The Secrets Behind DB Startup Parameters

Paul Koufalis, White Star Software

pk@wss.com

- So many parameters – what do they mean?
- Which should I use?
- Which shouldn’t I use?

- What are the secret magic values?

White Star Software
www.wss.com

Who is this Paul Koufalis?

- Progress DBA and UNIX sysadmin since 1994
- Expert consulting related to all technical aspects of Progress and OpenEdge

- Wide range of experience
  - Small 10 person offices to 1500+ concurrent users
  - AIX, HPUX, Linux, Windows…if Progress runs on it, I've worked on it

White Star Software
www.wss.com

- The oldest and most respected independent DBA consulting firm in the world
- Four of the world's top OpenEdge DBAs
- Author of ProTop, the #1 FREE OpenEdge Database Monitoring Tool
  - http://dbappraise.com/protop.html

Agenda

- The Must-Haves: Screw these up and expect a call from Inigo Montoya*
- The Freebies: Like finding money on the sidewalk. Maybe it’s $1 or maybe it’s $100
- The fancy-pants: Well aren’t you Mr. DBA now using these parameters?
- The ??? parameters: People use these? Sometimes…
- The Gang-O-Useless: **Don’t** raise your hand if they look familiar!

DISCLAIMER

- This information is *PROBABLY MOSTLY* applicable to you

DON’T BE DUMB – DO YOUR OWN TESTING

White Star Software
www.wss.com

The Must Haves

White Star Software
www.wss.com

Database Buffers (-B)

- The DBA equivalent of buying flowers for your wife
- More is better…to a point
- Big -B transfers effect of bad code from disk to CPU
- Buffer Hit % can be misleading (95% is 5X worse than 99%)

Recommendations

- Keep the working set in memory
- Tuning requires LARGE increases to –B
  - 50K to 55K is not going to change anything
- Fix the code…PLEASE!!

White Star Software
www.wss.com

Spin and LRU Skips

- No brainer parameters
- Typically don’t require much *tuning*: just use them

Recommendations

- -spin: π * date of birth of DBA (h/t Dan F.)
- -lruskips 100

White Star Software
www.wss.com

Lock Table (-L)

- Total number of locks in the database across all users
- Memory cost is low: +/- 72 bytes per entry: -L 100K = 7 Mg
- Highly application dependant
- VERY high numbers in small/medium environments make me suspicious

Recommendations

- Monitor high water mark
- Increase # as HWM approaches –L value
- Watch out for sudden jump-ups after code promotion, M&A activity

User Connections

- -n: number of connections (NOT users)
- -Mi/Ma/Mn/Mpb: Network broker connection parameters
- -m3/ServerType: SQL vs 4GL server

Recommendations

- -n: not directly correlated to licence count so give yourself a good buffer
- -Mi 1 –Ma 5 is a good start
- -Mpb = max concurrent users (4GL or SQL) / -Ma plus a few
- -Mn = Sum of –Mpb plus a generous few extra
- -ServerType: segregate 4GL and SQL connections on separate –m3 brokers

Replication Buffers (-pica)

- If you are using OpenEdge Replication the pica parameter is ultra important
- If you fill pica you throttle all DB updates

Recommendations

- Use the max value for your version
  - 256 (9.1E, 10.0B, 10.1A)
    1024 (9.1e04)
    512 (10.0b05)
    8192 (10.1A02, 10.1B01)
  - *Until* 10.2B08
- 10.2B08/11.2+ the max value is 1M
  - Start with 32K

DB Structure

- OK – not technically start-up parameters!

Recommendations

- DB block size: 4 or 8 Kb
- AI/BI block size: 16 Kb
- BI Cluster Size: 4 Mg and up (load dependant)
- Variable length extents
  - Yes there is a “cost” : a few milliseconds a day
  - In most cases the management benefit outweighs the negligible cost

Helper Processes

- AIW/BIW/APW/WDOG
- These *should* be start-up parameters
  - Currently you still need to start them manually

Recommendations

- ONE of each
- Rare is the site that needs 2+ APW

The Freebies

White Star Software
www.wss.com

AI and BI Buffers

- The default value used to be something silly like 5
- Watch for “Empty Buffer Waits”

Recommendations

- 50 is probably good enough for most of you
- 100 almost definitely is
- Cost is negligible: 100 X 16 Kb X 2 (AI and BI) = 3.2 Mg

White Star Software
www.wss.com

AI and BI Stall and BI Threshold

- Why crash? Freeze instead
- -aistall: quiet the database when you run out of AI files
- -bistall: quiet the database when you hit the BI threshold size
- -bithold: Max size in Mg of BI file before emergency shutdown or stall

Recommendations

- Always use all three
- Set –bithold to about double your normal BI HWM
  - I.e. if your variable length BI file is normally 2 Gb, use –bithold 5000
  - Watch out for month-end/year-end type processing
- CAREFUL: You **ABSOLUTELY** need monitoring to alert you (ProTop)

White Star Software
www.wss.com

PIN Shared Memory (-pinshm)

- Another no-brainer: why would you ever want to swap out DB shared memory?
- Ignored on Windows and AIX
- Not as important as in the past as most servers have lots of memory

Recommendations

- Use it

Storage Object Cache Size (-omsize)

- Another no-brainer
- Eliminates object cache I/O and latch usage

Recommendations

- ***select count(*) from _storageobject*** and round up a bit

White Star Software
www.wss.com

AI File Management (-aiarcinterval –aiarcdir)

- You all have AI enabled right?
- Ancient versions of Progress required scripting
  - Copy AI file
  - Rfutil –C aimage new + rfutil –C aimage empty
  - AIMGT does all that automatically

Recommendations

- Enable: rfutil db –C aiarchiver enable
- -aiarcdir: local filesystem #1, local filesystem #2 (NO NFS mounts)
- -aiarcinterval: business dependant
  - 15 minutes is good
  - 2-4h if using OE Replication

The Fancy Pants

White Star Software
www.wss.com

Table and Index Statistics

- -basetable 0 –baseindex 0 –tablerangesize –indexrangesize
- Really these should be in the “must have” section …
  - I had to discuss it with my spiritual advisor
- By default Progress only captures statistics for the first 50 tables
- Why isn’t this automatic? Because the shared memory size calculation is done before the DB is brought up but you need the DB up to count the tables and indexes

Recommendations

- # of tables and indexes – round up

Windows Event Level (-evtlevel)

- Apply the Linux patch

Recommendations

- -evtlevel NONE
- ‘Nuf said

White Star Software
www.wss.com

Server –minport and -maxport

- -S parameter defines the broker listening port
- Broker spawns servers (_mprosrv –m1) that must also listen on ports
- If you don’t want to drive your firewall admin crazy, set minport and maxport

Recommendations

- Any range big enough to handle the –Mn servers
- Leave some room as other processes may consume ports in the range

Alternate Buffer Pool (-B2)

- Very cool tool to pin database objects in shared memory
- If you do it right, there is no LRU latch
- Put frequently accessed small tables here (and their indexes!)

Recommendations

- Big enough to fit ALL the record and index blocks of all the tables assigned to –B2

prefetchDelay/prefetchFactor/prefetchNumRecs/prefetchPriority

- New network connection parameters
- Have done some benchmarking – dramatic improvements in certain cases
- prefetchPriority: Prefer filling network messages with records over polling
- -prefetchDelay: if not set, first message contains one record
- -prefetchNumRecs/-prefetchFactor: How much to fill a network message

Recommendations

- -prefetchPriority 100 –prefetchDelay
- -prefetchNumRecs 128
- –prefetchFactor 100 didn’t seem to work as well as prefetchNumRecs on my limited testing

White Star Software
www.wss.com

Message Buffer Size (-Mm)

- Default is 1024
- Must be the same everywhere (until 11.5.1 supposedly)

Recommendations

- -Mm 8192 for MTU 1500
- Probably larger for Jumbo Frames

Message Buffer Size

Server Activity

Srv	Type	Port	Cnx	Max	LogRd	v	QryRcvd	RecRcvd	MsgRcvd	r/msg	RecSent	MsgSent	s/msg	MB Sent	MB Rcvd	RcvdSz	SendSz
>9999	Total	0	25	0	1032433		14575	0	14695	0	513708	14575	35	111.18	1.52	109	7999
71	Auto	20080	0	5	46205		549	0	555	0	22983	549	42	4.18	0.06	109	7989
51	Auto	20060	0	5	44690		552	0	560	0	22229	552	40	4.20	0.06	109	7979
11	Auto	20017	1	5	43050		470	0	477	0	21413	470	46	3.58	0.05	109	7992
8	Auto	20013	1	5	42705		568	0	573	0	21247	568	37	4.34	0.06	109	8004
20	Auto	20026	1	5	42159		569	0	574	0	20977	569	37	4.34	0.06	109	7996
30	Auto	20037	0	5	42057		554	0	559	0	20927	554	38	4.23	0.06	109	7998
42	Auto	20051	0	5	41575		571	0	575	0	20688	571	36	4.37	0.06	109	8012
21	Auto	20027	1	5	41408		547	0	552	0	20598	547	38	4.17	0.06	109	7998
45	Auto	20054	0	5	40761		588	0	592	0	20285	588	35	4.48	0.06	109	8000
65	Auto	20074	0	5	39974		600	0	605	0	19890	600	33	4.58	0.06	109	7998
25	Auto	20031	1	5	39948		587	0	591	0	19879	587	34	4.48	0.06	109	8005
19	Auto	20025	1	5	39352		554	0	559	0	19579	554	35	4.22	0.06	109	7993
5	Auto	20008	1	5	38899		551	0	556	0	19358	551	35	4.21	0.06	109	7998
75	Auto	20084	0	5	38749		567	0	571	0	19281	567	34	4.33	0.06	109	8009
26	Auto	20032	1	5	38706		588	0	593	0	19260	588	33	4.49	0.06	109	8000
69	Auto	20078	0	5	38198		554	0	559	0	19007	554	34	4.23	0.06	109	7997
48	Auto	20057	0	5	37976		590	0	593	0	18901	590	32	4.50	0.06	109	8005
39	Auto	20047	0	5	37784		535	0	539	0	18801	535	35	4.08	0.06	109	8000
68	Auto	20077	0	5	36648		554	0	558	0	18237	554	33	4.23	0.06	109	7998
12	Auto	20018	1	5	36535		604	0	606	0	18184	604	30	4.61	0.06	108	8010

- ProTop says we're filling our –Mm 8192
- Aggregate MB sent matches NIC speed (1 Gbit)

The “Why Would I Use These?”

White Star Software
www.wss.com

-DbCheck and -MemCheck

- Consistency check on all DB blocks when written
- Consistency check on memory operations
- Sorry – not sure of effect on performance

Recommendations

- Not sure yet
- I want to say “yes turn them on” but then again I haven’t seen a lot of consistency issues

Excess Shared Memory (-Mxs)

- Broker calculates shared memory requirements based on start-up parameters
- Also adds on a little extra something-something…just in case
- Useful for _proutil db –C increaseto

Recommendations

- I don’t generally use this parameter

Semaphore Sets (-semsets)

- No one has ever given me a good explanation with respect to this parameter
- No one seems to know how to prove that you need more
- Documentation says:

> *When more than 1,000 users connect to a single database, there might be high contention for the semaphore set. If there is a lot of semaphore contention on a system, using multiple semaphore sets helps alleviate this contention and improve performance on high user counts.*

Recommendations

- There seems to be some rule-of-thumb: 1 per 100 concurrent users

White Star Software
www.wss.com

Shared Memory Segment Size (-shmsegsize)

- Specify max size of shared memory segment
- If you don’t specify it, Progress will ask the OpSys for the biggest shmseg possible
- Not sure why you would want to specify a smaller shmseg size

Recommendations

- Ignore

White Star Software
www.wss.com

Latch Spin Tuning

- Initial latch sleep time
- Maximum latch sleep time

Recommendations

- Leave them as is on modern versions of Progress

The Gang-O-Useless

White Star Software
www.wss.com

Direct IO (-directio)

- Even when it was useful (v9) it was maybe only relevant on AIX
- Officially it was supposed to instruct the O.S. not to buffer data in the FS cache
- The idea being no need to double buffer

- In reality data is buffered everywhere
  - SAN
  - FS
  - DB

- NEW: Dan Foreman mentioned that he found ONE WEIRD case where it helped

Delayed BI File Write (-Mf)

- Maximum age of BI notes in BI buffers
- Once upon a time there was a reason to increase this on very busy systems
- Not so much today
- Leave the default of 3 seconds

Cluster Age Time (-G)

- Now called “Before Image Truncate Interval”
- Number of seconds before DB reuses a BI cluster
- Used to be 60 seconds
- Just ignore it now – default is 0

DB Buffer Hash Table Entries (-hash)

- To find a block in –B you need to search in the hash table
- Waaaaaay back Progress didn’t calculate this correctly for large –B
- Recommendation was first prime number > -B X 0.25
- This has been fixed for a long time
- Ignore this parameter

No Crash Protection (-i)

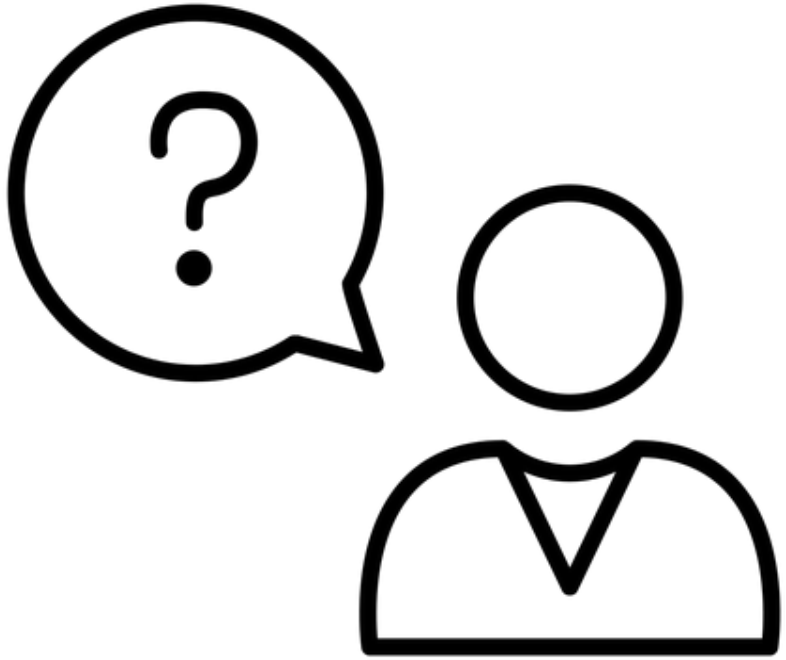
- DB writes fewer BI notes
- Great for loading data (proutil load –i)
- Not so great for a production DB
- Any problem pretty much means DB is garbage
  - I.e. restore from backup

- -r (Buffered BI I/O) is similar: don’t use it

LRU Skips – Alternate Buffer Pool (-lru2skips)

- The main use of –B2 is data that should stay in memory
- If size of B2 > size of data, LRU never used
- I.e. no need to use lru2skips!

Q&A



pk@wss.com

White Star Software
www.wss.com