

White Paper

By: Ulf Heinrich, SOFTWARE ENGINEERING and SEGUS Inc

Next Level DB2 Backup and Recovery Management

Objective

Ideally, all work in a computer center should be based on Service Level Agreements (SLA) that all interested parties in a company have agreed to.

The SLA for database backups and recovery management is also known as Recovery Time Objectives (RTO). Therefore, the RTO is a critical metric that must be considered when determining whether an SLA can be met. In an ideal computing world, all backups are always available and current. But, there is no such thing as an ideal world. Even if all backups are available and current, most data centers have no idea whether the duration of a recovery will actually support the business needs.

This paper explains why such SLAs are difficult to meet with current processes, and it provides a new perspective on how to implement and reach RTOs with reliable and cost-efficient processes. Just as questionable as backup processes and recovery times, the costs that are involved with today's recovery processes are not transparent and are always rising. This paper explains how to achieve higher recoverability and availability at a much lower cost.

Current backup and recovery processes

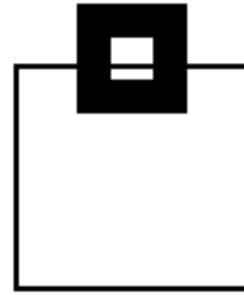
There is a misguided notion that recovering small objects, especially those with minimal updates, will always be fast. But this does not take into consideration situations where updates to a single page cover several archive logs. Recovering a large object with high update rates, all of which are covered by the active log, can be much faster.

On the other hand, what about those large objects with high update rates that contain your most critical enterprise data? Although a well-adjusted procedure to back them up every night is fine and good, large volumes of transactions during the day may lead to the fact that the update rates are already exceeded by midday. Being forced to recover using excessive log applies leads to a significantly increased recovery time.

The root cause of the above mentioned problems lies in the absence of the appropriate information stored in DB2 internal repositories like SYSLGRNX or unknown details affecting the recovery-status of a DB2 database, combined with old-fashioned backup strategies. Usually, only technical triggers for the generation of copy utilities are considered, e.g. pending states, update rates, or elapsed time between copies. And, these backups are done in batch. This means that copies are done for complete databases, or at least for object groups, all based on the same thresholds.

The drawbacks of current backup strategies are:

- Expected recovery times cannot be ascertainable using such a backup strategy.
- Similar DB2 objects have, in fact, completely different internal DB2 processes and therefore have completely different recovery times.
- A single, although perfect, strategy for one DB2 object is simply unfeasible in today's large systems.
- If no procedures are in place to provide for "just-in-time" backups during the day, you can kiss your SLA goodbye when extensive log applies take too long.



Systematic, repeated checks of the recovery state are not or very rarely done in reality. The estimated recovery state is based upon guesses, hunches, and experience that might be compromised due to unusual workloads, changes to the infrastructure or different DB2 behaviors.

At minimum, the consequences are:

- The DBA tends to go for a higher backup rate so that he is "on the safe side". This overloads the system, the environment, and the backup pools; while creating unnecessary costs. An example of this is a full image copy for all objects on a certain day-of-the-week or incremental copies after the first update.
- An unknown number of objects with unexpectedly long recovery times still exist. When a recovery must be done, the availability of the system is negatively impacted.

At worse, missing backups can result in permanent loss of data or unavailable resources can make any RTO impossible. The financial loss to the business can be significant.

What can be done?

An intelligent use of real-time statistics, combined with monitoring and reporting all recovery aspects of the DB2 environment and using, e.g., IBM's online COPY, will insure just-in-time backups that support business availability requirements. Additionally, there must also be a way to answer the age-old management question "How long would an outage last?".

Such a backup and recovery process must be an ongoing, proactive, daily process, which only automation can provide. Recovery AssuranceExpert for DB2 z/OS checks all components of availability and recoverability for all DB2 subsystems enterprise-wide. Features include:

1. Detailed checks of relevant DB2 availability or recovery affecting factors and recommended best practices for a recovery, i.e., ZPARMS, buffer pools, coupling facility, etc.

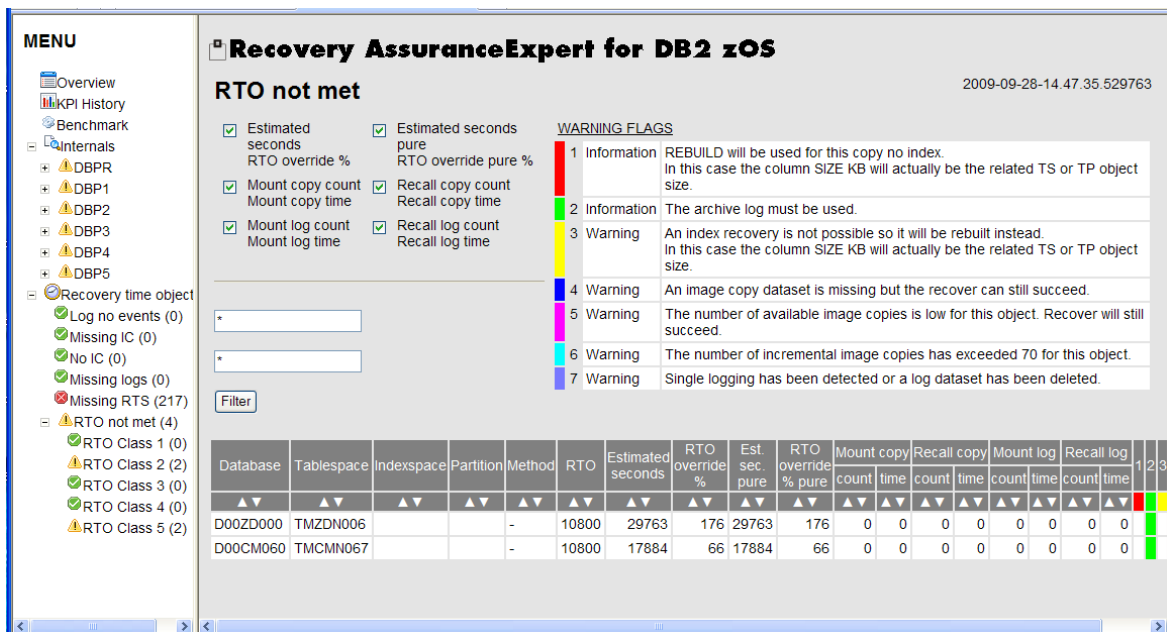
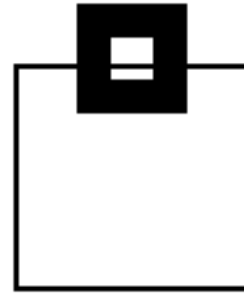


Figure: Drill-down technical reports provide warnings and errors at the object level



2. Determination of recovery times based upon chosen benchmarks and interpolated recovery-times for all objects in a DB2 database and comparison of these times to customer-specific Recovery Time Objectives (RTO):

- RTOs are dependant on the size of the DB2 object.
- Benchmarking measures the I/O performance of DB2 logs, backups, table and index spaces for VSAM and sequential data sets on tape and disk.
- In lieu of using technical thresholds for backups, Recovery AssuranceExpert can determine copy needs based on RTOs, when combined with an on-demand maintenance solution like RealTime DBAExpert.

Recovery AssuranceExpert supports all of this with simple operations and minimal resource usage. It supports interfaces to system management solutions, like Tivoli, as well as mainframe, Excel, and web-based reports.

Recovery AssuranceExpert delivers its results in two different forms. Technical reports provide detailed information at the object level. Easy drill down reports and indicators pinpoint any object that is unrecoverable or whose RTO cannot be met.

A management report provides a key performance indicator that represents the total system in regards to recovery and availability. An overall recovery indicator can be used in automation and reporting systems to issue an alert when any unrecoverable situation exists.

Next level DB2 backup and recovery management

Clear knowledge about the recovery status of an object enables an SLA-based, dynamic backup process exploiting Recovery AssuranceExpert. Creating copy strategies based on RTOs reaches a superior level of quality for copy und recovery management that fits exactly to the business needs. Such a level is:

Autonomic–self-regulating process without „guesswork“.

Object oriented–Highest granularity on the object level for absolute minimal costs.

SLA compliant–Fastest possible availability in recovery cases.

Conclusion

The theoretical problems of the backup process suddenly become very real indeed when a large recovery problem arises. Downtime for the actual problem is not known and nearly impossible to quantify.

Recovery AssuranceExpert monitors all relevant factors affecting a recovery and answers questions on a multi-departmental level. Checking RTOs with Recovery AssuranceExpert provides an efficient and inexpensive solution to satisfy „Business-Driven“ backup strategies, which also demands transparent and predictable recovery management. And, Recovery AssuranceExpert answers the big question, how long an outage may last before it happens.

About the Author: Ulf Heinrich is Solutions Delivery Director at SOFTWARE ENGINEERING GmbH. He specializes in DB2 recovery issues and database maintenance, focusing on the growing requirement for 24x7 operations and cost reduction. As a consultant at large customer sites, he has implemented database maintenance procedures and recovery strategies, and also experienced the pitfalls of recovery scenarios under real-world recovery pressure. His activities cover EMEA, as well as North America through SE's U.S. subsidiary, SEGUS Inc.