Stop Wasting Time With Utilities!

—

Jim Dee

Chief Architect for DB2, BMC

IBM Information Champion 2015

May / 2015

Objectives

1. Answer the question: What is the ideal frequency of execution for Reorg, for Copy, for Runstats, and for Unload and Load?
2. Can we achieve the ideal frequency of execution? (Hint: there wouldn't be a presentation if the answer were yes.) If we cannot, how can we plan to get as close as possible to that ideal?
3. We will look at the details of Reorg, Runstats, and Copy from an operational perspective.
4. We will look at Real Time Statistics and how to exploit them to complement an effective utility execution strategy.

Objective

Answer one simple question:
“What is the ideal frequency of execution of Reorg, Copy, and Runstats, and how close can we get to it?

Close to Ideal Timing

Want to run Reorg and Runstats just often enough to maintain application performance

- For static SQL, no point to Runstats if not rebinding

Want to run Copy just often enough to:

- Keep log apply manageable during a potential recovery
- Avoid COPYP status

Why Do We Run Reorgs?

Online Schema Change materialization

Extent consolidation? Seriously?

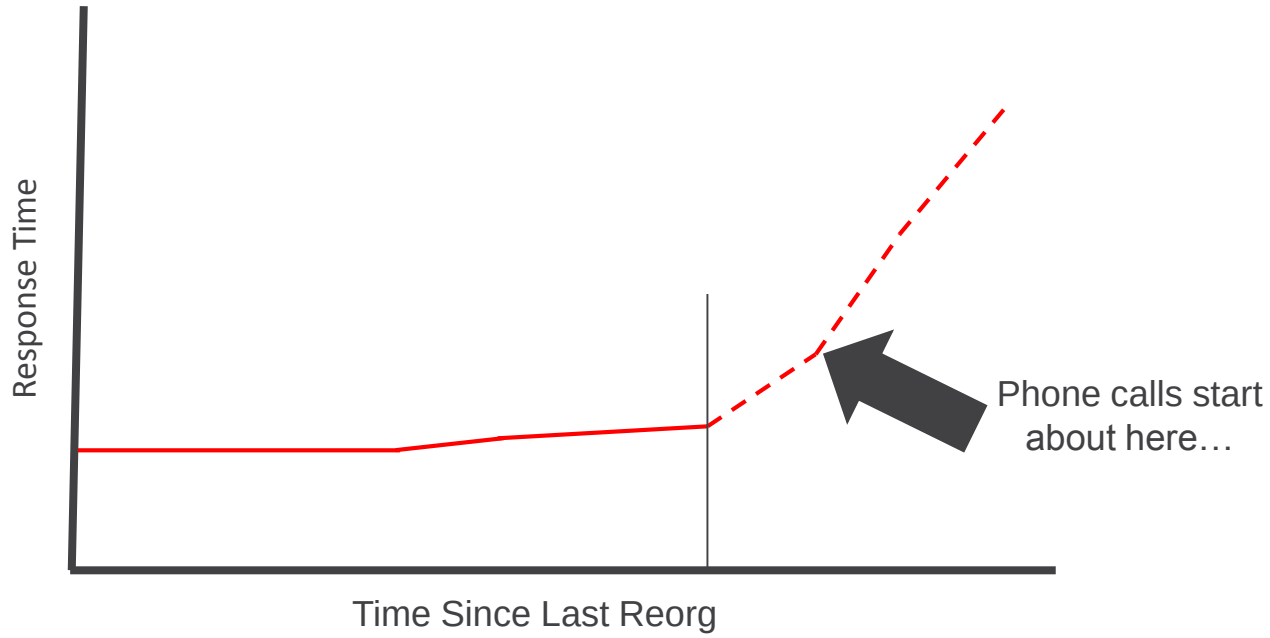
- 23,000 and change isn't enough?
- No performance impact today

Making available the space for pseudo-deleted rows or index entries

Rebalancing or freeing space by partition

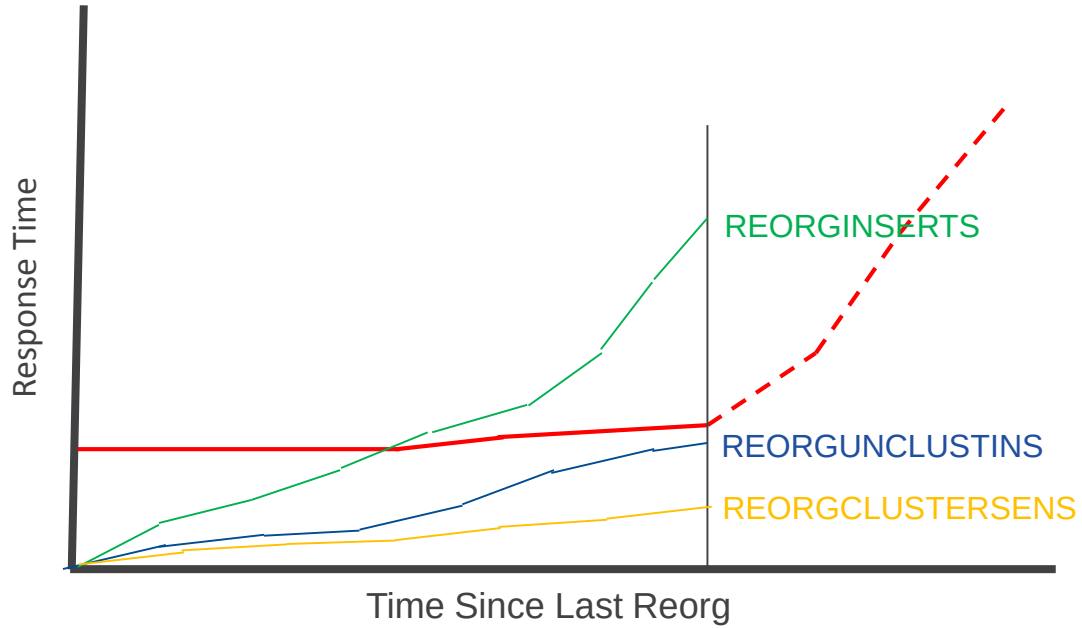
Which leaves us with performance…

Perfect Time to Run Reorg

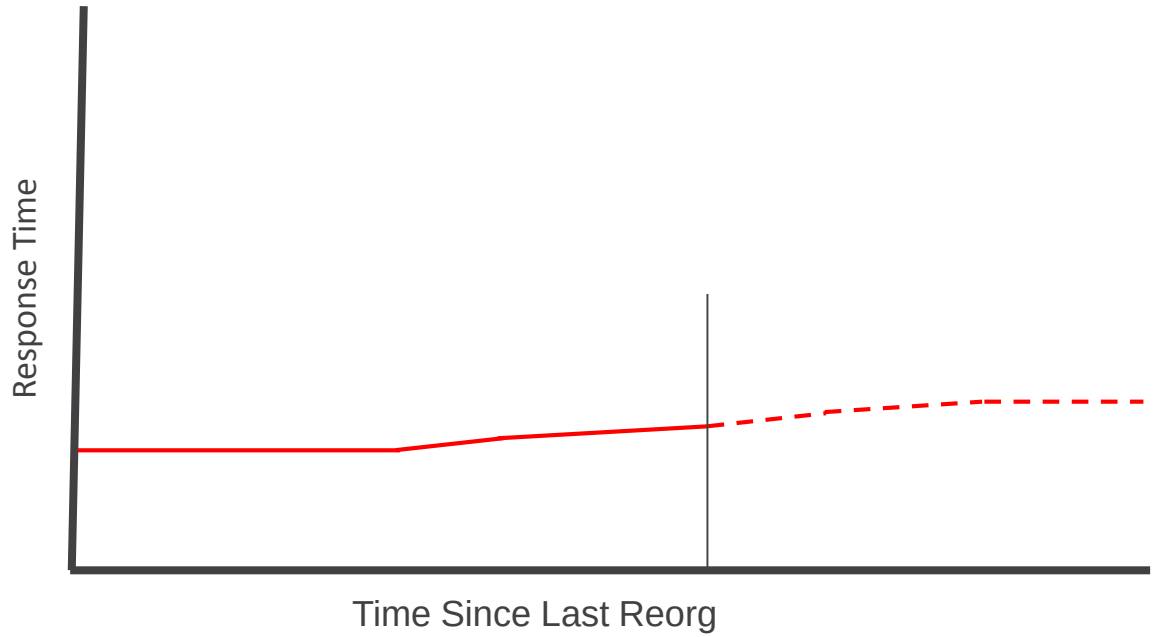


- Ideal input would be performance trend for every SQL statement against the object

Performance Data is Expensive to Collect



Finally…



- Is this slow growth trend in response time reflecting data volume growth, or will a reorg restore performance?

REORG…INDEX or TABLESPACE?

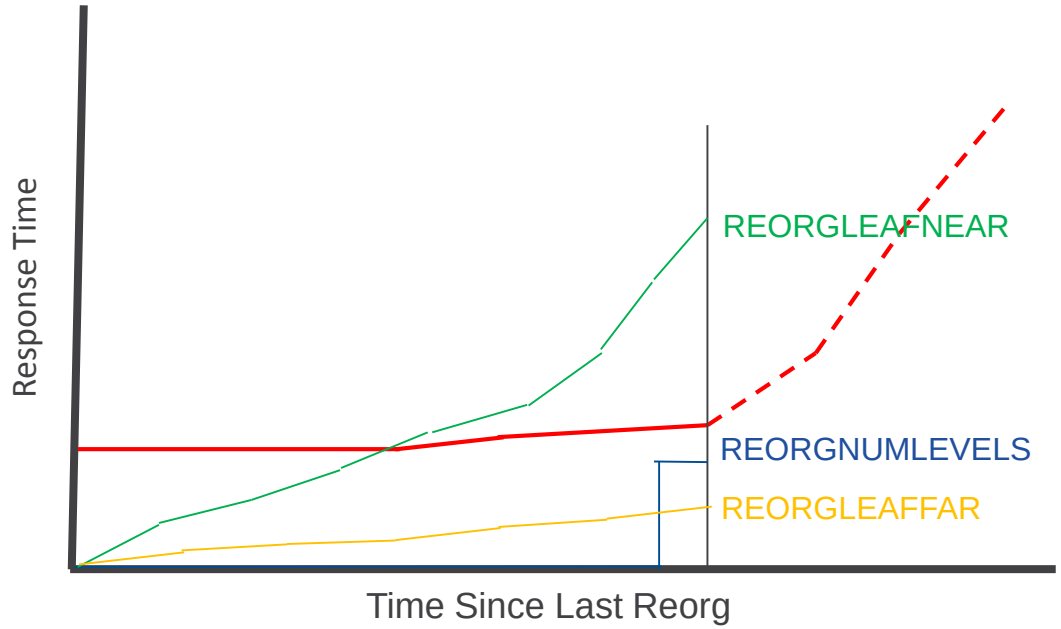
Random access can be vulnerable to disorganized indexes

REORG INDEX may solve your performance issue

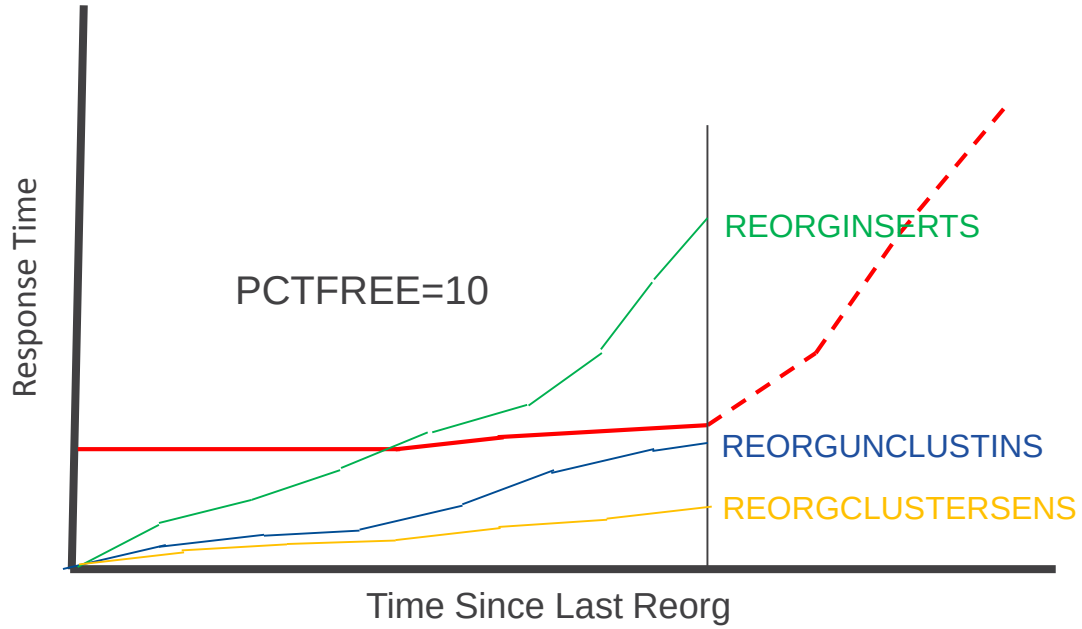
Do you run data sharing?

- How many page splits do you have?
- How much index contention do you see?
- MEMBERCLUSTER might solve that problem…but

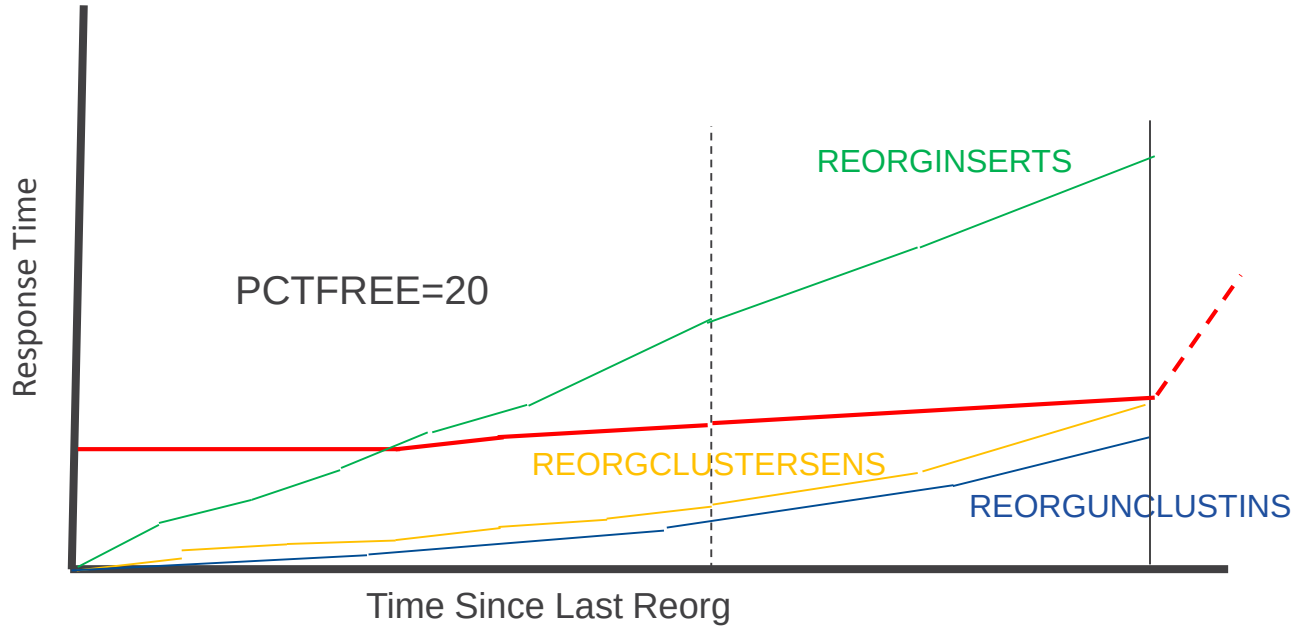
Performance Data is Expensive to Collect



Can You Trade DASD For Performance?



Can You Trade DASD For Performance?



Operational Considerations for Reorg

Want cost of reorg to be lower than application gains

Don’t run during peak of R4HA

Don’t run during high access period

Use thresholds on correct metrics to run reorgs when required but only when required

What is the Ideal Frequency for RUNSTATS?

- Assuming you've automated reorgs and you're maintaining your spaces…and your application performance…

- Almost Never!

Why Do We Run Runstats?

Schema changes (including adding or changing indexes)!

Some application code changes

Data volume changes

- Must be significant!

To add new statistics

- Which don't exist but are necessary

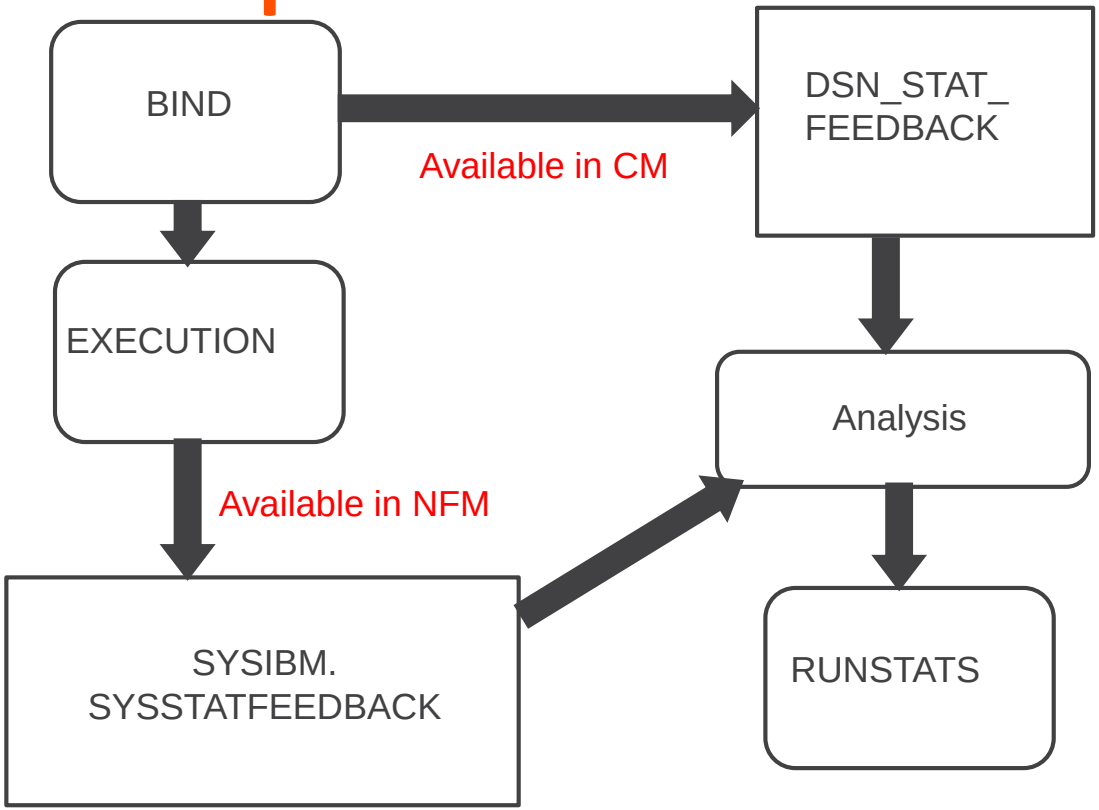
To exploit RUNSTATS improvements

- Like Enhanced statistics in DB2 9

Why run RUNSTATS if access paths won't change?

Must REBIND static packages

Optimizer Help for RUNSTATS



Operational Considerations for Runstats

RUNSTATS is not as disruptive as REORG, but it can cause an outage

Associated BIND or REBIND is more disruptive

- Exploit the increased concurrency in DB2 11!

Use Access Path Stability!

Consider saving old stats conveniently, so you can back out quickly if paths regress

Why Do We Run Copy?

Only purpose is to improve performance or avoid degradation…wait, no…

Purpose is to provide data to RECOVER utility

Secondary purpose is to allow updates after a REORG or LOAD

Secondary purpose could be as source for data migration

Frequency is a function of how much log would have to be applied during RECOVER

Is it Possible to Avoid A Copy?

What if the space has not been updated since the last copy?

- Not realistic often in production
- What about test systems?

You may be able to save time and storage space with incrementals

Do you want to copy indexes?

- Do you know at what point the savings in rebuild time outweigh the cost of making index copies?
- You still have to apply the log!

Test recovery!

- It’s the only way to know if your copies are frequent enough

What About Unload and Load?

Usually not scheduled utilities

Exception is periodic ETL

- One case where you may be executing many unloads or loads concurrently

You need to automate to get the timing right

MODIFY?

Should MODIFY RECOVERY be an automated utility?

- Necessary to keep your catalog uncluttered!
- Consider automation to clean up the MVS catalog (if you're not using GDG's) and to free tapes
- Consider RETAIN LOGLIMIT or RETAIN LAST
- Don't run MODIFY RECOVERY with REORG!

What about MODIFY STATISTICS?

- Again, keep your catalog uncluttered!
- AGE or DATE are a little clearer here

Characteristics of Automation Solution

- It should submit utilities as soon as possible after need is detected
- It should time execution to minimize application outages
- It should consider system load
- It should use LISTDEF and PARALLEL (where possible)
- Ideally, it should handle failure smoothly, and requeue failed utilities automatically

Summary

REORG is the biggest opportunity for automation

Be careful with RUNSTATS!

COPY automation is best done in the context of a recovery strategy

LOAD and UNLOAD can be automated…with restrictions

Thank You

Bring IT to Life.™

Jim Dee

BMC Software

jim_dee@bmc.com

Stop Wasting Time With Utilities