What should you be aware of before your manager asks you to recover lost data?

Michael Kalouš
CA Technologies

November 2018, Last update 2018-11-07
Session **IH**

What should you be aware of **before** your manager asks you to recover lost data?

Agenda

- Reasons and Resources for recovery
- What might need to be recovered?
- Recovery points
- Data recovery - SQL
- Rebuilding indexes vs. recovering indexes
- REPORT RECOVERY
- After recovery
- Speeding up the process
- Recovery with non matching OBIDs

Reasons for RECOVERY

- Good
  - (?)
- Bad and not so uncommon
  - Software problems and user errors
- Worse and not so common
  - Object accidentally dropped
  - Hardware error
  - Site disaster

Resources for Recovery

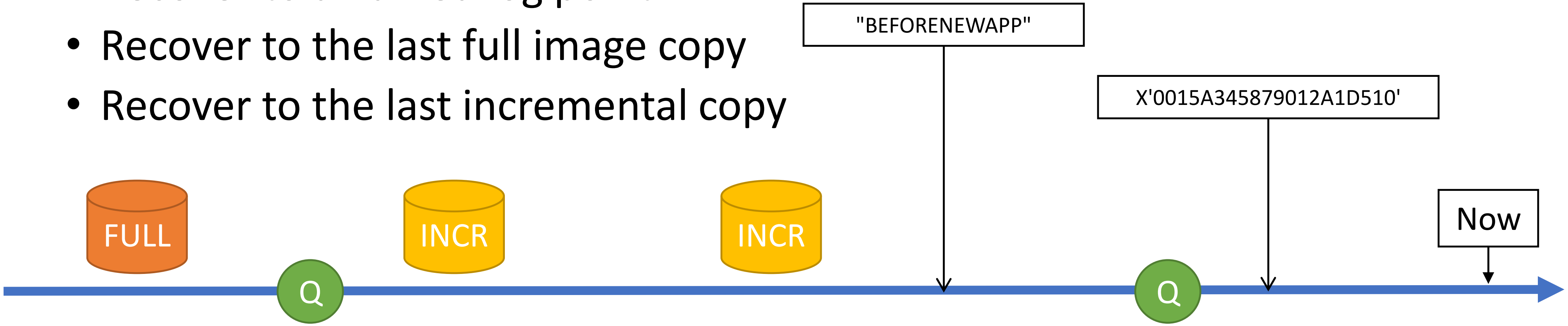
- Basics - DB2
  - SYSIBM.SYSCOPY, SYSIBM.SYSLGRNX
  - DB2 Log
  - Bootstrap Data Set (BSDS)
  - Utilities (RECOVER, REBUILD, REPORT RECOVERY)
- More
  - A log processing tool (it is not DSN1LOGP!)
  - Advanced recovery tool
  - Records created by the tools (tables, datasets).
    Do you have backups of these?

What might need to be recovered

- Database object(s)
- Application – we know PLAN/PACKAGE?
- DASD Volume, Storage Group
- Bootstrap Data Set (BSDS), Log data set
- DB2 catalog and directory, whole DB2 Subsystem (SLB)
- Dropped table or tablespace. Result of an "unfortunate accident". Fun begins when the object can no longer be created
- Object that is not logged (was that wise?)
- Data in the tables

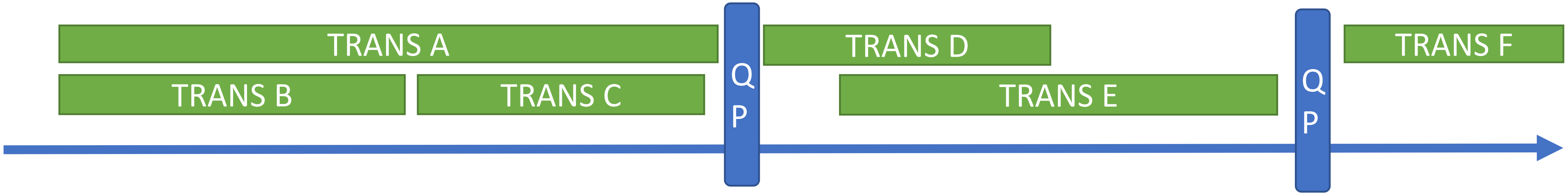
What is the PIT I'm looking for?
Is it consistent? Inflight transactions?

- Typical recovery points
  - Recover to the current point in time
  - Recover to a specific RBA or LRSN, or timestamp
  - Recover to a QUIESCE or a quiet point
  - Recover to a named log point
  - Recover to the last full image copy
  - Recover to the last incremental copy



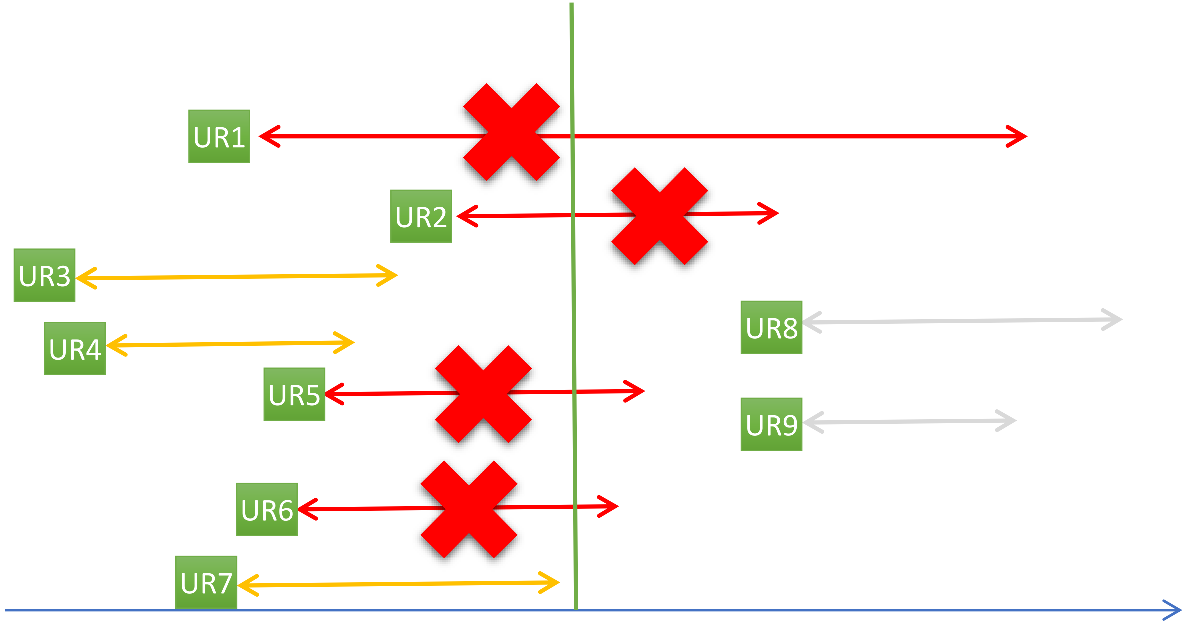
Quiet Point as a good recovery point

- Recovering to a point where there was no activity is beneficial
- Range of log where there is no activity against selected group of database objects is a quiet range
- Log processing tool can identify these ranges, identify their lowest RBAs and register them as ICTYPE='Q' (QUIESCE) rows



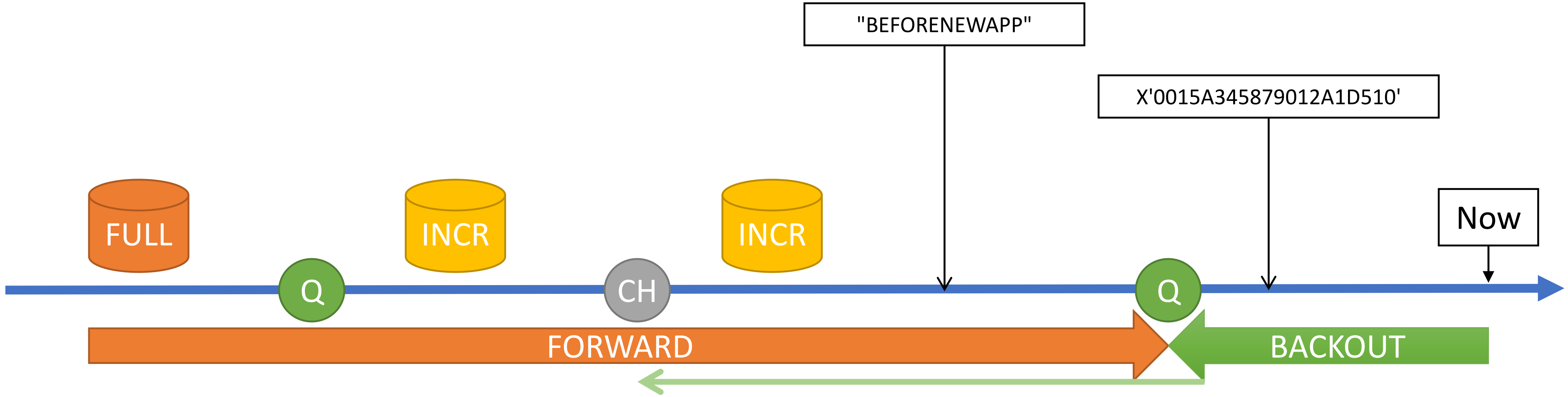
Inflight transactions
Recovery Point not Quiet

- You might not have a luxury to recover to a quiet point
  - RECOVER TORBA/TOLRSN X'.........'
  - Effects of unfinished transactions are rolled back by the RECOVER utility



Forward Recovery vs BACKOUT

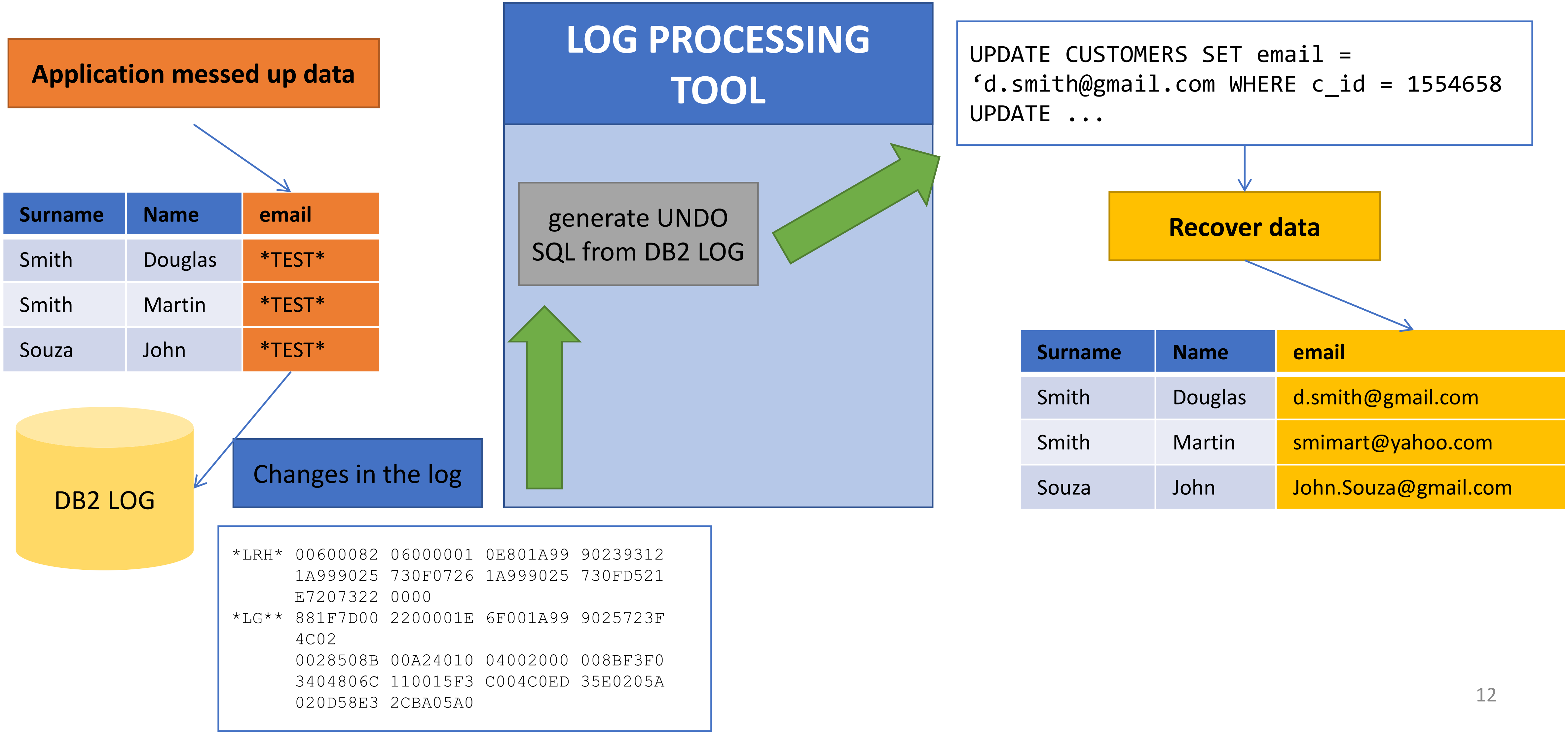
- When you have your recovery point (and your DB2 VSAM)
  - Forward. Merge Copies and log apply
  - BACKOUT. Read DB2 log and undo changes. If your copy is too old…



Data Recovery

- Principles
  - Application developers, while building the applications which access database objects, can mess up the data either on test subsystems or on production subsystems
    Anyone else with the database access can do "something wrong"
  - Database object is still there, online, only the data is wrong
  - DBA is required to resolve this situation and recover the data into its original state
- Resources you need
  - Log processing tool
  - DB2 log

Data recovery – UNDO SQL from DB2 LOG



Application messed up data

Surname	Name	email
Smith	Douglas	*TEST*
Smith	Martin	*TEST*
Souza	John	*TEST*

```
*LRH* 00600082 06000001 0E801A99 90239312
      1A999025 730F0726 1A999025 730FD521
      E7207322 0000
*LG** 881F7D00 2200001E 6F001A99 9025723F
      4C02
      0028508B 00A24010 04002000 008BF3F0
      3404806C 110015F3 C004C0ED 35E0205A
      020D58E3 2CBA05A0
```

```
UPDATE CUSTOMERS SET email =
‘d.smith@gmail.com WHERE c_id = 1554658
UPDATE ...
```

Recover data

Surname	Name	email
Smith	Douglas	d.smith@gmail.com
Smith	Martin	smimart@yahoo.com
Souza	John	John.Souza@gmail.com

Recovering an Application

- A business application is using a set of database objects. It might not be obvious that the database objects are related.
- Best if the development team knows what database objects are used, the knowledge is documented and the documentation is updated.
- The **SYSIBM.SYSPACKDEP** catalog table can help, if the answer is "I am not sure, but I know the package name".
- For given package name, you will get NAME, QUALIFIER, and TYPE of all objects involved (tables, views, indexes, and more)
- Static SQL only. For dynamic queries, **SYSIBM.SYSDYNQRYDEP**

Rebuild or Recover Index

- Rebuild
  - No need to keep image copies. This saves a lot of space
  - Is more CPU demanding (read TS, extract keys, SORT, assemble pages, write to the index space)
    Index on expression – evaluate the expression for every key – more CPU
  - After rebuild, the index records are in ideal positions, the index is fresh
  - Always available option. Index keys perfectly match data in the tablespace
- Recover
  - You must have a copy and keep the copy
  - Recovery requires I/O only. For very large indexes, it can be an advantage
  - Index records are in their original non-ideal positions
  - Recover to the same PIT as the tablespace. Recovery in parallel with TS

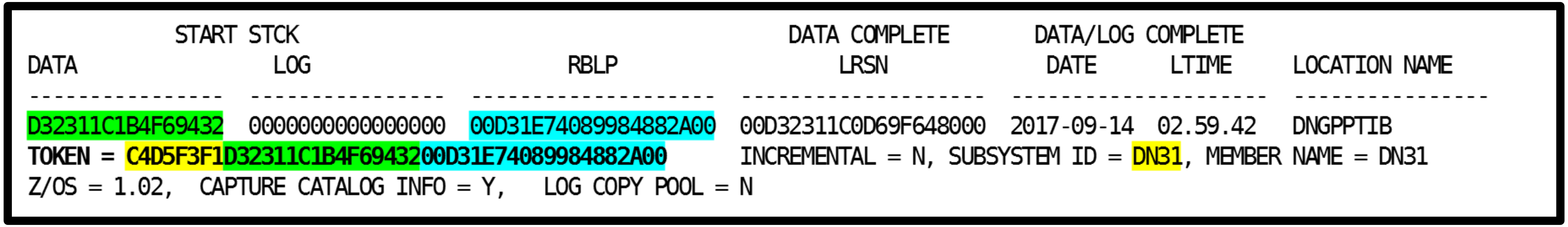
Rebuild or Recover Index

- When your tendency is REBUILD
  - Check indexes on large objects
  - Double check NPIs on large PBRs and PBGs
- When your tendency is RECOVER
  - Create copies together with tablespaces in one step
- Rebuilding with availability
  - REBUILD INDEX SHRLEVEL … CHANGE

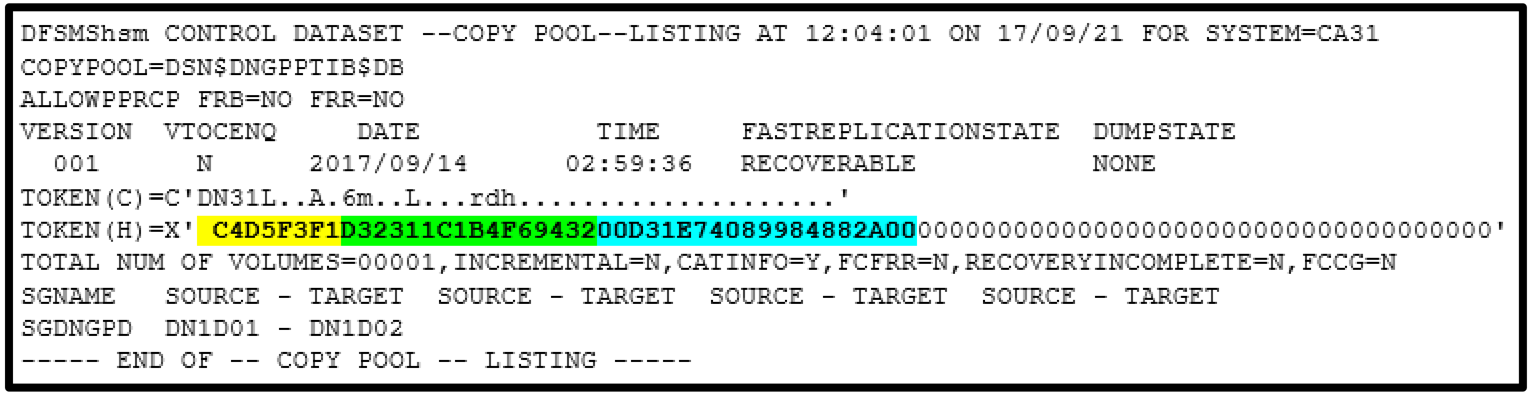
REPORT RECOVERY

- The recovery history from SYSCOPY, SYSLGRNX and BSDS
  - Copies
  - Other recoverable and non-recoverable events (LOAD, REORG, MODIFY, PIT RECOVERY, REBUILD)
  - Log inventory from BSDS
  - Log ranges
- Other
  - List of available system-level backups (SLBs)
- All in one dense report

REPORT RECOVERY – SLB example



```
            START STCK                                          DATA COMPLETE        DATA/LOG COMPLETE
DATA                    LOG                   RBLP                   LRSN            DATE       LTIME      LOCATION NAME
----------------   ----------------   --------------------   --------------------   ---------------------   ----------------
D32311C1B4F69432   0000000000000000   00D31E74089984882A00   00D32311C0D69F648000   2017-09-14   02.59.42   DNGPPTIB
TOKEN = C4D5F3F1D32311C1B4F6943200D31E74089984882A00       INCREMENTAL = N, SUBSYSTEM ID = DN31, MEMBER NAME = DN31
Z/OS = 1.02,   CAPTURE CATALOG INFO = Y,    LOG COPY POOL = N
```



```
DFSMShsm CONTROL DATASET --COPY POOL--LISTING AT 12:04:01 ON 17/09/21 FOR SYSTEM=CA31
COPYPOOL=DSN$DNGPPTIB$DB
ALLOWPPRCP FRB=NO FRR=NO
VERSION  VTOCENQ     DATE          TIME     FASTREPLICATIONSTATE  DUMPSTATE
  001      N      2017/09/14     02:59:36   RECOVERABLE           NONE
TOKEN(C)=C'DN31L..A.6m..L...rdh....................'
TOKEN(H)=X' C4D5F3F1D32311C1B4F6943200D31E74089984882A000000000000000000000000000000000000000'
TOTAL NUM OF VOLUMES=00001,INCREMENTAL=N,CATINFO=Y,FCFRR=N,RECOVERYINCOMPLETE=N,FCCG=N
SGNAME   SOURCE - TARGET  SOURCE - TARGET  SOURCE - TARGET  SOURCE - TARGET
SGDNGPD  DN1D01 - DN1D02
----- END OF -- COPY POOL -- LISTING -----
```

Can I avoid recovery?

- Yes, under some conditions
  - PIT
  - You still have DB2 VSAM datasets and you know they are intact
  - No changes since the indicated log point
  - SCOPE UPDATED
- Recovery avoidance saves resources

Ready to make decisions?

- Recovery Point
  - To which point to recover
- Recovery utility vs. Log Processing Tool
  - Quick direct access to pages
  - Generating SQL from DB2 log and executing it
- Standard forward recovery vs. BACKOUT
  - Merge copies and apply log
  - Process log and undo changes
- Rebuild indexes or recover them from image copy
- Advanced recovery tools can help
  - Offer possible options for recovery
  - Rate amount of work for each option

How I can speed up the process of recovery? Technical Aspects

- General Rule
  - A lot depends on the backup side, how much resources you are ready to spend?
- Reduce work for the tools in advance
  - Create incremental copies to reduce log apply
  - Merge incremental image copies to reduce number of copies
  - Borrow some IMS practice and accumulate changes for critical objects
  - Consider enabling DCC for your log processing tool
- Make most of the I/O subsystem
  - Use DASDs
  - Backup on partition level and use options for parallel processing
  - Consider stacking image copies on tapes using LBI
  - Fast replication capabilities

How I can speed up the process of recovery? Procedure and Organization

- Have a recovery procedure outlined and in writing
- Define roles and have answers to the question
  - "Who can make a decision to which point in time we are heading with the recovery"
  - "What is the scope of recovery?"
- Ensure involved workers have required authorizations
- Allocate resources to practice and test the recovery procedure

After Recovery

- Phew, the data is back, but remember …
  - Deal with dependent object statuses.
    You can observe RECP, RBDP and CHKP.
  - Do not stray from the access path.
    Some of DB2 statistics can be inaccurate. Gathering statistics for tablespaces and indexspaces is needed to prevent access path degradation.
- Before declaring victory
  - Create new backups
  - Let users check the data and run their applications

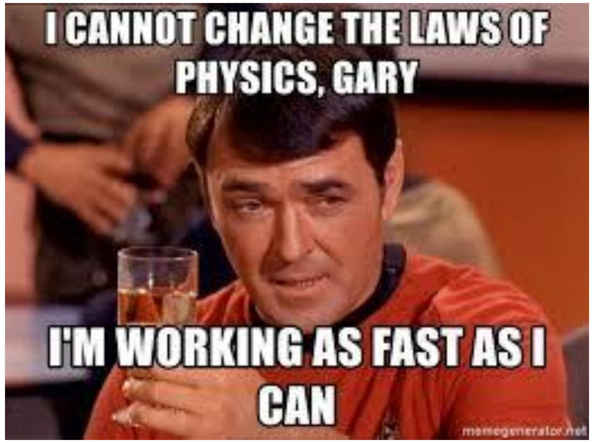
Final Topics

Reasons for RECOVERY

- Good
  - Practicing. Simulation and Estimation
- Simulation, fully featured "dry recovery"
  - Everything is recovered is in the real run, except
    - Database objects are live and not touched
    - The resulting DB2 VSAMs are "aside" the original ones
  - You have a solid evidence of
    - Recoverability
    - Time needed to recovery
- Estimation
  - A "lightweight" version of simulation
  - Information on recovery resources is analyzed and time needed to recover is estimated

Service Level Agreement (SLA)

- In how much time the application must be up and running when something goes wrong.
- Keep in mind
  - Tapes and disks still have limited speed v << c
  - Giant objects will take time to recover despite high image copy frequency
  - Humans need time to make decisions


I CANNOT CHANGE THE LAWS OF PHYSICS, GARY
I'M WORKING AS FAST AS I CAN
memegenerator.net

- Fast replication solutions (FlashCopy, SLB) can be an answer when traditional copies are out of breath. Availability x COW.
- Simulate and estimate to check if you fit

Recovering Objects with Non-Matching OBIDs

- Non-matching OBIDs ?!
  - At the time of full copy, the object had different OBID than at the time of recovery
- When this is needed?
  - When recovering a dropped database object (ouch)
  - When migrating data from one object to another object using an image copy
- What the recovery tool must handle
  - Change the OBIDs in the pages, when needed
  - Identify log records to apply with old OBIDs

Dropped Table

- Implications of dropping a table are very serious
- Much is lost
  - The data, associated objects
  - Information in the catalog (SYSTABLES, SYSCOLUMNS, etc.)
  - Data in SYSIBM.SYSCOPY, SYSIBM.SYSLGRNX
  - Statistics, Authorizations, Triggers
  - Packages are invalidated
- Something is still left
  - Image copy datasets
  - Records in the DB2 log.

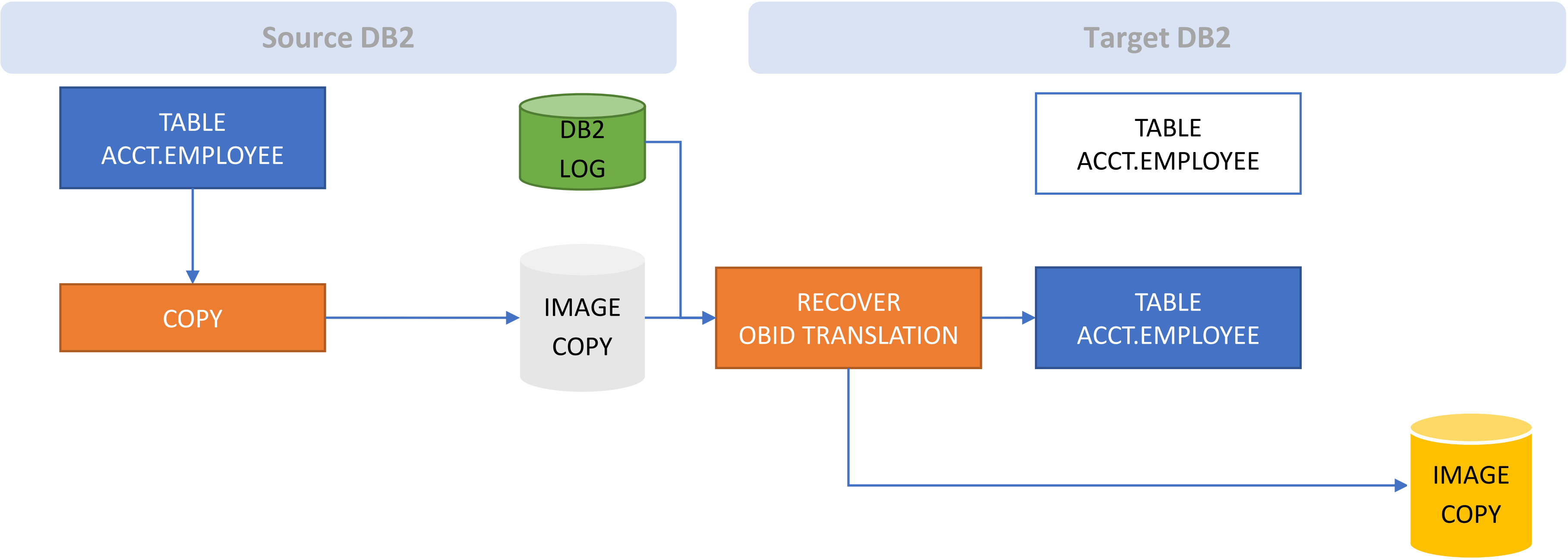
Dropped Table

- Recovering the dropped table
  - Recovery is officially described in 11 manual steps.
- Advanced recovery or log processing tool helps with the process
  - DB2 logs the changes to the catalog tables (SYSTABLES, SYSCOLUMNS,…,SYSCOPY)
  - DDL (table + assoc. objects) is reconstructed and available copies are identified
  - Recovery step is prepared (with or without log apply)
  - REDO SQL is generated from the log and executed (without log apply)
  - Indexes are rebuilt
  - Statistics restored or regenerated. Invalidated packages are rebound
- All you need to know is the name of the object and approximate time range when it was dropped

Migrating Data with RECOVER

- Make image copy a dual-use technology
  - Backup and a convenient source for data migration
- Pros and Cons
  - Faster than a sequence of UNLOAD and LOADs
    Operating on page level
    No temporary unload dataset is needed
  - More sensitive to differences between the source and target objects
    Some differences can be prohibitive
- Is a non-recoverable event
  - Image copy is recommended

Migrating Objects with Non-Matching OBIDs



Conclusions

- Review your backup strategy.
  Do you have a strategy for everything that might need recovery?
- Know in advance your application, dependencies, critical objects.
  This way, you will know what to recover
- Roles in the process should be defined. Who makes the decisions?
- Remember that recovery requires resources – CPU, tape units, disk space, resources for SORT
- Know strengths and weaknesses of your tools
- Get a log processing tool
- JCLs should can be prepared in advance
- ….. and PRACTICE and TEST your recovery

Questions

We want your feedback!

- Please submit your feedback online at ….
  - http://conferences.gse.org.uk/2018/feedback/IH

- Paper feedback forms are also available from the Chair person

- This session is **IH**