should be backed up initially and after any change.
– MS SQL Server databases:
  · Back up both system and user databases.
  · Have a separate maintenance plan for system databases, i.e., master, model, msdb. Master supports only full backups; tempdb backup is not required, as it gets rebuilt during SQL Server startup.
  · Back up all user databases. Set up all user databases for full recovery model, and back up both database and transaction logs.

• **Determine the appropriate backup type to use for your data.**
– Oracle databases:
  1. Logical backups—This type of backup is performed through Oracle utilities "exp." From version 10g onward, Data Pump can also be used. The whole database, individual schemas, tables or tablespaces can be backed up. Restore is done using "imp" or Data Pump. With such backups, recovery to the point of failure is not possible.
  2. Physical offline or cold backups—The database must be shut down and a copy must be made of all essential data files and other components of the database.
  3. Physical online or hot backups—This method enables the database to be backed up while the database is up and running. The following points should be kept in mind while doing online backups:

- Either put the tablespaces in backup mode and back up the associated data files using an OS copy command, or use RMAN, a robust tool provided by Oracle for backup and recovery with version 8.x onward. Oracle adds new functionality to this tool with each version. RMAN can use the database control file to keep its catalog, or the DBA can setup schema for each database, in a separate database for RMAN catalogs.
- The DBA must review and keep in mind the RMAN compatibility matrix for the database being backed up/restored as well as the RMAN executable and RMAN Catalog Database/Schema.
- DBAs must familiarize themselves with full, incremental and differential backups and set these up using RMAN scripts. DBAs must review their RDBMS edition, e.g., incremental backups are not possible in standard editions prior to Oracle 10g. To restore/recover a database to the point of failure or a previous point in time, the DBA must put the database in archive log mode and back up all archived redo logs.
- It is important not to forget to back up the RMAN catalog at the end of each backup. DBAs can do an export backup of RMAN catalog schema.

– SQL Server databases:

1. Logical backups—In SQL Server, individual schema objects can be backed up to flat files in any of the supported file formats. Then flat files can be restored using tools such as the bcp utility, the Import and Export Wizard, or the SQL Server Integration Services tools.
2. Physical backups—It is recommended that all user databases be set up for full recovery model, and both database and transaction logs should be backed up to restore/recover the database to the point of failure. DBAs should thoroughly familiarize themselves with database recovery models and full, differential and transaction-log backups, and set these up accordingly. File or filegroup backup strategy can be used if the databases to be backed up are very large databases (VLDBs) that are partitioned among multiple files.

- **Establish a strategy for handling VLDB backups**—In Oracle, the DBA can reduce the backup window for VLDBs by allocating multiple channels and fine-tuning backups, can save disk space by using compressed backups, and can block tracking with incremental backup techniques with the latest versions. The DBA must review the version and edition of the database to confirm availability of this option. If this does not do the trick, the DBA can consider setting up split mirror backups. For SQL Server, the DBA can partition the database among multiple files and use the file or filegroup backup strategy. Also, using multiple backup devices in SQL Server allows backups to be written to all devices in parallel.
- **Establish an appropriate backup schedule and window**—It is good practice to select a backup window at a point when the lowest amount of activity affects the database so that the backup does not reduce available database server resources and slow down the database user's activity. The DBA can tune the backup window by parallelizing backups using multiple channels; however, the DBA must review the version and edition of the database to confirm availability of this option. In the vast majority of cases, it is best to set up a weekly backup cycle starting with full backups on Friday night or Saturday morning and incremental/differential backups throughout the weekdays. Archive/transaction log backups can be scheduled for every few hours, depending on the volatility of the database.
- **Decide where to store backups**—Both Oracle and MS SQL Server databases can be backed up directly to tape or disk (locally or over the network), and then the backups can be archived to tape. It is good practice to back up to disk, transfer to tape and store tapes offsite for disaster recovery (DR). The backups to disk are faster; DBAs have more control and can better monitor these and, with this method, DBAs hold two sets of backups—one on disk, the other on tape. During restore, if backups are still on disk, it will be a faster restore, reducing mean time to recover (MTTR).

- **Develop a backup retention policy**—The backup retention policy relates to both the disk and tape rotation schedule and should be decided upon based on the SLA established with the business-user community. The data owner should specify the retention period for the data. The retention period may vary from months to years, depending on local laws. Accordingly, the DBA should be deleting old backups to create space for current backups. The data retention policy should be chosen carefully, making sure that it complements the backup media subsystem retention policy and requirements for the backup recovery strategy. If not using a catalog, the DBA must ensure that the control file record keep time instance parameter matches the retention policy.

EFFECTIVE BACKUP MANAGEMENT

After making a solid backup plan and completing initial work, the DBA should properly manage backups, keeping the following points in mind:

- **Automating backups**—For Oracle, either set backups through OEM or use an OS scheduling tool, and Spool output to a log file that can be reviewed for any errors. In SQL Server, use Maintenance Plans for scheduling backups.
- **Monitoring backups**—Set up monitoring using appropriate tools so that the DBA gets an e-mail or alert through a pager or cell phone for any failed backups, which should be rerun as soon as possible.
- **Backup logs and catalogs**—Review backup logs and backup catalog information periodically for any issues. Use RMAN reporting to show backup status. For Oracle, back up the RMAN catalog database by exporting all catalog schemas periodically as well as by doing an export backup of RMAN catalog schema at the end of each backup. For SQL Server, backup system databases, especially master and msdb.
- **Database catalog maintenance**—With Oracle databases, use "delete obsolete" to remove backups that are outside the organization's retention policy. If obsolete backups are not deleted, the catalog will continue to grow and performance will become an issue. Cross-checking (cross-check backup) will check that the catalog/control file matches the physical backups.
- **Validating backups**—Validate and verify backups without doing actual restores.
- **Setting up dependencies**—When backing up to disk, archive these backups to tape as soon as backup to disk completes. Set up a process so that disk backups get transferred to tape without loss of time.

BACKUP RESTORATION TESTING

Imagine the following scenario: A flood has hit the area in which a company's headquarters resides, and the entire IT infrastructure has been damaged, but not destroyed. Before the event, the DBAs performed backups to the backup media, following all of the processes noted previously in this article, and had these stored offsite. In the enterprise's most recent IT audit, the auditor rated the backup process as "effective."

> "Backups are of no use if the IT team cannot restore the data to the system at the time of need."

The backup media from the offsite storage is retrieved and loaded. A message appears on-screen that states that the backup media are "unreadable" due to integrity issues. What could have happened?

Many things could have happened. However, it is clear that a critical step did *not* happen. The restoration from the backup media was never really tested. The control was marked as effective because a backup process was in place and being performed. In addition, no errors were ever received when the enterprise backed up to the backup media.

Backups are of no use if the IT team cannot restore the data to the system at the time of need. A DBA should formulate a detailed strategy for this task:

1. **Databases restore testing**—There should be a requirement to test database restores from disk as well as from tape backups.
2. **Validating restores where possible**—The DBA can validate and verify backups without doing actual restores. Validating backups using the "restore validate database" command will do everything except actually restore the database. This is the best method to determine if the backup is good and usable before being in a situation in which it becomes critical.
3. **Refreshing nonproduction databases from production backups**—It is good practice to periodically build nonproduction databases from production backups using appropriate backup/restore utility commands as a restore practice.
4. **Performing annual/biannual restore testing from tape as part of audit**—The DBA will have to explain the process

through a narrative, preserve logs and take screenshots to show this type of restore testing.

5. **Actual restores**—During actual restores, the DBA should back up the database before doing the restore. Depending on the type of loss and backups available, the DBA must decide on whether to go for complete (point-in-time) or incomplete recovery. Incomplete recovery can be time-based, cancel-based or change-based.
6. **Strategy to recover from database corruption**—For Oracle databases, the DBA can turn on block checking using appropriate parameters to detect the presence of corrupt blocks in the database. This has a slight performance overhead, but will allow early detection of corrupt blocks caused by underlying disk, storage system or input/output (I/O) system problems. By default, RMAN also checks for corrupt blocks during backup. In later versions of Oracle, RMAN can be used to repair corrupted blocks in the database.

BACKUP AND RECOVERY SLA

The DBA team must draft a backup and recovery SLA, covering details of backup procedures and including a timeline for recovery, and have management sign off on it. The SLA does not assist in the recovery process itself, but sets the user community's (and management's) expectations for the recovery process, which may provide the team more time to complete the restore process.

DISASTER RECOVERY PLAN

The DBA should take care to ensure that databases are included as a key element in the company's overall DRP. All stakeholders need to understand the elements of the recovery plan and in what order the IT team will restore the databases. The business must provide its input at this stage so that the most business-critical applications are available as soon as possible.

DATABASE AND OS BACKUP AND RECOVERY TOOLS

It seems obvious, but DBAs play the final and most important role in the process in that they must keep their knowledge of backup and recovery tools for RDBMSs up to date. During the actual restore event, DBAs will not have time to figure out any advancements in backup and recovery tools.

CONCLUSION

The primary responsibility of the database administration team is to review all types of RDBMSs in the enterprise and to develop a comprehensive backup plan to conduct effective backup management by proactively monitoring backups, getting alerted for failed backups and rerunning these seamlessly, without loss of time. It is good practice to back up data to physical disk and to then archive the data to tape for disaster recovery purposes.

Once an approach has been established, it is imperative to test data restoration periodically as part of the backup and restore strategy, and to review all options before executing the actual restoration/recovery. It is important to confirm that the DBA team is abreast of the latest backup and recovery tools and to ensure that the team has a clearly documented process in place with clear responsibilities. If DBAs maintain proper backups, monitor these proactively and can provide assurance of the recovery of data up to the point required by the business, they have done a major part of the job for which they were hired.

IT auditors can assist data administration teams in strengthening their controls and data recovery processes by validating the DBA operations, including the testing of the recovery of data. This continuous, proactive and cooperative effort between internal audit and the DBA team can provide assurance to management that, in the event of a disaster, the business's data can be recovered.