

BUFFERPOOL Tuning The Next Generation

Roy Boxwell, Software Engineering GmbH

P06







Agenda

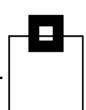
- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A







Agenda



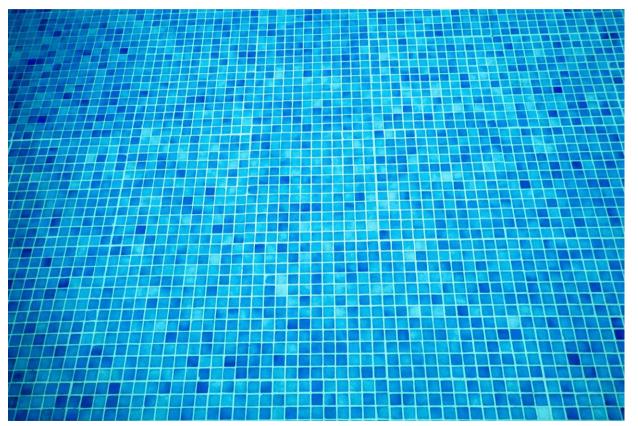
- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A







Bufferpools:









Real bufferpools (BPs) have existed since the get-go of DB2 (when the B was big!)

The idea, back in the day, was to have two sizes of BP for matching the two sizes of DB2 pages - namely 4KB and 32KB. DB2 started *very* small - we only had FOUR pools! Three 4KB and a single 32KB:

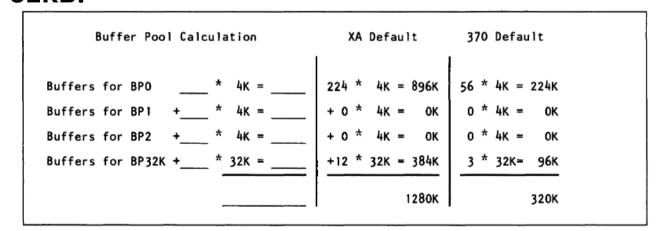


Figure 29. Buffer Pool Size Calculation









In DB2 V3.1 we got the boost of BP0 to BP49 and BP32K, BP32K1 to BP32K9. BP0 had 2,000 pages as a default and BP32K only 24(!) pages as a default...



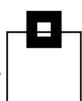
In DB2 V6.1 IBM introduced the zero buffers... BP8K0 – BP8K9 and BP16K0 – BP16K9 thus endearing themselves into the heart of all future DBAs!



In DB2 V8.1 BP0 default jumped to 20,000, BP8K0 was 1,000, BP16K0 was 500 and BP32K was 250



In Db2 11 the B went lower case...

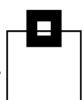


In fact, the defaults have not changed at all so they must be good and stable, yes?









In fact, the defaults have not changed at all so they must be good and stable, yes?

```
DSNTTP1
                 INSTALL DB2 - BUFFER POOL SIZES - PANEL 1
===>
Enter 4 KB buffer pool sizes in number of pages.
           ==> 20000
                            18 BP17 ==> 0
                                                      35 BP34 ==> 0
    2 BP1
           ==> 0
                             19 BP18 ==> 0
                                                      36 BP35 ==> 0
                                                      37 BP36 ==> 0
           ==> 0
                             20 BP19 ==> 0
           ==> 0
                            22 BP21 ==> 0
           ==> 0
                                                      39 BP38 ==> 0
                            23 BP22 ==> 0
           ==> O
                                                      40 BP39 ==> 0
                            24 BP23 ==> 0
           ==> 0
                                                      41 BP40 ==> 0
                            25 BP24 ==> 0
                                                      42 BP41 ==> 0
           ==> 0
                            26 BP25 ==> 0
                                                      43 BP42 ==> 0
   10 BP9 ==> 0
                                                      44 BP43 ==> 0
                                                      45 BP44 ==> 0
   12 RP11 ==> 0
                            29 BP28 ==> 0
                                                      46 BP45 ==> 0
  13 BP12 ==> 0
                            30 BP29 ==> 0
                                                      47 BP46 ==> 0
                                                      48 BP47 ==> 0
   14 BP13 ==> 0
                            31 BP30 ==> 0
  15 BP14 ==> 0
                            32 BP31 ==> 0
                                                      49 BP48 ==> 0
  16 BP15 ==> 0
                            33 BP32 ==> 0
                                                      50 BP49 ==> 0
                            34 BP33 ==> 0
   17 BP16 ==> 0
PRESS: ENTER to continue
                            RETURN to exit
                                             HELP for more information
```









In fact, the defaults have not changed at all so they must be good and stable, yes?

```
DSNTIP1
                 INSTALL DB2 - BUFFER POOL SIZES - PANEL 1
===>
                      DSNTIP2
                                        INSTALL DB2 - BUFFER POOL SIZES - PANEL 2
Enter 4 KB buffer poo
          ==> 20000
                      ===>
    1 BP0
    2 BP1
           ==> 0
                      Enter 8 KB, 16KB, and 32 KB buffer pool sizes in number of pages.
    3 BP2
           ==> O
                           1 BP8K0 ==> 2000
                                                     11 BP16K0 ==> 500
           ==> 0
                                                                                  21 BP32K
                                                                                            ==> 250
                           2 BP8K1 ==> 0
                                                     12 BP16K1 ==> 0
                                                                                 22 BP32K1 ==> 0
           ==> O
   6 BP5
           ==> O
                           3 BP8K2 ==> 0
                                                     13 BP16K2 ==> 0
                                                                                 23 BP32K2 ==> 0
          ==> 0
                           4 BP8K3 ==> 0
                                                     14 BP16K3 ==> 0
                                                                                 24 BP32K3 ==> 0
    8 BP7
          ==> 0
                                                     15 BP16K4 ==> 0
                                                                                 25 BP32K4 ==> 0
                           5 BP8K4 ==> 0
                                                                                 26 BP32K5 ==> 0
                           6 BP8K5 ==> 0
                                                     16 BP16K5 ==> 0
          ==> 0
                           7 BP8K6 ==> 0
                                                     17 BP16K6 ==> 0
                                                                                 27 BP32K6 ==> 0
                           8 BP8K7 ==> 0
                                                     18 BP16K7 ==> 0
                                                                                 28 BP32K7 ==> 0
   12 RP11 ==> 0
                           9 BP8K8 ==> 0
                                                     19 BP16K8 ==> 0
                                                                                 29 BP32K8 ==> 0
  13 BP12 ==> 0
                          10 BP8K9 ==> 0
                                                     20 BP16K9 ==> 0
                                                                                  30 BP32K9 ==> 0
  14 BP13 ==> 0
  15 BP14 ==> 0
                                      4-KB BUFFFR POOL FOR USFR DATA ===> BP1
                                                                                        BP0
                                                                                                - BP49
  16 BP15 ==> 0
                                      8-KB BUFFER POOL FOR USER DATA
                                                                                        BP8K0

    BP8K9

  17 BP16 ==> 0
                                     16-KB BUFFER POOL FOR USER DATA

    BP16K9

                                                                                        BP16K0
                                     32-KB BUFFER POOL FOR USER DATA ===> BP32K

    BP32K9

PRESS: ENTER to cont
                          35 DEFAULT BUFFER POOL FOR USER LOB DATA

    BP32K9

                          36 DEFAULT BUFFER POOL FOR USER XML DATA
                                                                                        BP16K0 - BP16K9
                          37 DEFAULT BUFFER POOL FOR USER INDEXES
                                                                       ===> BP0

    BP32K9

                                                                      HELP for more information
                              ENTER to continue
                                                    RETURN to exit
```









Ok, so we have a nice set of default sizes but what are they used for?



Well, even today on our "fake" DASD there is a nasty thing called I/O and I/O is slow!



The best I/O does not cause a disk seek at all, in fact that is the entire point of a BP. A piece of required data is found in memory so that the data is instantly available to the application process.





Why do we have so many BPs in Db2 for z/OS? Why not just one huge area of RAM stuffed full of data?



Well, the answer to that is "Horses for Courses"

The performance of any given BP is strongly related to the applications running and using it. Think of a process that is sequentially reading through a table for summation purposes. It reads data but will *never* want to read it again.



Is this "good" for the BP?





Why do we have so many BPs in Db2 for z/OS? Why not just one huge area of RAM stuffed full of data?



Well, the answer to that is "Horses for Courses"

The performance of any given BP is strongly related to the applications running and using it. Think of a process that is sequentially reading through a table for summation purposes. It reads data but will *never* want to read it again.



Is this "good" for the BP? Nope.



Now imagine an application process that is randomly reading data through an index. It fetches the leaf and non-leaf pages into the BP as it needs them and then carries on.



When it needs that "used" leaf page it will find it again in the BP. Is this a "good" use of the BP?





Now imagine an application process that is randomly reading data through an index. It fetches the leaf and non-leaf pages into the BP as it needs them and then carries on.



When it needs that "used" leaf page it will find it again in the BP. Is this a "good" use of the BP? Yes.





Now imagine an application process that is randomly reading data through an index. It fetches the leaf and non-leaf pages into the BP as it needs them and then carries on.



When it needs that "used" leaf page it will find it again in the BP. Is this a "good" use of the BP? Yes.



If both of these applications share the <u>same</u> BP this is obviously not good, but this is what most, if not all, Db2 shops do!





Now imagine a sort – Yes I am talking about DSNDB07 usage here! You might not know it but SORT requires a BP as well. What are the odds of a repeated reread in a sort pool?



You can imagine they are pretty low!

Sort should *always* be in its own little/large pool with VPSEQT set to 99%





What about LRU, FIFO, NONE when stealing pages? "It depends" raises its ugly head here!



LRU is least recently used - so that the "stalest" pages can be got rid off and replaced with newer ones and is, naturally, the default.



NONE is great for "permanent in-memory data" Xref tables etc.

FIFO is great if you really do not care about the least recently used logic. It is really a niche use case though.





Most shops have no personnel to look at, monitor, tune and change the BPs at their site. The expertise is "graying" and lots of people are "afraid" of changing something so crucial as a BP.

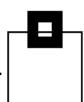


This is not healthy!

BP usage always changes over time! (Death by cut-and-paste is a classic...)







The Db2 Directory and Catalog are the most important part of any Db2 system. They are the meta-data repositories of the entire system and contain *everything* you need to run, but these objects go into only the listed default BPs and I will bet that most shops also use these BPs for application data – This I call BP Pollution.

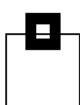


The first thing you must do is move all non-Db2 objects *out* of BP0, BP8K0, BP16K0 and BP32K.



This can be tough, but it must be done! It is useless starting to tune BPs when their usage is completely broken!





A standard set of "Rules of Thumb" is:

- 1) The Db2 Directory and Catalog on their own
- 2) Application tablespaces and indexes kept apart
- 3) LOB and XML on their own
- 4) Sort on its own
- 5) In-memory tables (PGSTEAL(NONE) on their own)
- 6) Randomly accessed data on their own
- 7) Sequentially accessed data on their own
- 8) GBP correctly sized
- 9) The rest...







Agenda

- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A





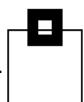


The answer is a massive









The answer is a massive











It is often written that the quickest and best results of any type of tuning are indeed with BP and GBPs.



SQL Tuning is still, obviously, required but the big system wide ROI can be had in the BP and GBP area.





Using large frame sizes (1Mb or 2GB) saves up to 4%



If you can imagine a system where it is actively paging, actively reading and rereading data and index pages all the time just because it cannot find the data in the BP it is clear that you can



"Tune the SQL until you die, it will not get faster!"







Another very popular problem is the:

"Stuff it in BP8K0, I know that BP exists!"



This is especially popular for COMPRESS YES indexes – Thus solving one problem and introducing another, even worse, problem at the same time!



Naturally, the BP problem is not seen and everyone wonders why performance tanks every now and again...



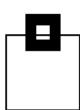
Agenda

- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A









There is an ALTER command...









There is an ALTER command...

Now in Db2 for LUW you would just throw a few more CPUs or GBs of RAM at the problem and walk away a happy bunny. On z/OS it is a little bit more tricky but I know of shops that have basically done the same!









There is an ALTER command...

Now in Db2 for LUW you would just throw a few more CPUs or GBs of RAM at the problem and walk away a happy bunny. On z/OS it is a little bit more tricky but I know of shops that have basically done the same!



You must find the current state of your BPs.







There is an ALTER command...

Now in Db2 for LUW you would just throw a few more CPUs or GBs of RAM at the problem and walk away a happy bunny. On z/OS it is a little bit more tricky but I know of shops that have basically done the same!



You must find the current state of your BPs.



You must then ALTER them to do what they should be doing!



Agenda

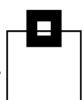
- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A







What about GLOBAL BUFFERPOOLS?



These are the forgotten "zombies" of the Db2 for z/OS BP world!

They are incredibly important for the well-running of any datasharing system and basically require the same style of tuning as normal – local – BPs.

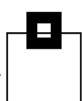


They have sizes and thresholds just like local BPs but they have normally been completely forgotten about in most shops!





What about GLOBAL BUFFERPOOLS?



These are the forgotten "zombies" of the Db2 for z/OS BP world!

They are incredibly important for the well-running of any datasharing system and basically require the same style of tuning as normal – local – BPs.



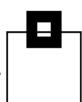
They have sizes and thresholds just like local BPs but they have normally been completely forgotten about in most shops!



You must find the current state of your Global BPs.



What about GLOBAL BUFFERPOOLS?



These are the forgotten "zombies" of the Db2 for z/OS BP world!

They are incredibly important for the well-running of any datasharing system and basically require the same style of tuning as normal – local – BPs.



They have sizes and thresholds just like local BPs but they have normally been completely forgotten about in most shops!



You must find the current state of your Global BPs.



You must then ALTER them to do what they should be doing!

Agenda

- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A









We now know that we have a problem!

How can we actually visualize this and do stuff?



The archaic way that 3270 green screen outputs data and/or the extremely detailed obscure formulae that must be used and/or the different sources of data that must be trawled, all make it "non trivial"...







We now know that we have a problem!

How can we actually visualize this and do stuff?



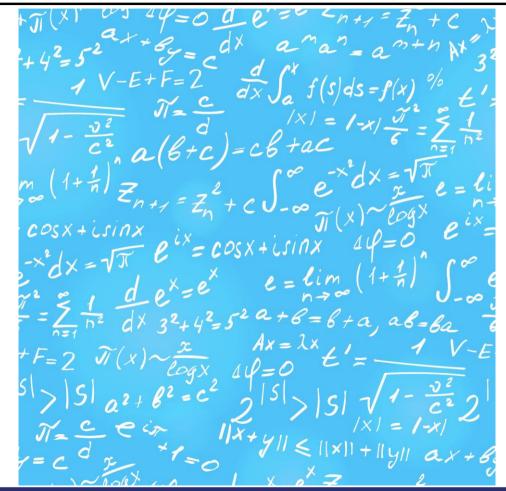
The archaic way that 3270 green screen outputs data and/or the extremely detailed obscure formulae that must be used and/or the different sources of data that must be trawled, all make it "non trivial".



How "non trivial"???













You have to get the correct info from 100's of metrics to then calculate and check all the performance-relevant thresholds:

Frame boundary

Frame sizing

Frame size with LFAREA

System residency

Random residency

Sequential residency

PGFIX(NO) used

Page-ins for read required

Page-ins for write required

DMTH Threshold hit

Prefetch disabled







You have to get the correct info from 100's of metrics to then calculate and check all the performance-relevant thresholds:

Random Sync I/O

Bufferpool paging

System Hit Ratio

Application Hit Ratio

No. of page updates for each page written

No. of pages written for each write I/O

Page arrival rate

Prefetch size

VPSEQT should be changed

DWQT hit rate / Second

VDWQT hit rate / Second







You have to get the correct info from 100's of metrics to then calculate and check all the performance-relevant thresholds:

GBP Hit ratio

GBP Writes failed

GBP Reclaims for directory entries

GBP Cross Invalidations (XI) due to directory reclaims

GBP Snapshot

GBP Castout

GBP Sizing and Duplexing

GBP Sync read XI miss ratio - high

GBP Sync read XI data not returned per day - high

GBP Sync read XI data not returned per second – high







Not only that, but you have to know which threshold and/or which value to ALTER to actually do the corrective action!



This is all pretty nasty work that someone probably did way back in the 1990's but since then it has not been updated...







Not only that, but you have to know which threshold and/or which value to ALTER to actually do the corrective action!

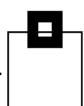


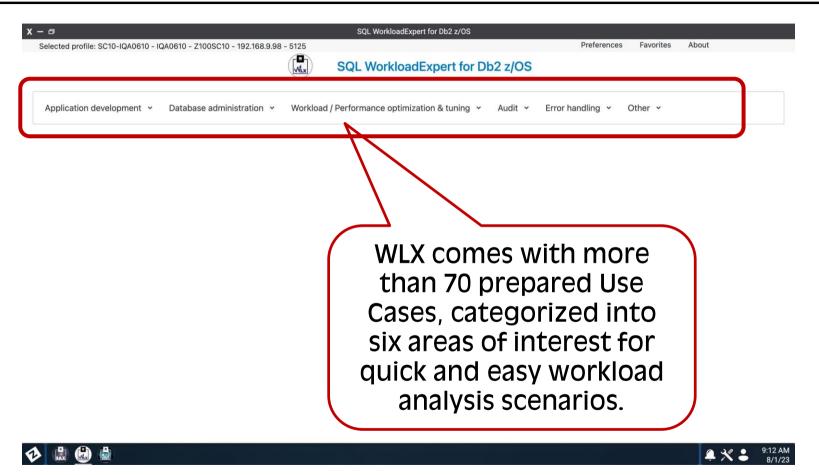
This is all pretty nasty work that someone probably did way back in the 1990's but since then it has not been updated...







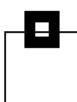


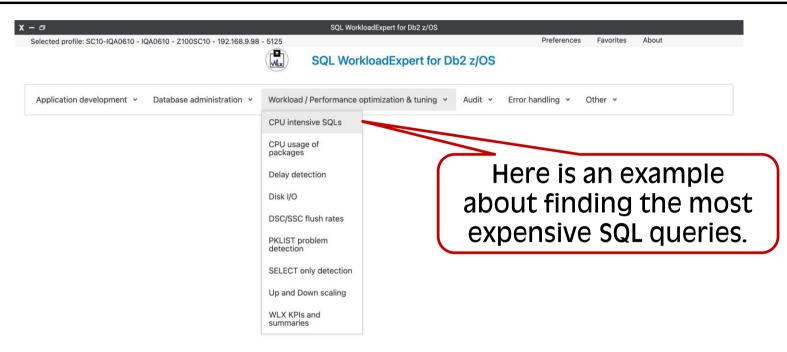














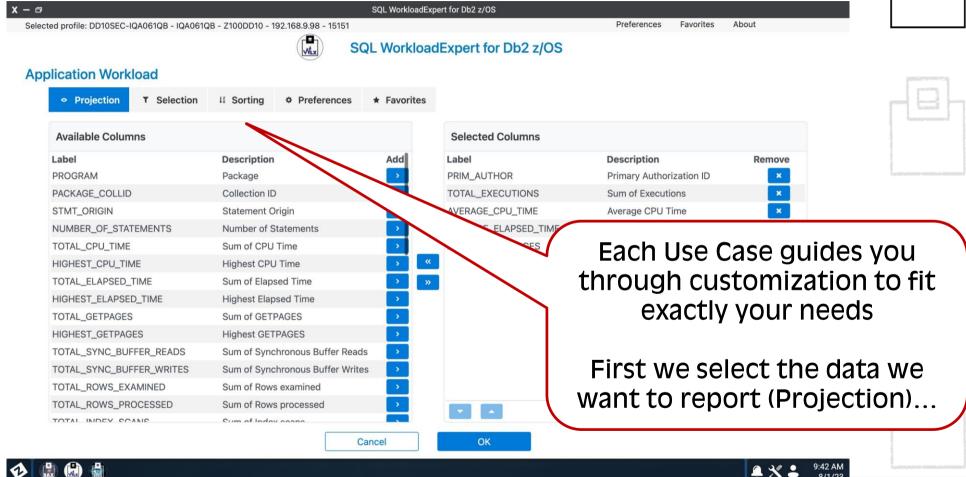






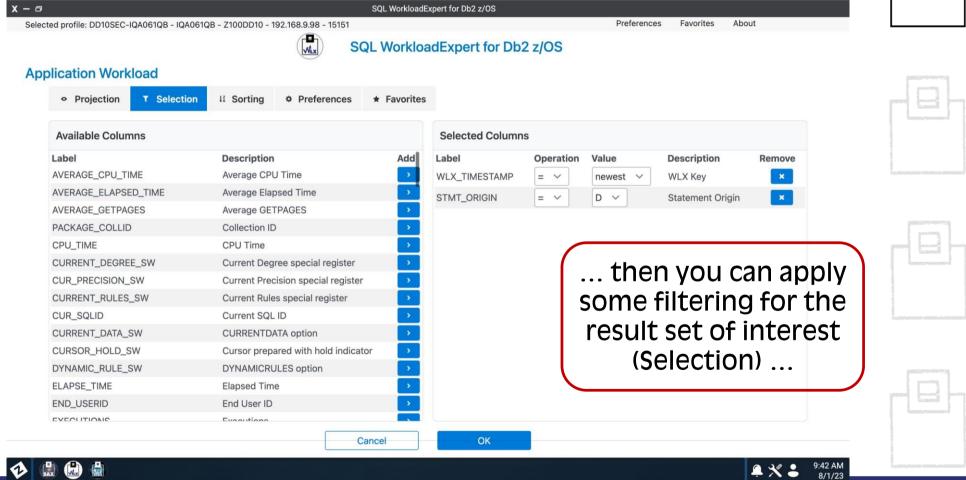
© 2023 SOFTWARE ENGINEERING GMBH





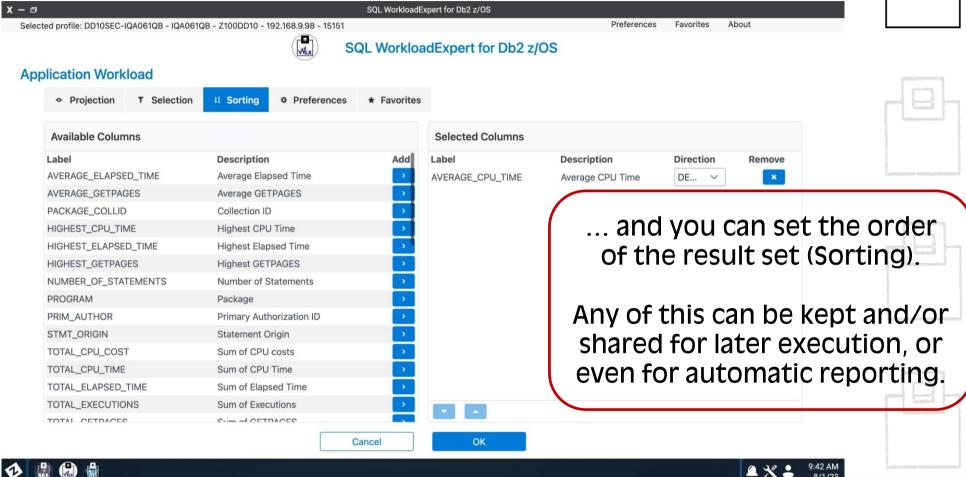
© 2023 SOFTWARE ENGINEERING GMBH





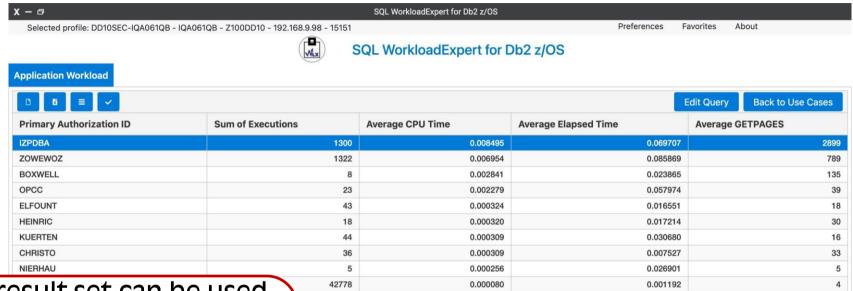
© 2023 SOFTWARE ENGINEERING GMBH





© 2023 SOFTWARE ENGINEERING GMBH





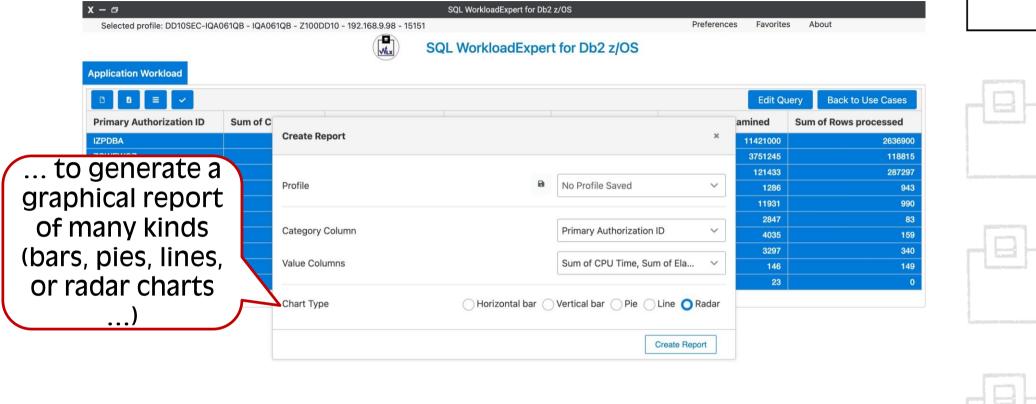
The result set can be used for further in-depth drill down analysis, cross-reference reporting, exporting into a pdf/excel, or





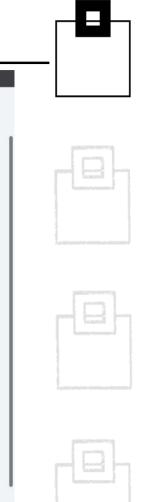
9:41 A

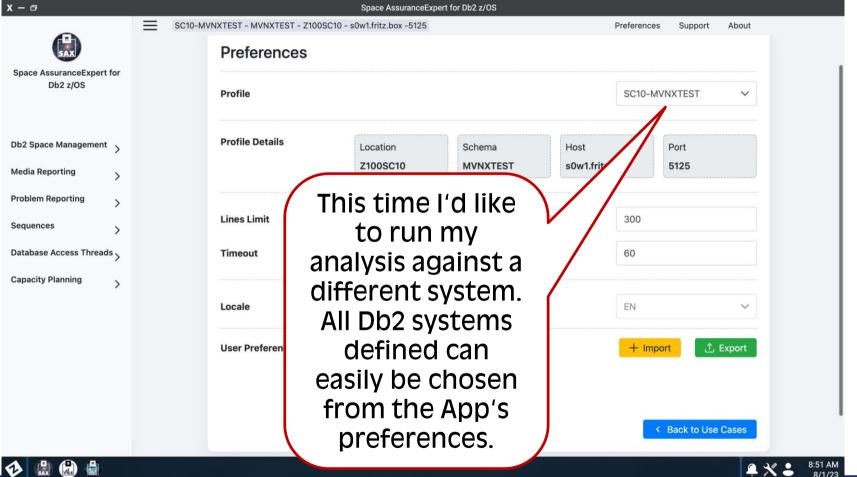






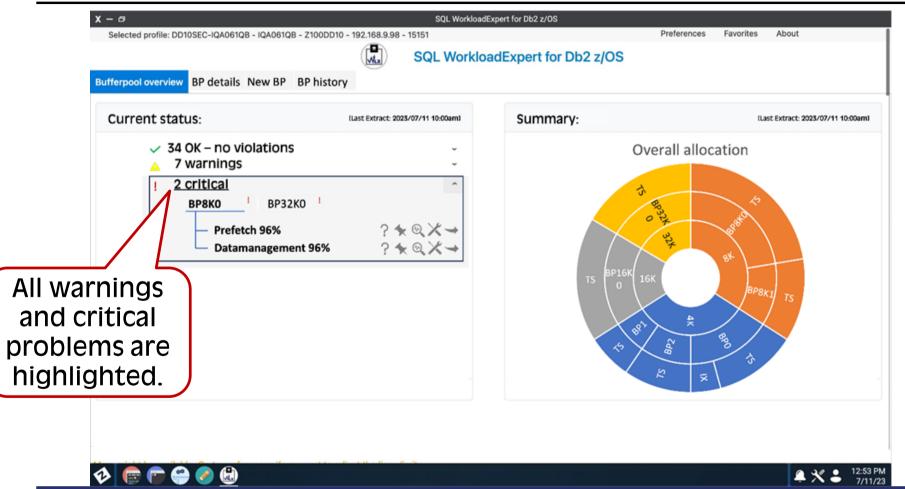






© 2023 SOFTWARE ENGINEERING GMBH 52



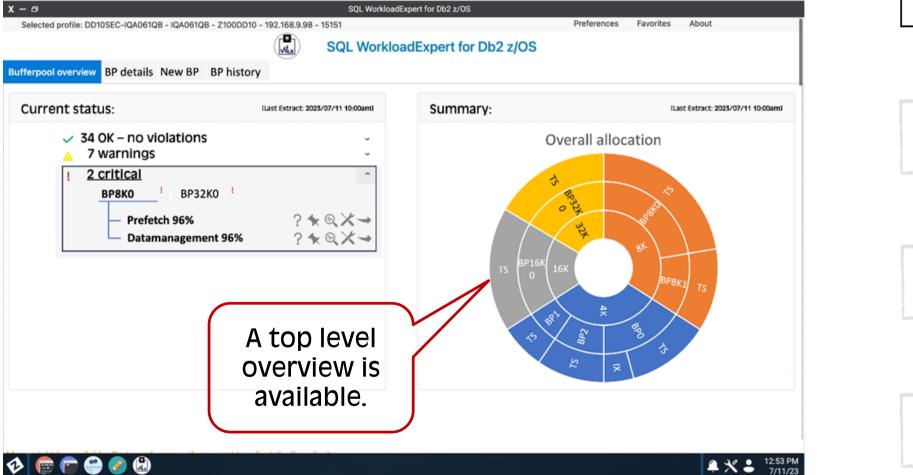




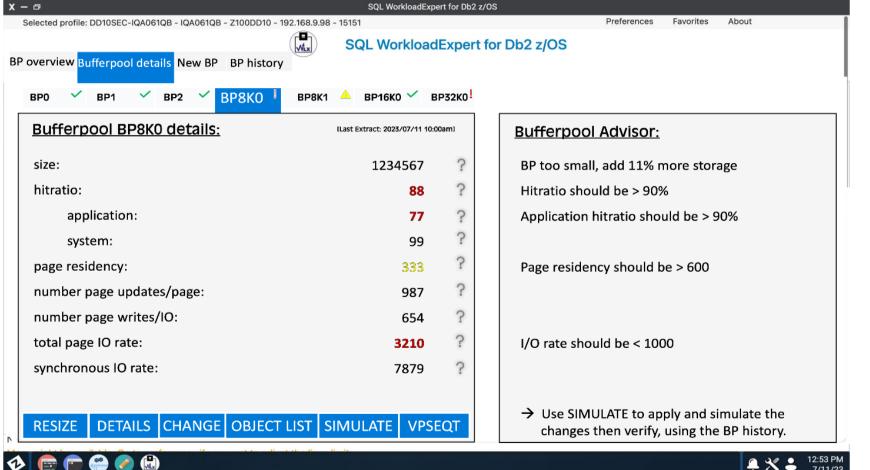










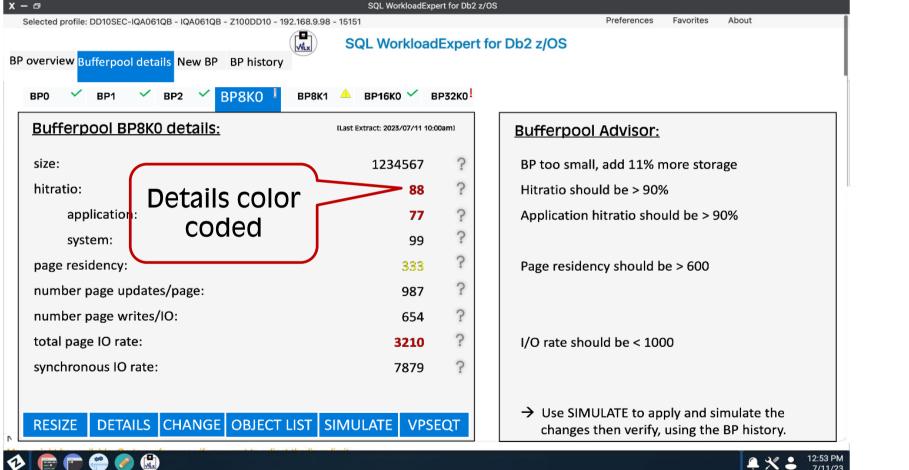






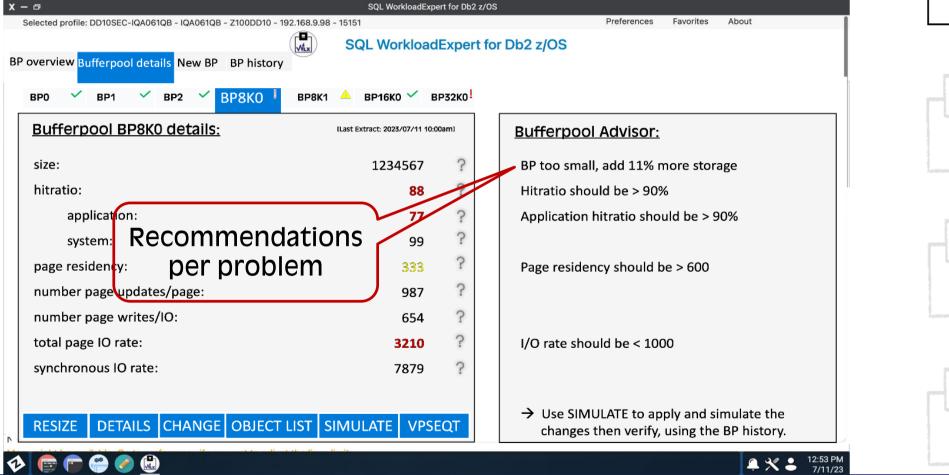






7/11/23





With one click you generate all of the ALTERs you require and, naturally, we do not execute them!

You can "batch up" the changes and issue them at a quiet time.

Please remember:

Measure your system first.

Do the ALTERs – Perhaps not *all* of them at once!

Measure your system.

You *should* see a system-wide improvement and/or application-level improvement.

This is all done *without* even knowing the applications that are running!







Agenda

- What use are BUFFERPOOLs?
- Is it worth tuning?
- How do you tune them?
- What about GLOBAL BUFFERPOOLS?
- The modern way to visualize BP/GBP problems
- Q&A







Questions & Answers









