

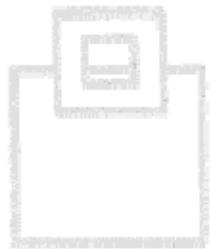
Are you a Runstats Master?

Roy Boxwell – Senior Software Architect

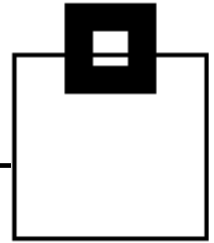
 **SOFTWARE ENGINEERING GMBH**

GSE UK Conference 31st October – 1st November 2006

SOFTWARE ENGINEERING



SOFTWARE ENGINEERING – What's new 2006:



Anniversary:

“We are in the Tools Business” – since 1986

Current seminars & Workshops:

- DB2 Migration to V8 – Be Prepared



Current development highlights:

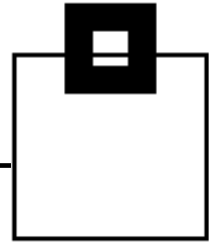
- DB2 V8 EarlyPrecheck ✓ (GA since January 2006)
- DB2 V9 Test & Development (Research since Feb. 2005)
- Feeding the Optimizer – Best in, Best Out
- **Recovery HealthCheck**



SOFTWARE ENGINEERING

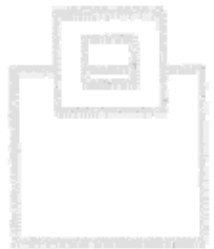


SOFTWARE ENGINEERING – What's new 2006:



New  **HealthCheck series™ for DB2 z/OS**

product family:

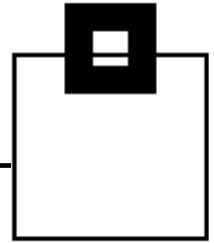



-  **Statistics HealthCheck**
-  **Recovery HealthCheck**
-  **Exception HealthCheck**

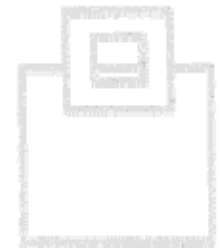
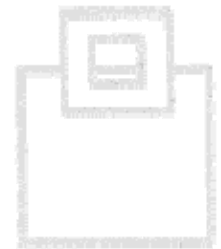


SOFTWARE ENGINEERING

AGENDA

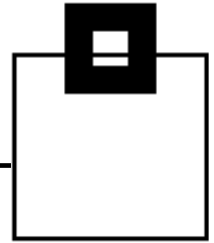


- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- RUNSTATS walk through
- SYSCOLDIST explained
- RUNSTATS Q & A
-  **Statistics HealthCheck**
- Notes & Questions



SOFTWARE ENGINEERING

Basic RUNSTATS knowledge

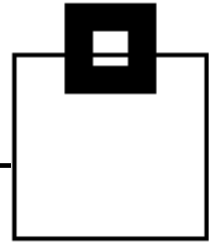


- The RUNSTATS utility
 - gathers summary information about the characteristics of data in table spaces, indexes and partitions
 - Invalidates the dynamic statement cache
 - Optionally
 - Reports the statistics
 - Updates the DB2 catalog
 - Updates the DB2 catalog history tables



SOFTWARE ENGINEERING

Basic RUNSTATS knowledge



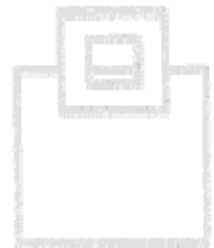
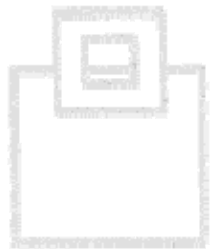
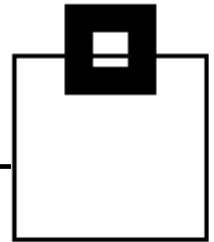
- **Different types of RUNSTATS**
 - RUNSTATS tablespace
 - RUNSTATS index
 - REORG / LOAD with Inline RUNSTATS
- **Different types of statistics**
 - **Pure access path statistics**
 - Those used by BIND in its process of optimization to determine access path
 - **Parallelism access path statistics**
 - Those used by BIND in its process of optimization to determine the degree of parallelism
 - **Space statistics**
 - Those used by the DBA to monitor space usage; to assist in capacity planning; to determine frequency of reorg; etc.



Catalog tables used for access path

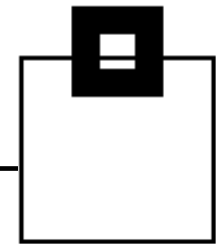
- SYSIBM.SYSCOLDIST
- SYSIBM.SYSCOLSTATS *
- SYSIBM.SYSCOLUMNS
- SYSIBM.SYSINDEXES
- SYSIBM.SYSINDEXPART
- SYSIBM.SYSROUTINES
- SYSIBM.SYSTABLES
- SYSIBM.SYSTABLESPACE
- SYSIBM.SYSTABSTATS

* degree of parallelism only



SOFTWARE ENGINEERING

Columns used for access path



SYSCOLDIST 3.1

- **CARDF** 5.1
- **COLGROUPOCOLNO** 5.1
- **COLVALUE**
- **FREQUENCYF**
- **NUMCOLUMNS** 5.1
- **STATSTIME**
- **TYPE** 5.1

SYSCOLUMNS

- **COLCARDF**
- **HIGH2KEY**
- **LOW2KEY**

If not marked then
table/column used
in DB2 2.3

SYSINDEXES

- **CLUSTERING*** 3.1
- **CLUSTERRATIOF**
- **FIRSTKEYCARDF**
- **FULLKEYCARDF**
- **NLEAF**
- **NLEVELS**

SYSINDEXPART

- **LIMITKEY*** 4.1

SYSTABLES

- **CARDF**
- **EDPROC***
- **NPAGESF**
- **PCTROWCOMP** 3.1

SYSROUTINES 6.1

- **CARDINALITY***
- **INITIAL_INSTS***
- **INITIAL_IOS***
- **INSTS_PER_INVOC***
- **IOS_PER_INVOC***

SYSTABSTATS 3.1

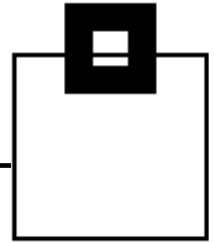
- **CARDF** 4.1
- **NPAGES** 4.1

SYSTABLESPACE

- **NACTIVEF**

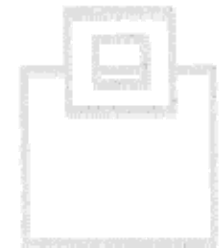
* This column is not
updated By RUNSTATS

So what?



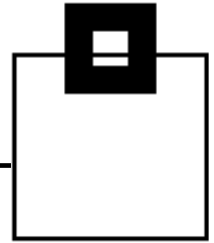
OK, we now know all the info that DB2 uses to choose its access path. What can we do with this info?

- We can change it to, hopefully, improve an SQL.
- We can delete some of it to, hopefully, improve an SQL.
- We can insert into it to, hopefully, improve an SQL.
- We can ignore it and simply trust that DB2
„Knows what it is doing...”
- We can mess it all up.
- We can use clever add-ons to *really* mess it all up!
- We can use freeware to see what state all these statistics are currently in. Good, Bad, or Awful.

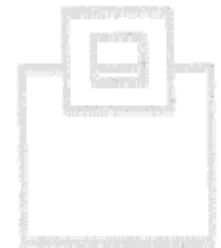
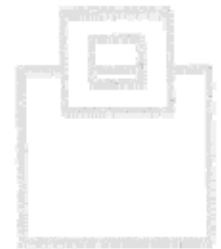


SOFTWARE ENGINEERING

AGENDA

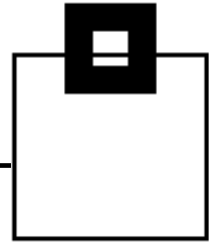


- RUNSTATS basics and DB2 Optimizer relevant statistics
- **IBM recommendations through the ages**
- RUNSTATS walk through
- SYSCOLDIST explained
- RUNSTATS Q & A
- Statistics HealthCheck
- Notes & Questions



SOFTWARE ENGINEERING

IBM Recommendations DB2 2.3



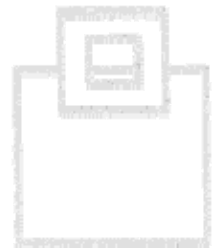
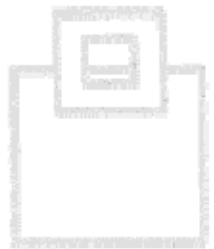
Correlations in the catalog (DB2 Administration Guide)

Relationships exist among certain columns of certain tables:

- Columns within SYSCOLUMNS
- Columns in the tables SYSCOLUMNS and SYSINDEXES
- Columns in the tables SYSCOLUMNS and SYSCOLDIST (well actually SYSFIELDS in those days!)

If you plan to update some values, keep in mind the following correlations:

- COLCARDF and FIRSTKEYCARDF/FULLKEYCARDF
- COLCARDF, LOW2KEY and HIGH2KEY. For non-default COLCARDF

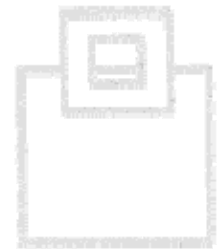
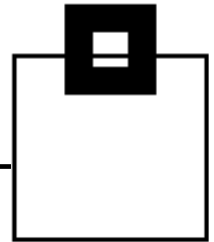


SOFTWARE ENGINEERING

IBM Recommendations DB2 3.1

Correlations in the catalog (DB2 Administration Guide)

No change

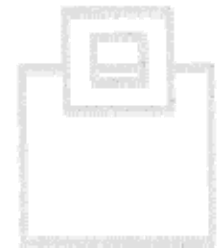
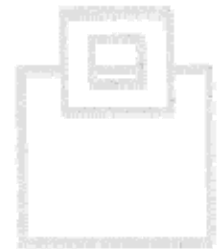
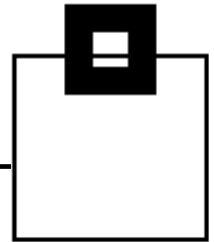


SOFTWARE ENGINEERING

IBM Recommendations DB2 4.1

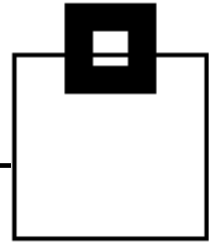
Correlations in the catalog (DB2 Administration Guide)

No change



SOFTWARE ENGINEERING

IBM Recommendations DB2 5.1



Correlations in the catalog (DB2 Administration Guide)

Relationships exist among certain columns of certain tables:

- Columns in the tables SYSCOLUMNS, SYSCOLDIST, and SYSINDEXES



If you plan to update some values, keep in mind the following correlation:

- CARDF in SYSCOLDIST. CARDF is related to COLCARDF and FIRSTKEYCARDF and FULLKEYCARDF. It must be the minimum:
 - A value between FIRSTKEYCARDF and FULLKEYCARDF if the index contains the same set of columns
 - A value between $\text{MAX}(\text{colcardf of each col})$ and the product of all the columns COLCARDFs in the group

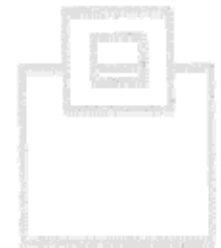
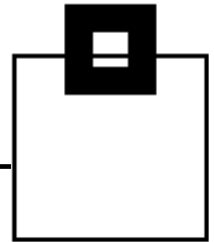


SOFTWARE ENGINEERING

IBM Recommendations DB2 6.1

Correlations in the catalog (DB2 Administration Guide)

No change

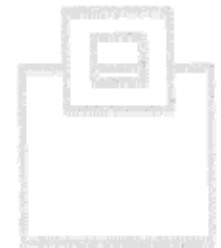
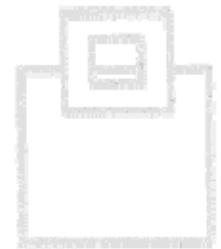
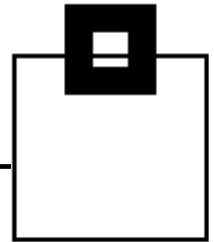


SOFTWARE ENGINEERING

IBM Recommendations DB2 7.1

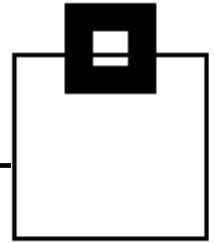
Correlations in the catalog (DB2 Administration Guide)

No change



SOFTWARE ENGINEERING

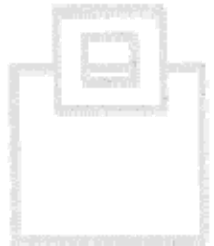
IBM Recommendations DB2 8.1



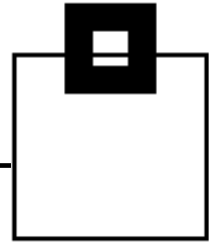
Correlations in the catalog (DB2 Administration Guide)

If you plan to update some values, keep in mind the following correlations:

- The COLCARDF, LOW2KEY, and HIGH2KEY gained a note:
„If the COLCARDF is 1 or 2 DB2 uses LOW2KEY and HIGH2KEY as domain statistics to generate frequencies“
- CARDF in SYSTABLES. CARDF must be equal or larger than any other cardinalities, such as COLCARDF, FIRSTKEYCARDF, FULLKEYCARDF, and CARDF in SYSCOLDIST
- FREQUENCYF and COLCARDF or CARDF. The number of frequencies collected must be less than or equal to COLCARDF for the column or CARDF for the column group
- FREQUENCYF. The sum of frequencies collected for a column or column group must be less than or equal to 1



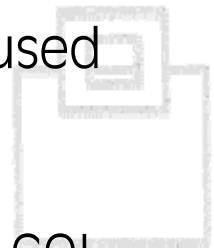
IBM Recommendations DB2 8.1



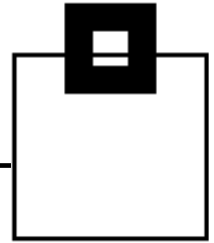
In the „Filter factors and catalog statistics“ chapter
(DB2 Administration Guide)

Recommendation: If query performance is not satisfactory, consider the following actions:

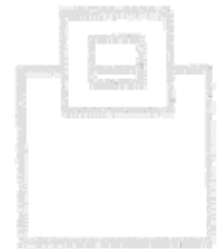
- Collect cardinality statistics on all columns that are used as predicates in a WHERE clause.
- Collect frequencies for all columns with a low cardinality that are used as COL op literal predicates.
- Collect frequencies for a column when the column can contain default data, the default data is skewed, and the column is used as a COL op literal predicate.
- Collect KEYCARD on all candidate indexes.
- Collect column group statistics on all join columns.



AGENDA

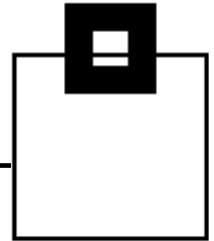


- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- **RUNSTATS walk through**
- SYSCOLDIST explained
- RUNSTATS Q & A
- Statistics HealthCheck
- Notes & Questions



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 7.1 & 8.1



Object Details:

Database	ROYTEST2
Tablespace	ROYTEST2
Table	ROYTEST2
Columns	COL1, COL2, COL3, COL4, COL5 all CHAR(16) NOT NULL
Indexes (All Unique)	COL1, COL2 & COL1, COL3 & COL2 & COL1

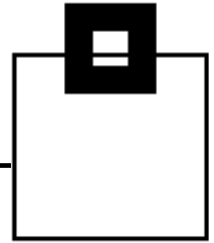


All RUNSTATS run with UPDATE ACCESSPATH.



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 7.1 & 8.1



RUNSTATS ROYTEST2.ROYTEST2

Updates in the DB2 catalog (as listed in the RUNSTATS output):

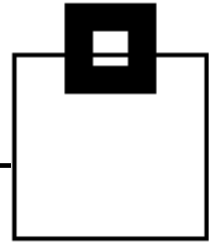
- SYSTABLES
- SYSTABLESPACE

Probably the worst RUNSTATS you could actually do!



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 7.1 & 8.1

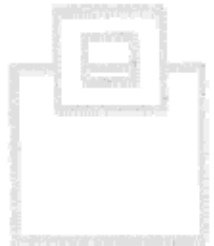


RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL)

Updates in the DB2 catalog (as listed in the RUNSTATS output):

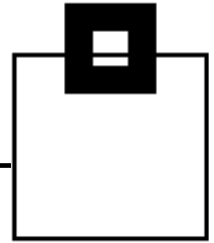
- SYSTABLES
 SYSCOLUMNS - all columns
- SYSTABLESPACE

Still pretty bad!



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 7.1

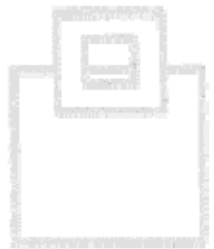


RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL) INDEX(ALL)

Updates in the DB2 catalog (as listed in the RUNSTATS output):

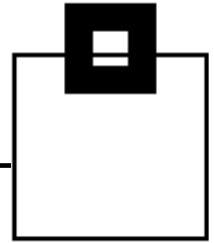
- SYSTABLES
 - SYSCOLUMNS not for COL1
- SYSTABLESPACE
- SYSINDEXES for Index 01
 - SYSCOLUMNS for COL1

- SYSINDEXES Repeated block for indexes 02 & 03
 - SYSCOLUMNS for COL2 & COL3
 - SYSCOLDIST for COL2 & COL3 – Frequencies only



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 8.1



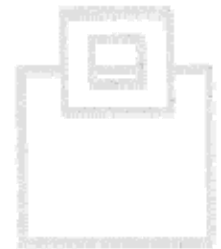
RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL) INDEX(ALL)

Updates in the DB2 catalog (as listed in the RUNSTATS output):

- SYSTABLES
- SYSCOLUMNS not for COL1
- SYSTABLESPACE

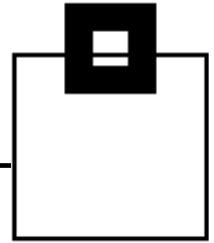
- SYSCOLUMNS for COL1, COL2 & COL3
- SYSINDEXES for Index 01 – 03

- SYSCOLDIST for COL2 & COL3 – Frequencies only



SOFTWARE ENGINEERING

RUNSTATS walk through DB2 7.1

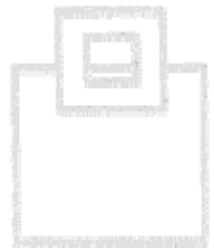
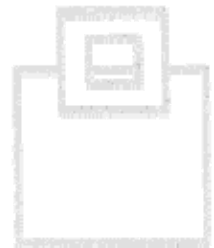


RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL) INDEX(ALL KEYCARD)

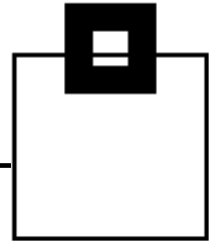
Updates in the DB2 catalog (as listed in the RUNSTATS output):

- SYSTABLES
 - SYSCOLUMNS not for COL1
- SYSTABLESPACE
- SYSINDEXES for Index 01
 - SYSCOLUMNS for COL1

- SYSINDEXES Repeated block for indexes 02 & 03
 - SYSCOLUMNS for COL2 & COL3
 - SYSCOLDIST Frequencies for COL2 & COL3
 - Cardinality for COL3 on COL3, COL2



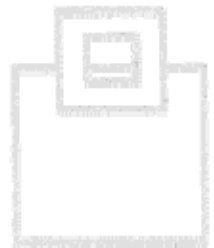
RUNSTATS walk through DB2 8.1



RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL) INDEX(ALL KEYCARD)

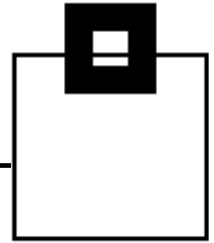
Updates in the DB2 catalog (as listed in the RUNSTATS output):

- SYSTABLES
- SYSCOLUMNS not for COL1
- SYSTABLESPACE
- SYSCOLUMNS for COL1, COL2 & COL3
- SYSINDEXES for Indexes 01, 02 & 03
- SYSCOLDIST Frequencies for COL2 & COL3
Cardinality for COL3 on COL3, COL2



SOFTWARE ENGINEERING

AGENDA

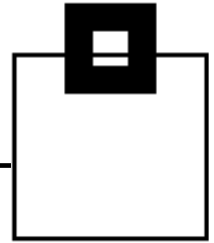


- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- RUNSTATS walk through
- **SYSCOLDIST explained**
- RUNSTATS Q & A
- Statistics HealthCheck
- Notes & Questions



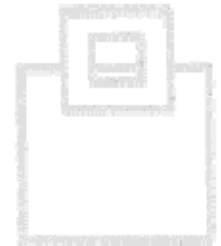
SOFTWARE ENGINEERING

SYSCOLDIST contents explained



The SYSCOLDIST is used for two separate functions:

- Frequencies
- Cardinalities



The column TYPE contains either a „C“ or a „F“ (new in V8 is „N“ for non-padded frequency values)

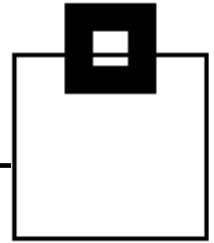


Regardless of the TYPE value the columns TBOWNER, TBNAME, NAME, COLGROUPOCOLNO, NUMCOLUMNS and STATSTIME are used for the same purpose.



SOFTWARE ENGINEERING

SYSCOLDIST contents explained



TBOWNER and **TBNAME** are self explanatory.

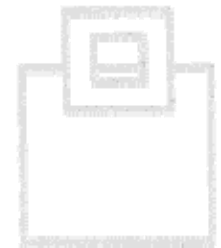
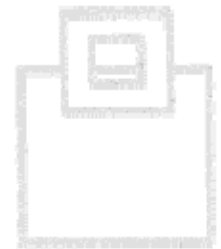
NAME is the always just the first column name.

These three columns are also the non-unique index.

COLGROUPOCOLNO for a single column object is an empty string, for a multi-column object it contains a string of two byte smallint fields which contain the column numbers from the original table.

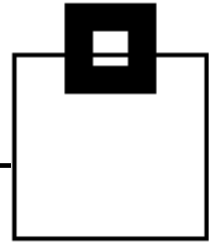
NUMCOLUMNS is the number of columns in this group.

STATSTIME is the time when RUNSTATS inserted this entry. It is also used when there is a complete duplicate so that DB2 uses the last inserted value.



SOFTWARE ENGINEERING

SYSCOLDIST contents explained

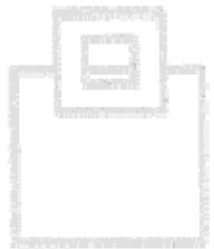


For frequency data the following columns are used:

COLVALUE contains the actual data from the 1 – n columns in the group simply concatenated together. This data might be readable or it might not. Caution must be used when inserting or changing this data as numeric and date, time forms must be the internal DB2 format and not the external format (E.g. The high bit must be flipped for numbers) and watch out for NULLable columns!

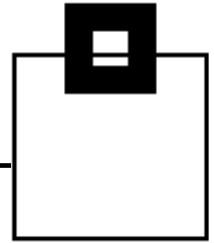
FREQUENCYF contains a floating point value between 0.0 and 1.0 which is the frequency that this value occurs for this group in the table.

If you have V8 you can also use **MOST**, **BOTH**, or **LEAST** on your **RUNSTATS** to collect the most, both, or the least often.



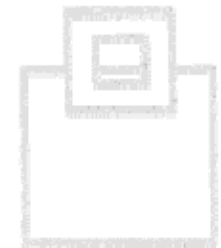
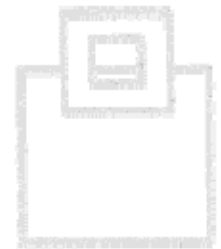
SOFTWARE ENGINEERING

SYSCOLDIST contents explained



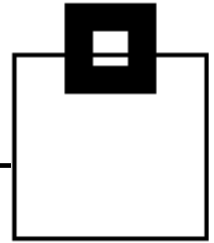
For cardinality data the following column is used:

CARDF contains a floating point value which is how many different values for this group occur in the data.



SOFTWARE ENGINEERING

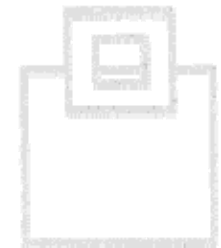
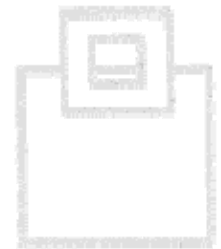
SYSCOLDIST contents explained



SYSCOLDIST in a nut shell

- Frequencies are good for - COL op literal
- Cardinalities are good for - everything!

And remember that a frequency without a cardinality is like a warm beer. Nice to look at, perhaps, but the optimizer will not touch it!



SOFTWARE ENGINEERING

SYSCOLDIST contents explained

In this example you can see the importance of cardinality and frequency. If no frequency data existed in the SYSCOLDIST DB2 would assume that all values are equally distributed.

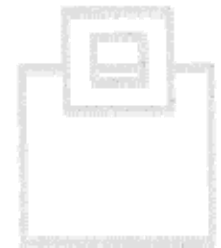
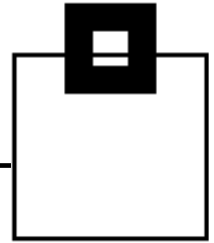
Here the default filter factor would be 1/5 (1/COLCARDF) or 20%

Category	Frequency
Infant	5%
Child	15%
Adolescent	25%
Adult	40%
Senior	15%

This would lead the optimizer to under-estimate by 50% for ADULT and to over-estimate by 400% for INFANT.

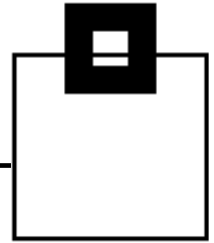
AGENDA

- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- RUNSTATS walk through
- SYSCOLDIST explained
- **RUNSTATS Q & A**
- Statistics HealthCheck
- Notes & Questions



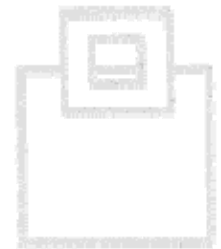
SOFTWARE ENGINEERING

Runstats Q & A



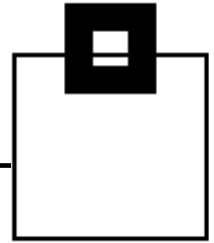
We now know which data is used and where, so now comes a list of RUNSTATS questions:

- `FREQVAL NUMCOLS 3 COUNT 10` – What does this do?
- `FREQVAL NUMCOLS 3 COUNT 0` - What does this do?
- Use of `COLGROUP` in DB2 V8
- Use of `SAMPLE`
- Use of `REOPT(ONCE)`
- Does use of `REORG INDEX` with inline statistics cause problems?
- What happens to frequencies & cardinalities when not specified in the RUNSTATS utility run?



SOFTWARE ENGINEERING

Runstats Q & A



FREQVAL NUMCOLS 3 COUNT 10

According to the DB2 documentation the NUMCOLS is the number of leading index columns to sample so you actually need to give:



RUNSTATS ROYTEST2.ROYTEST2 TABLE(ALL) INDEX(ALL KEYCARD

FREQVAL NUMCOLS 1 COUNT 10

FREQVAL NUMCOLS 2 COUNT 10

FREQVAL NUMCOLS 3 COUNT 10)

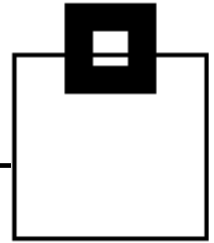


To get the results you expect!



SOFTWARE ENGINEERING

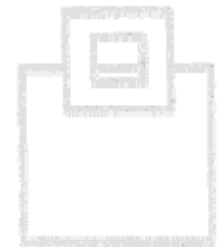
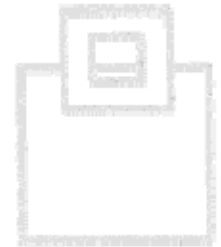
Runstats Q & A



FREQVAL NUMCOLS 3 COUNT 0

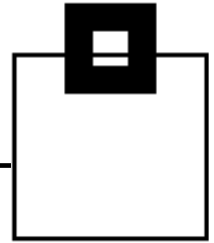
This is more of a „hidden feature“ and it deletes all of the multi-column frequencies.

However it does **NOT** delete any entries created by COLGROUP processing in V8.



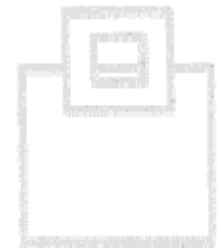
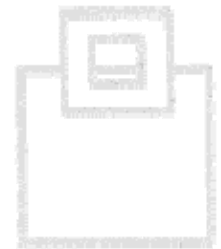
SOFTWARE ENGINEERING

Runstats Q & A



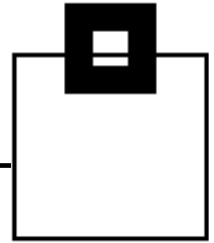
Use of COLGROUP in DB2 V8

This is a very powerful addition to DB2 V8 as it completely removes the need for the old DSTATS program. DSTATS had serious performance problems and column type limitations which made its use complex, time consuming, and error prone.



SOFTWARE ENGINEERING

Runstats Q & A



Use of SAMPLE

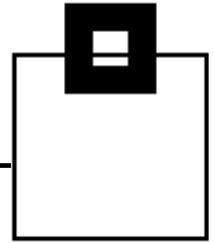
According to the DB2 Utilities documentation this is only for non-indexed columns, however from my tests in V7 and V8 it has a major impact on indexed columns as well.

The problem with SAMPLE is that even using SAMPLE 100 DB2 still gets it wrong...

Col	1	5	10	20	21	22	23	24	25	99	100
TBNAME	62	121	162	178	178	180	180	178	180	180	180
CLNAME	11647	11647	5442	4854	11647	5057	5340	3293	6707	7210	7168

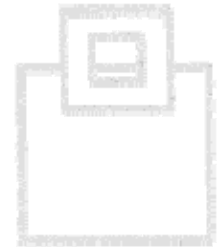
The actual values were 181 and 6871. So be very careful when using this!

Runstats Q & A

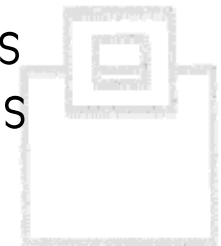


Use of REOPT(ONCE) in DB2 V8

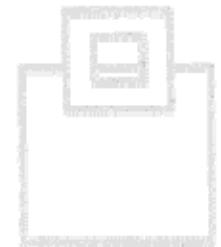
This is a very interesting addition to DB2 V8 as it enables DB2 to do its dynamic SQL mini-bind only once. This can be very good for performance... or not...



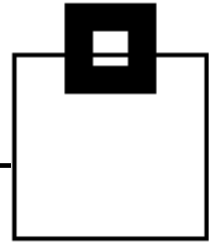
If you are a SAP customer then the change that SAP did in mySAP ERP (ecc5.0) to use the REOPT(ONCE) could be a CPU killer! The problem is that the first run SQL might not actually reflect the normal SQL that is executed over the day. The way out of this problem is to actually delete all frequency records from the SYSCOLDIST for the relevant tables.



SOFTWARE ENGINEERING



Runstats Q & A



Does use of REORG INDEX with inline statistics cause problems?

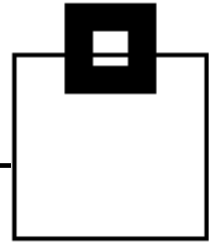
YES!

The problem is that the inline stats will **only** update the index statistics and **not** any of the table ones. This leads, very quickly, to the statistics „drifting apart“. The solution here is either run a tablespace runstats after an index reorg, manually update the table statistics, or do not use inline statistics.



SOFTWARE ENGINEERING

Runstats Q & A



What happens to your frequencies and cardinalities when you do a RUNSTATS run without using FREQVAL and/or KEYCARD?

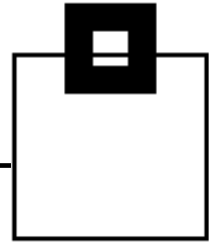
For frequencies the leading column information is replaced (and then the COLGROUPOCOLNO is correctly set to an empty string) and all the other data stays in the catalog. This can be a good and a bad thing! The cardinalities stay there! Again this can be a good and a bad thing!

This means that if you have ever run a RUNSTATS with either FREQVAL or KEYCARD and since then *without* these keywords you probably have „interesting“ data in the SYSCOLDIST.

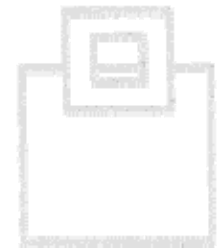
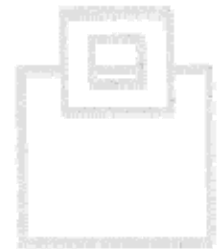


SOFTWARE ENGINEERING

AGENDA

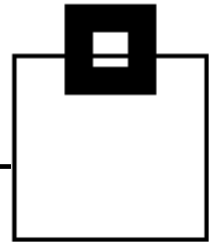


- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- RUNSTATS walk through
- SYSCOLDIST explained
- RUNSTATS Q & A
- **Statistics HealthCheck**
- Notes & Questions

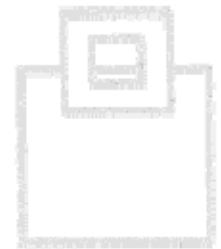


SOFTWARE ENGINEERING

📄 **Statistics HealthCheck**



Now that you have learned everything that RUNSTATS does,
the question is:

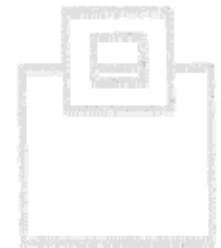


How do I know that the statistics I have are correct?

The answer is:

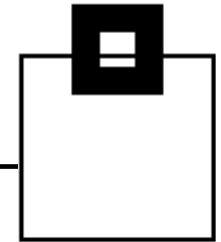


Run 📄 **Statistics HealthCheck**



SOFTWARE ENGINEERING

Statistics HealthCheck

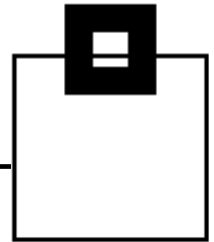


Here is the list of critical violations:

```
Critical violations
Frequency < 0 . . . . . : 0
Frequency > 100 . . . . . : 0
Frequency count > cardf . . . . . : 0
Frequency count > colcardf . . . . . : 0
Frequency sum < 0 . . . . . : 0
Frequency sum > 100 . . . . . : 0
SYSCOLDIST cardf outside allowable range . . . . . : 1
SYSCOLUMNNS - low2key high2key empty . . . . . : 0
SYSCOLUMNNS aggregate colcardf < SYSCOLDIST cardf : 1
SYSCOLUMNNS 1st cardf <> SYSINDEXES firstkeycardf : 0
SYSCOLUMNNS 1st cardf <> SYSINDEXES fullkeycardf : 0
SYSINDEXES 1 col ix firstkeycardf <> fullkeycardf: 0
SYSTABLES cardf < SYSCOLDIST cardf . . . . . : 0
SYSTABLES cardf < SYSCOLUMNNS colcardf . . . . . : 0
SYSTABLES cardf < SYSINDEXES firstkeycardf . . . . . : 0
SYSTABLES cardf < SYSINDEXES fullkeycardf . . . . . : 0
SYSTABLES no RUNSTATS . . . . . : 72
```

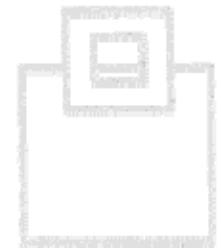
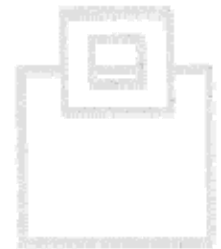


Statistics HealthCheck



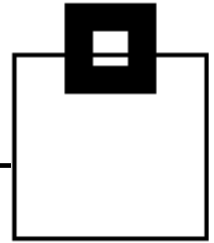
Here is an example from the detail report:

```
VIOLATION  DA.S1N PA.T1N
I - CLUSTERING IX NOT CLUSTERED . : PA.X1NPK
S - MISSING INDEX KEYCARD . . . . : PA.X1N022
S - MISSING INDEX KEYCARD . . . . : PA.X1N042
S - MISSING LEADING COLUMN FREQ . : PA.X4010E ADRESS_STREET
S - MISSING LEADING COLUMN FREQ . : PA.X1N017 LBS
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N017 NUMCOLS 00002
S - MISSING LEADING COLUMN FREQ . : PA.X1N022 POSTCODE
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N022 NUMCOLS 00002
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N022 NUMCOLS 00003
S - MISSING LEADING COLUMN FREQ . : PA.X1N042 ADRESS_CITY
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N042 NUMCOLS 00002
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N042 NUMCOLS 00003
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N042 NUMCOLS 00004
W - MISSING INDEX FREQVAL NUMCOLS : PA.X1N042 NUMCOLS 00005
```



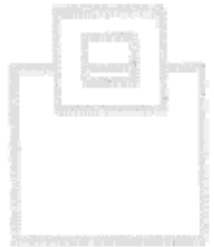
SOFTWARE ENGINEERING

Garbage in the SYSCOLDIST



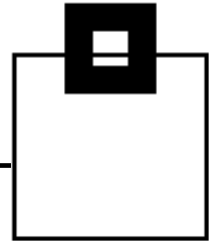
A colleague is on holiday and the telephone rings with end-users or management saying that the system performance is bad. You grab your RUNSTATS JCL, submit the job and all is well.

Or is it?



SOFTWARE ENGINEERING

Garbage in the SYSCOLDIST



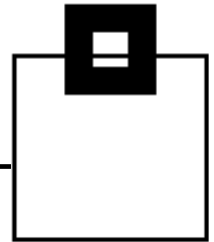
Here is the query that you have found:

```
SELECT A.ICTYPE , A.TIMESTAMP
FROM SYSIBM.SYSCOPY A
WHERE A.ICTYPE IN ('D','I','F','S','W','Y','R','X','Z')
      AND A.ICBACKUP NOT IN ('RP','RB')
      AND A.DBNAME = ?
      AND A.TSNAME = ?
      AND (A.DSNUM = ? OR (? > 0 AND A.DSNUM = 0))
      AND A.TIMESTAMP = (SELECT MAX(B.TIMESTAMP)
                          FROM SYSIBM.SYSCOPY B
                          WHERE B.ICTYPE IN
                                ('D','I','F','S','W','Y','R','X','Z')
                                AND B.ICBACKUP NOT IN ('LB','RP','RB')
                                AND B.DBNAME = ?
                                AND B.TSNAME = ?
                                AND (B.DSNUM = ? OR (? > 0 AND B.DSNUM = 0))
                          )
WITH UR
;
```



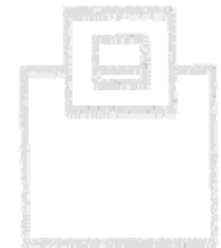
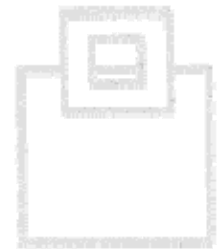
SOFTWARE ENGINEERING

Garbage in the SYSCOLDIST

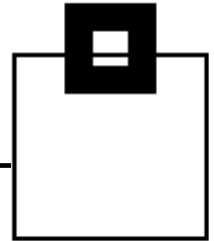


This is the actual RUNSTATS you should run...

```
RUNSTATS TABLESPACE DSNDB06.SYSCOPY
TABLE (SYSIBM.SYSCOPY)
COLGROUP (ICTYPE, TIMESTAMP, TSNAME)
COLGROUP (DBNAME, TIMESTAMP)
COLGROUP (ICTYPE, TSNAME)
COLGROUP (DBNAME, TIMESTAMP, TSNAME)
COLGROUP (TIMESTAMP, TSNAME)
COLGROUP (ICBACKUP) FREQVAL COUNT 10
COLGROUP (ICTYPE) FREQVAL COUNT 1
COLGROUP (DBNAME, ICTYPE, TIMESTAMP, TSNAME)
COLGROUP (ICBACKUP, TSNAME)
COLGROUP (DBNAME, ICBACKUP, TSNAME)
COLGROUP (ICBACKUP, TIMESTAMP, TSNAME)
COLGROUP (DBNAME, ICTYPE, TSNAME)
COLGROUP (ICTYPE, TIMESTAMP)
COLGROUP (DSNUM) FREQVAL COUNT 1
COLGROUP (DBNAME, ICTYPE, TIMESTAMP)
COLGROUP (DBNAME, ICBACKUP, TIMESTAMP)
COLGROUP (ICBACKUP, TIMESTAMP)
COLGROUP (DBNAME, ICBACKUP, TIMESTAMP, TSNAME)
SORTDEVT SYSDA
SHRLEVEL CHANGE REPORT YES
```

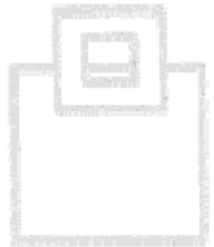


Garbage in the SYSCOLDIST



This is the actual RUNSTATS you should run...

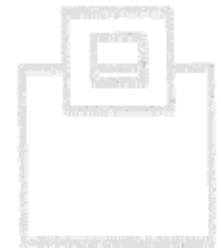
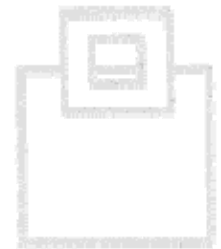
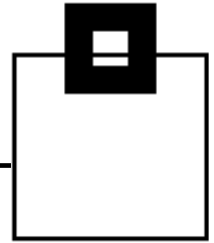
```
RUNSTATS TABLESPACE DSNDB06.SYSCOPY  
TABLE (SYSIBM.SYSCOPY)  
COLGROUP (ICBACKUP) FREQVAL COUNT 10  
COLGROUP (ICTYPE) FREQVAL COUNT 1  
COLGROUP (DSNUM) FREQVAL COUNT 1  
SORTDEVT SYSDA  
SHRLEVEL CHANGE REPORT YES
```



SOFTWARE ENGINEERING

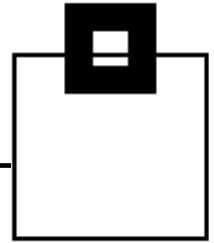
AGENDA

- RUNSTATS basics and DB2 Optimizer relevant statistics
- IBM recommendations through the ages
- RUNSTATS walk through
- SYSCOLDIST explained
- RUNSTATS Q & A
- Statistics HealthCheck
- **Notes & Questions**



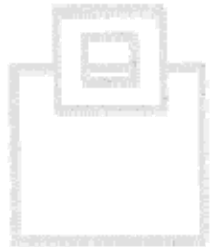
SOFTWARE ENGINEERING

Notes and Future development



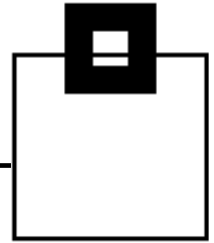
The  **Statistics HealthCheck** is freeware and so has certain limitations:

- A maximum of 100 indexes on any one table
- A maximum of 2000 SYSCOLDIST Cardinality entries for a table
- A maximum of 2000 SYSCOLDIST Frequency entries for a table
- No unicode table names, index names, or column names (long is OK!)



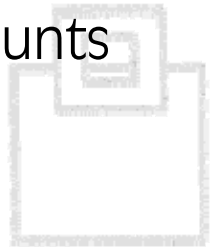
SOFTWARE ENGINEERING

Notes and Future development



Some possible future enhancements:

1. Add optional link to the Real-time Statistics tables to check that TOTALROWS and TOTALENTRIES is correctly mirrored in the CARD counts
2. SAP support
3. Direct link to our **Maintain** thresholds for RSA etc
4. Better input parameters for DB's to check or exclude
5. Generate corrective RUNSTATS statements
6. Generate corrective EXEC SQL statements for COLCARDF and SAMPLE problem

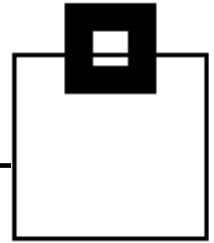


Go to www.seg.de and then the **Statistics HealthCheck** forum for more details.



SOFTWARE ENGINEERING

Questions for you



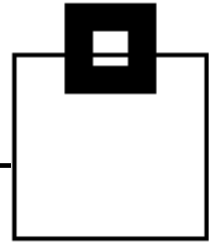
Now I have some questions for you all!

- How many of you use SAP on z/OS?
- How many of you use the DSNACC.EXCEPT_TBL to limit RUNSTATS?
- How many of you use a third-party or home-made tool to limit the execution of RUNSTATS?
- How many of you manipulate statistics in the Catalog?



SOFTWARE ENGINEERING

The End



Many thanks for your attention and if you have any questions I will be pleased to answer them!



By the way, you are all

official

RUNSTATS Masters!



SOFTWARE ENGINEERING

