



IDUG VIRTUAL
2021 NA Db2 Tech Conference

Session D06
The Data Pioneer: Recovery
Martin Hubel, MHC Inc.

Db2 for LUW

Data pioneers are those of us that have to establish Db2 environments from scratch. Understanding recovery is a necessary skill for the Data Pioneer, and it arguably the most important skill. This presentation covers the determination of your requirements for recovery, provides a detailed overview of Db2's recovery mechanism, and various scenarios that should be covered in any recovery plan.

Agenda

- History
- The Importance of Backup & recovery
- Tasks and priorities
- Types of recovery and scenarios

Pioneers: A Lot of Work Ahead

- No services existed at all
 - Surveyors laid out the land allotments
 - Roads were planned (maybe) but not built
 - No fire department or any infrastructure
- Pioneers had to be self-sufficient
 - You have your own tools, or what you make
- Pioneers had to solve many problems
 - Experience taught many lessons
- When you start from scratch you have to do everything yourself



Fast Forward to Today

- A new business or division has started
 - Whoopee! Success!
- Developers have created a new application
 - Wow, it's so cool!
 - They worked so hard and quickly!
- Users and customers are coming on board
 - So many customers
- Demands on the database are growing
 - Let's hire a DBA

DBA Priorities

- Data integrity
 - Ensure availability & recoverability
 - Protect data
- Developer support
 - Data loads
 - SQL support
- Performance
 - System performance
 - Application performance

Selling Integrity to Management & Developers

- Level of availability required?
 - Cost per minute of outage for web application?
- How large is the batch window?
 - How long does the biggest job run?
- For customer data:
 - What is the cost if data is unrecoverable?
 - Can you afford to lose a customer?
 - What steps should you take to ensure customer data is recoverable?

Selling Recovery to Management

- Maximum length for an outage before severe financial impact?
 - What is the maximum acceptable recovery time?
- Cost and risk analysis:
 - Company business lost
 - Production capacity lost
 - Personnel downtime
 - Value of lost data

DBA Areas of Interest

- This list is in priority order for day 1 if you were the first DBA involved:

1. *Backup and recovery*
2. Naming standards
3. Security
4. Data movement
5. Support of developer practices
6. Data quality
7. Design of future databases
8. Performance

The first job of a DBA needs to be availability of the database, so recoverability is job 1.

Performance is a rewarding and engaging task, but on this list it is last. The database has to be up and developers productive before we concern ourselves with performance.

One-time Activities for Recovery*

- Take a backup if there are none
 - Review ability to recover
- Develop and use naming standards
- Develop scripts (short list with ties to recovery tasks)
 - Back up
 - Data loads / migration
 - Restore / redirected restore
 - Cleanup – old backups & archive logs
 - Db2 can automate some utilities
 - Schedule via cron or Db2's ATS (Administrative Task Scheduler – usually only Windows)
- * Adjust as necessary

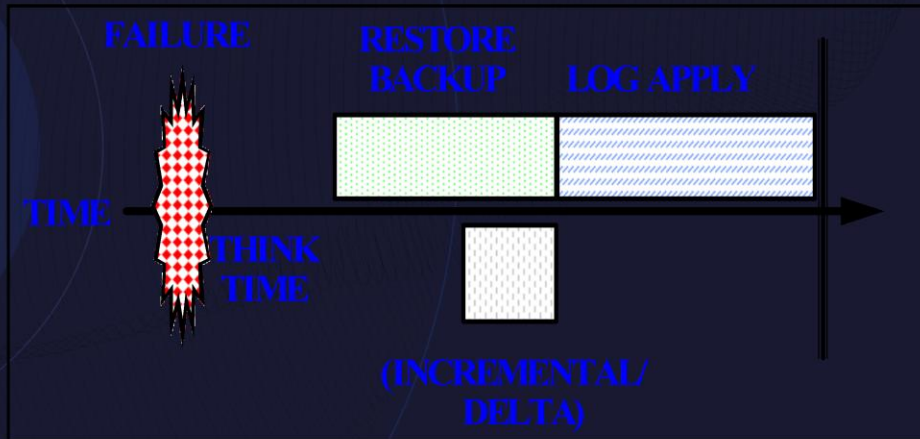
You may be surprised at the number of shops that are at risk by not having a usable backup. Some of my best stories relate to the superhuman efforts to recover from this situation. It is important to not only take a backup, but then go further to ensure that you know how to and can recover using this backup.

Scripts help to automate and standardize your backup and recovery tasks.

Backup Strategy

- Your backup strategy will determine your recovery options
 - Ensure you have backups that work
 - Backup should be a regularly scheduled activity
- Recovery may be required occasionally
 - Hardware failure
 - Recover to as close as possible to the point of failure
 - Application or software failure
 - Often to a chosen point in time before the failure
 - Disaster recovery
 - Offsite or cloud
- Test your recovery scenarios prior needing them
 - Failure to recover in production is never acceptable

Recovery Elements



The time required to recover is based on a number of factors. Think time includes the time to identify that a failure has occurred, the type of failure, and the type of recovery needed.

Strategies for Recovery

- A recovery plan should include procedures for all situations
 - Recover an application as quickly as necessary during normal recovery
 - Procedures must be tested
- Set logretain to yes
 - Allows for log archiving
- Use dual recovery structures
 - Raid or log mirroring – keep at least 2 generations of backups
- System information is critical
 - Must be copied as frequently as most critical user data

Recovery Tips

- Test your recovery procedures regularly
- If unsure, make a backup prior to recovery - things can get worse
 - Make a copy after recovery if you think it might happen again
- When using disk or file backup utilities:
 - Stop databases and applications first
 - Updated data is often in memory and not yet written to disk
 - Utilities outside of Db2's control do not allow forward recovery using Db2 logs

Types of Application Recovery

- Transaction recovery/backout
- Application backout
- Accidentally deleted / changed data
- Out of space
- Media failure / data inconsistencies
- To prepare & mitigate, learn about:
 - Recovery to current
 - Point-in-time recovery
 - Checkpoint/restart

This is a list of the things that can go wrong with application. Plan for these things to happen, and learn how to recover.

System Recovery

- These errors affect all users and applications
 - May need to recover the entire database
 - Note: situations like a loss of power require restart but not recovery
 - Db2 restart happens automatically when Db2 is started
- Needed as a result of:
 - Configuration errors
 - Operator errors
- To prepare and mitigate:
 - Dual recovery structures
 - Physical separation of files
 - Operator training

System recovery

Disaster Recovery

- Situations where the data center or offices become unavailable
 - System is down
 - Network issues
 - Fire, flood etc.
 - DBAs must actively participate in disaster recovery planning
 - Restore system from offsite / cloud backup
 - Even web services can have issues
 - <https://www.vox.com/2017/3/2/14792636/amazon-aws-internet-outage-cause-human-error-incorrect-command>

DB2 Diagnostic Information

- **db2diag.log**
 - File in db2dump directory – see dbm for location, size and retention options
 - Use head/ tail/grep from command line
 - db2diag command allows for searching of the db2diag.log
- **LIST HISTORY Command**
 - History file - recovery events include: full database and table space level backup, incremental backup, restore, and rollforward operations
 - Additional logged events include: create, alter, drop, or rename table space, reorganize table, drop table, and load
 - db2 list history dropped table all for db sample

DB2 Backups

- Simplest backup command
 - `db2 backup db sample`
 - This command takes a full DB copy to disk
- Can backup at database or tablespace level
 - Offline (DB unavailable) or Online (DB available)
 - Delta (last changes) or incremental (since full)
 - Consider backups to cloud locations
- Buffers and parallelism:
 - Do not specify and Db2 will choose an optimal value
 - Specify your own values if you know better
 - Must have sufficient `UTIL_HEAP_SZ`
- Restore time has been consistently slightly longer than backup time

Online Backup

- Log files needed for recovery can be included in the backup file
 - `BACKUP DB db TO path INCLUDE LOGS`
- To restore using the backed up logs -
 - `RESTORE DB db FROM path LOGTARGET logpath`
- To recover from the logs only -
 - `RESTORE DB db LOGS FROM path LOGTARGET logpath`

Log Management

- Log Archive Methods - valid values of LOGARCHMETH1/2 are:
 - OFF (circular logging)
 - LOGRETAIN (no log archiving)
 - USEREXIT (old user exit functionality)
 - DISK:/location
 - TSM or TSM:mgmtclass
 - VENDOR:/vendor/library
- db2tapemgr manages storing log files on tape
- More DB parameters to deal with failures

DB2 Restore & Rollforward

- Restore backup files
 - `db2 restore db sample`
 - Roll forward log processing is then done to current or another point
- Recovery options:
 - Database level
 - Tablespace level allows online restore
 - Apply incremental copies to full automatically
 - Can do ROLLFORWARD as a second step
- Will create DB if it does not exist
- Can redirect restore to another DB

Rollforward – To Where and When?

- **END OF LOGS: (Apply as many changes as possible):**
 - Rollforward will apply all available logs associated with the backup
 - Archived logs retrieved unless NORETRIEVE
- **Point-in-time (PIT): (Apply changes up to a specified time):**
 - Coordinated Universal Time (UTC) or LOCAL TIME. Format: yyyy-mm-dd-hh.mm.ss.nnnnnn
- **END OF BACKUP: (Apply as few changes as possible):**
 - DB recovered from an online DB backup then do ROLLFORWARD to earliest consistent point
- **Table space point-in-time considerations:**
 - Minimum roll forward time maintained for each table space
 - Table spaces are placed in backup pending when the roll forward completes

Recovery scenarios

- Almost all recovery scenarios require logretain yes
 - Circular logging is for DB where a total rebuild is expected
- Hardware or system software failure
 - Data and units of recovery should be OK
 - Roll forward to end of logs following restore
- Application / user / operations failure
 - Potential for incorrect updates
 - Some updates may be wrong
 - Need to decide on recovery time and choose that during recovery
 - Need to backup database afterwards if point in time chosen

Database Redirected Restore

- A database restore operation uses a database backup image to recreate a database in a new location:
 - Restore a backup image to a target machine that is different from the source machine
 - Restore your table space containers into a different physical location
 - Restore operation failed because one or more containers are inaccessible
 - Redefine the paths of a defined storage group
 - Use redirected restore as part of disaster recovery
 - Convert older DMS table spaces to automatic storage
- Consider `db2relocatedb` when moving or copying DB2 DB with non-DB2 utilities

HADR : High Availability Disaster Recovery

Primary



Processes all transactions.

Sends log entries to standby.

Log Records



Standby



Cloned from primary.

Receives and stores log entries.

Reapplies log entries

- **Very fast failover, transparent to applications**
 - Easy to set up and administer
 - Multi-site capability (cope with full site failure)
 - Low performance impact
 - **Software upgrades without an outage**

PureScale(tm)

- Premium solution for availability
- Multiple Db2 servers joined via a CF
- Benefits
 - Can lose a server with no loss of data or availability
 - Never need an outage for maintenance
 - 24x365 solution
- Getting easier to setup and administer with each Db2 release

Summary

- DBA who cannot recover their database are in serious trouble
 - Must learn and choose the facilities best suited to your environment
 - Practice your recoveries
 - Think outside the box
 - Something unexpected might take longer
- Most application recoveries are point in time
- Adapt your strategies to include the cloud
- Almost everyone uses HADR
 - PureScale is suitable for some customers

Martin Hubel is an independent consultant and has worked extensively with DB2 since 1985. Martin develops and teaches DB2 advanced courses and is recognized as a leading authority in the field. He has been using DB2 on Linux, Unix, and Windows since 1993 and has participated in several beta test programs for these platforms. He is an IBM Gold Consultant, a Lifetime IBM Champion, and a member of the IDUG Volunteer Hall of Fame.



Speaker: Martin Hubel
Company: MHC Inc.
Email Address: martin@mhubel.com

Don't forget to fill out a session evaluation!