



Using onstats for Performance Tuning and Troubleshooting

Suma Vinod
IBM

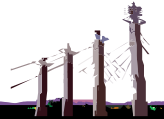
Session D04
Wednesday, April 28, 2008 • 14:10 a.m. – 15:10 a.m.

2008 IIUG Inform*i*x Conference



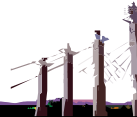
Agenda

- Performance Problems
- Engine Hangs
- Long Checkpoint duration
- Other Useful onstat commands
- What is new in IDS 11 and IDS 11.50?



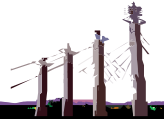
Performance Issues

- Improve my system
- Why did my system suddenly slow down?
- A particular query performance not acceptable



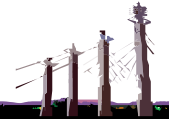
Understand the system

- How long did SQL statements take?
- How much resources of each category did a statement take?
 - Disk IO
 - CPU
 - Memory
- How long and how many times did we wait on each resource?
 - Locks
 - Disk IO
- Index Scan or Sequential scan
If Index scan, is it using the right index



Information

- onstat outputs
 - onstat -p
 - onstat -g ppf
 - onstat -D
 - onstat -g iof
 - onstat -F
- Sysmaster database
- Sqexplain output



onstat -p

IBM Informix Dynamic Server Version 10.00.FC6WM1 -- On-Line -- Up 05:25:50 -- 637504 Kbytes

Profile							
dskreads	pagreads	bufreads	%cached	dskwrits	pagwrits	bufwrits	%cached
315375	1567961	51147331	99.39	60026	107548	2074544	97.15

isamtot	open	start	read	write	rewrite	delete	commit	rollbk
48082667	1667795	3443060	30155453	652615	274949	95613	21782	1029

Percentage of reads from shared memory relative to disk reads. A low value indicates that too few BUFFERS are configured. Should be > 95% for OLTP systems

ixda-RA	idx-RA	da-RA	RA-pgsused
79790	895	58226	128752

Percentage of writes from shared memory relative to disk writes. LRU and CKPTINVL affect this



onstat -p

IBM Informix Dynamic Server Version 10.00.FC6WM1 -- On-Line -- Up 05:25:50 -- 637504 Kbytes

Profile

dskreads pagreads
315375 1567961

isamtot
48082667

If ratio between seqscans and isamtot > 1/%,
check index usage

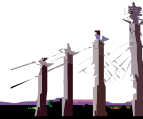
gp_read gp_write gp_rewrt gp_del gp_alloc gp_free gp_curs
2 0 0 0 0 0

ovlock ovuserthread ovbuff usercpu syscpu numckpts flushes
0 0 0 732.82 209.55 6 12

bufwaits lokwaits lockreqs deadlks dltouts ckpwaits compress
75324 0 35652918 0 0 5 36241

seqscans
74792

ixda-RA idx-RA da-RA RA-pgsused lchwaits
79790 895 58226 128752 3532



onstat -p

IBM Info Number of times we had to wait on a buffer in the buffer pool. This can indicate that a single page is being altered too much. Also this can occur if the page is being flushed to disk too often (LRU_MIN set to 0).

Profile
dskread 315375
isamtot 48082667 7667795 3443060 30155453 652615 274949 95613 21782 1029

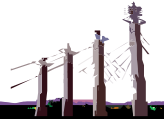
gp_read gp_write gp_rewrt gp_del
2 0 0 0 0

ovlock ovuserthread ovbuff userrep
0 0 0 732.82 209.55

bufwaits lokwaits lockreqs deadlks
75324 0 35652918 0

ixda-RA idx-RA da-RA RA-pgsused lchwaits
79790 895 58226 128752 3532

Number of times lock request issued on a locked table/page/row. If ratio between lokwaits/lockreqs is too high, then applications may be single-threading.



onstat -p

IBM Informix Dynamic Server Version 10.00.FC6WM1 -- On-Line -- Up 05:25:50 -- 637504 Kbytes

Profile

dskreads	pagreads	bufreads	%cached	dskwrits	pagwrits	bufwrits	%cached
315375	1567961	51147331	99.39	60026	107548	2074544	97.15

isamtot	open	start	read	write	rewrite	delete	commit	rollbk
48082667	1667795	3443060	30155453	652615	274949	95613	21782	1029

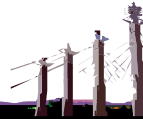
gp_read	gp_write	gp_rewrt	gp_del	gp_alloc	gp_free	gp_curs
2	0	0	0	0	0	0

ovlock	ovuserthread
0	0

bufwaits	lokwaits	lock
75324	0	3564

ixda-RA	idx-RA	da-RA	RA-pgsused	lchwaits
79790	895	58226	128752	3532

If ratio between rollbk and commit > 1%,
need to examine the application to see why
there are a lot of rollbacks



onstat -g ppf

Partition profiles

partnum	lkrqs	lkwts	dlks	touts	isrd	iswrt	isrwt	isdel	bfrd	bfwrt	seqsc	rhitratio
0x100001	0	0	0	0	0	0	0	0	0	0	0	0
0x100002	18	0	0	0	16	0	0	0	33	0	2	88
0x100003	6	0	0	0	2	0	0	0	8	0	0	63
0x100071	107	0	0	0	127	0	0	0	293	12	1	98
0x100072	22	0	0	0	21	0	0	0	138	45	0	97
0x100073	3	0	0	0	3	0	0	0	11	0	0	73
0x100074	2	0	0	0	2	0	0	0	20	8	0	80
0x100075	1	0	0	0	1	0	0	0	7	0	0	58
0x100076	0	0	0	0	0	0	0	0	4	0	0	75

Partnum of the
table/fragment

Activity on the
table/fragment

Ratio of buffer
reads that did not
require a disk
read



www.iiug.org



onstat -g ppf

Partition profiles

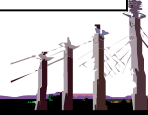
partnum	lkrqs	lkwts	dlks	touts	isrd	iswrt	isrwt	isdel	bfrd	bfwrt	seqsc	
	rhitratio											
0x100001	0	0	0	0	0	0	0	0	0	0	0	0
0x100002	18	0	0	0	16	0	0	0	33	0	2	88
0x100003	6	0	0	0	2	0	0	0	8	0	0	63
0x100071	107	0	0	0	127	0	0	0	293	12	1	98
0x100072	22	0	0	0	21	0	0	0	138	45	0	97
0x100073	3	0	0	0	3	0	0	0	11	0	0	73
0x100074	2	0	0	0	2	0	0	0	20	8	0	80
0x100075	1	0	0	0	1	0	0	0	7	0	0	58
0x100076	0	0	0	0	0	0	0	0	4	0	0	75

High numbers here
indicate that the
application needs
some work

High numbers indicate
that adding an index
might help



www.iiug.org



onstat -D

Dbspaces
address number

address	number	pages	chunks	status	name
a13e150	1	1	2	N	informix dbspace2
a13ea88	2	1	4	N	informix dbspace4
a13ec08	4	1	5	N	
a13ecc8	5	1	6	N	
a13ed88	6	1	7	N	
a13ee48	7	1	8	N	
a13ef08	8	1	1	N	

8 active, 2047 maximum

If too high, split data on those chunks

Avoid having the two
chunks on the same
device

Chunks

address chk/dbs offset

address	chk/dbs	offset	page Rd	page Wr	pathname
a13e210	1	32	21176	11196	/work5/PROD/chunks/rootdbs
a13e388	2	8	13	22	/work5/PROD/chunks/dbspace1
a13e488	3	3	2	0	/work5/PROD/chunks/dbspace3
a13e588	4	4	13	23	/work5/PROD/chunks/dbspace4
a13e688	5	5	14	22	/work5/PROD/chunks/dbspace5
a13e788	6	6	13	21	/work5/PROD/chunks/dbspace6
a13e888	7	7	11	21	/work5/PROD/chunks/dbspace7
a13e988	8	8	14	22	/work5/PROD/chunks/dbspace1

8 active, 2047 maximum

page Rd page Wr

pathname

/work5/PROD/chunks/rootdbs
/work5/PROD/chunks/dbspace1
/work5/PROD/chunks/dbspace3
/work5/PROD/chunks/dbspace4
/work5/PROD/chunks/dbspace5
/work5/PROD/chunks/dbspace6
/work5/PROD/chunks/dbspace7
/work5/PROD/chunks/dbspace1

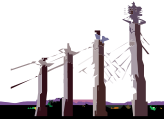


onstat -g iof

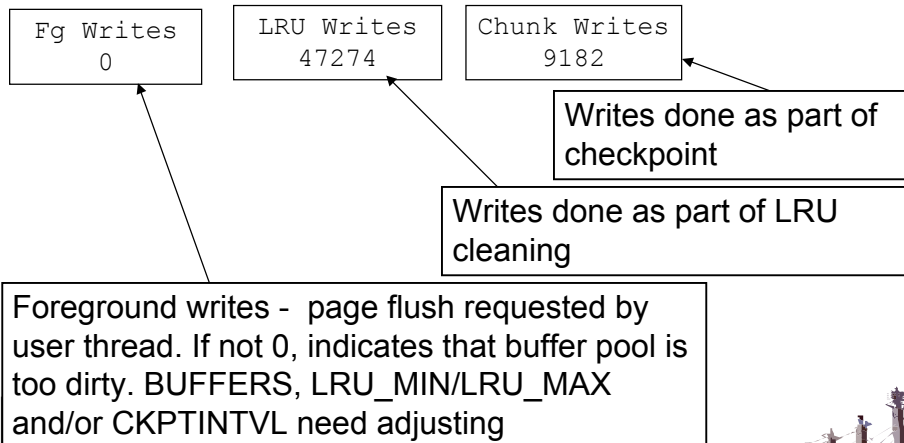
AIO global files:

gfd	pathname	totalops	dskread	dskwrite	io/s
3	RootChunk	11771	8723	3049	3.27
4	Chunk1	25341	25341	4611	8.32
5	Chunk2	12420	5436	6984	3.45
6	SBChunk	260	1120	140	0.35
7	Chunk3	108	107	1	0.03

Need to consider moving data from this chunk as there is a lot of I/O. If it is a single table, consider fragmentation.

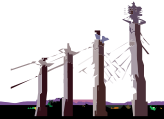


onstat -F



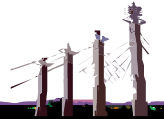
Types of Hangs

- Network/Connectivity related hangs
- Resource Hangs
 - A single thread waiting on a resource
 - Multiple threads waiting on a common resource
- Server Blocked - Deadlock situation



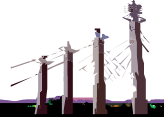
Network/Connectivity Hangs

- Consider
 - the current protocol (TCP, shared memory, streams ???)
 - the current INFORMIXSERVER setting
 - DBSERVERALIASES
- Is the problem specific to just one protocol? Are shared memory connections fine, but TCP connections hung?
- Do other server names of the same protocol experience the same problem?



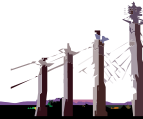
Listen Threads and Poll threads

- The listen thread
 - Responsible for accepting new connections and starting a corresponding session for this new connection
- The poll thread
 - monitors for incoming events
 - tells the listen thread when new connection requests have arrived
 - notifies existing connections when new messages arrive



Network Thread Names

Protocol	Poll Thread	Listen Thread
TLI	tlitcpoll	tlitcplst
Sockets	soctcpoll	soctcplst
Streams	ipcstrpoll	ipcstrlst
Shared Memory	sm_poll	sm_listen

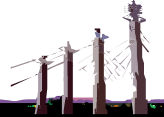


Typical Network Thread States

- Consider this onstat -g ath output:

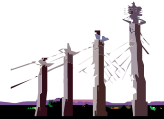
```
Threads:
tid      tcb      rstcb    prty status          vp-class    name
9        b338d70  0        2    running          9tli       tlitcpoll
12       b37a698  0        3    sleeping forever    1cpu       tlitcplst
13       b308300  0        3    sleeping forever    3cpu       tlitcplst
```

- This output represents typical states for these threads
- Poll threads are generally running
- Listen threads are commonly found sleeping forever
- Poll threads are constantly checking for incoming events. It will wake the listen thread if a new connection indication has arrived
- If a hang has been reported and one or all of these threads are consistently in some other state, this may be an outward indication of the problem



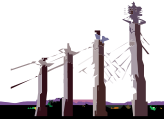
Information Needed

- Information about threads – onstat –g ath
- Network onconfig parameters
- Sqlhosts
- DBSERVERNAME, DBSERVERALIASES
- Network onstats
 - onstat –g ntu
 - onstat –g ntm
 - onstat –g ntt



Case study - New connections sometimes hang

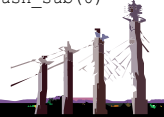
- Network onconfig parameter
- Sqlhosts
- onstat -g ath
- Network onstats



onstat -g ath

Threads:

tid	tcb	rstcb	prty	status	vp-class	name
7	11ac0a438	11a1de028	4	sleeping secs: 1	1cpu	main_loop()
8	11abbcb8e	0	2	running	1cpu	sm_poll
9	11ab416e0	0	2	running	14tli	tlitcpgpoll
10	11ac56da8	0	2	running	15tli	tlitcpgpoll
11	11ac72da8	0	2	running	16tli	tlitcpgpoll
12	11ac8fda8	0	2	running	17tli	tlitcpgpoll
13	11acabda8	0	2	running	18tli	tlitcpgpoll
14	11acc8da8	0	2	running	19tli	tlitcpgpoll
15	11ace4da8	0	2	running	20tli	tlitcpgpoll
16	11ad01da8	0	2	running	21tli	tlitcpgpoll
17	11ad1dda8	0	3	sleeping forever	3cpu	sm_listen
18	11ab4bb00	0	2	sleeping secs: 1	3cpu	sm_discon
19	11ab50b00	0	3	cond wait closing	1cpu	tlitcplst
20	11ad7f028	0	3	sleeping forever	3cpu	tlitcplst
21	11ad9f2d0	0	3	sleeping forever	4cpu	tlitcplst
22	11adbf4f0	0	3	cond wait closing	5cpu	tlitcplst
23	11ade3248	0	3	sleeping forever	6cpu	tlitcplst
24	11ade3980	11a1de840	2	sleeping forever	3cpu	flush_sub(0)



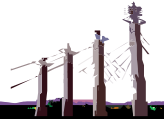
SQLHOSTS and SERVICES

- SQLHOSTS

```
kj10_online    ontlitcp 9.11.1.111 kj10_net  
kj10_iponline  ontlitcp netis01    kj10_ipnet  
kj10_wonline   ontlitcp netis01_w   kj10_web  
kj10_remotel  ontlitcp 9.11.1.111 kj10_remotel
```

- Services

```
kj10_net 3010/tcp # netis01  
kj10_web 3015/tcp # netis01 web Port  
kj10_ipnet 3016/tcp # netis01 ipload Port
```



onstat -g ntm

Network mailbox information:

box	netscb	thread name	max	received	in box
max in box	full	signal			
5	11ac3daf0	tlitcpoll	10	991565	0
3	0	yes			
6	11ab56d18	tlitcpoll	10	212234	0
3	0	yes			
7	11ac76550	tlitcpoll	10	212236	0
3	0	yes			
8	11ac58cf0	tlitcpoll	10	212279	0
3	0	yes			
9	11acaf550	tlitcpoll	10	213133	0
5	0	yes			
10	11ac91cf0	tlitcpoll	110	212145	73
73	2	yes			



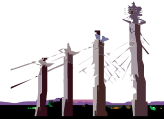
-g ntu and -g ntt

- Individual thread network information (basic):

netscb	type	thread	sid	fd	poll	reads	writes	q-nrm	q-pvt	q-exp
11baf01b8	tlitcp	unknown	848350	280	10	0	0	0/ 0	0/ 0	0/ 0

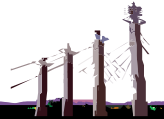
- Individual thread network information (times):

netscb	thread	name	sid	open	read	write	address
11ccfc9a8	sqlexec		848452	00:25:01			00:25:01
11baf01b8	tlitcp		848350	00:17:20			



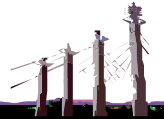
Listener Thread Stack– onstat –g stk

```
0x1007fd42c oninit :: yield_processor_mvp
0x1007f4a98 oninit :: mt_wait
0x100894b58 oninit :: disctli
0x10087fb04 oninit :: tlDiscon
0x10087c82c oninit :: slSQIdiscon
0x100876bc8 oninit :: pfDiscon
0x10086d844 oninit :: cmDiscon
0x10086a000 oninit :: ascAbort
0x10086a354 oninit :: ascInd
0x100841cdc oninit :: ASF_Call
0x10084996c oninit :: sql_listener
0x1007fdbb0 oninit :: startup
```



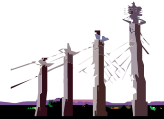
Symptoms

- The symptom of this case is that no new connections are able to be made on a particular listen thread.
- The listen thread has sent mail to the poll thread and it is waiting in cond wait closing state.
- The poll thread should pick up the mail message and signal the listen thread that it can continue – **this is not happening.**



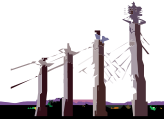
Symptoms

- The hang up seems to be related to the poll thread. Since the poll thread is running on a NET vp, we don't have to necessarily worry about threads running on CPU VPs.
- One of the poll threads would seem to be the crux of the problem, so get its stack.
- The onstat -g stk does not work on running threads and so use pstack.



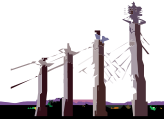
pstack for the poll thread

```
/usr/proc/bin/pstack 1939
1939:  oninit -v
      ffffffff7e7a2ff8 getmsg  (19, 141c6d9e8, 141c6d9d0, 141c6d9fc)
      ffffffff7e7a20e4 getmsg (19, 141c6d9e8, 141c6d9d0, 141c6d9fc, 4, 0) +
      28
      ffffffff7f32dc28 _t_is_ok (100dda080, ffffffff7f49c388, 2, 0,
      100dda080, 19) +
124
      ffffffff7f34cba0 _tx_snddis (2, 100dda118, 1, 8, 5000, 141c6dd70) + 448
      000000010083e00c tli_mail_close (100db65a8, 141d45960, 100da0a18,
      100d0b270,
      66000) + 240
      000000010083d6d4 tli_mail_process (10083d474, 141d45960, 1, 100d0d620,
      100d0da98) + 224
      000000010083a8c0 tli_poll_thread (1, 100dbdd88, 100c0d248, 140f14080,
      3,
      7d9ec8) + 268
      00000001007d42b0 startup (100db67b8, 7, 100dbdde8, 140f12b48,
      100bbb5a0,
      140f12ad0) + fc
      000000010082471c net_startup_step (0, 0, 0, 0, 0, 0) + 310
```



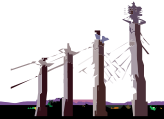
Solution

- The pstack info is repetitive for the poll thread and indicates that it may be hung within the TLI library call, `t_snddis`. This is a call that is made to abort the request associated with the new connection that the listener thread is attempting to reject.
- SUN Defect



Resource Hangs

- Isolate the source of the hang
- Start with onstat -g ath and onstat -u outputs

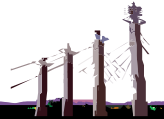


Listing User threads – onstat -u

Userthreads

address	flags	sessid	user	tty	wait	tout	locks	nreads	nwrites
ca0ea1c	Y--P--D	13	informix	-	a07d804	0	0	0	0
ca0e4e8	Y--P---	16	informix	7	d5cacc8	0	1	0	0

- Examine the *wait* column which identifies the address of a resource that this thread is currently waiting on.
- The *wait* column address may be associated with items like locks, conditions, buffers or perhaps mutexes.
- Examine onstat -k, onstat -g con, onstat -b and onstat -g lmx respectively to check for these items.



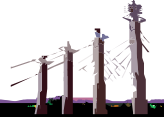
Finding the wait resource

- The `onstat -g con` output that corresponds to prior `onstat -u`

Conditions with waiters:

cid	addr	name	waiter	waittime
554	a07d804	bp_cond	36	63
533	d5cacc8	sm_read	40	247

- The thread associated with session 13 is waiting on the *bp_cond* condition and the thread associated with session 16 is waiting on *sm_read* condition.
- You may need to look for the address of the *wait* resource within the output of `onstat -a` or `onstat -g all`.



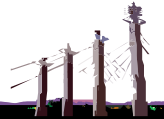
Listing all threads – onstat –g ath

- The onstat -g ath can also be helpful in identifying what type of resource a thread is waiting for

Threads:

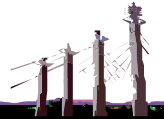
tid	tcb	rstcb	prty	status	vp-class	name
36	d6064e8	ca0ea1c	1	cond wait bp_cond	1cpu	bf_priosweep()
40	d5cad20	ca0e4e8	1	cond wait sm_read	1cpu	sqlxec

- The thread id associated with session 13 of the previous onstat output is 36. The thread id associated with session 16 of the previous onstat output is 40.
- onstat -g ath also shows that thread 36 is waiting on bp_cond and thread 40 is waiting on sm_read condition.



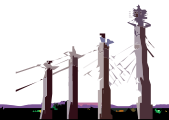
Resource Hangs – More to consider

- Are threads waiting on a resource held by another user thread?
- What is the thread doing that holds the resource...check onstat -g ses output for the corresponding session?
- Is the resource consistently held by the same thread or is it just a 'hot' resource desired and held by many different threads?



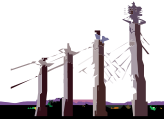
Server Blocked

- Deadlock
 - A single thread in critical section waiting for a resource
 - Incorrect alarmprogram errors
 - A thread waiting for a resource another thread is holding and the second thread in turn is waiting for a resource the first thread owns



Case Study

- Server blocked at checkpoint
- No threads are running



Case Study – Server Blocked at Checkpoint

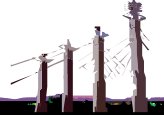
- IBM Informix Dynamic Server Version 9.40.FC4X4 -- On-Line (CKPT REQ) -- Up 4 d ays 11:27:14 -- 4316160 Kbytes Blocked:CKPT

- `onstat -g ath`

```
6                               3 sleeping 5cpu main_loop()
532883 1b3f02428 1a532e618 2 running 10cpu sqlexec
```

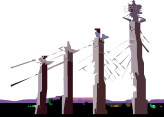
- `onstat -g ses`

tid	name	rstcb	flags	curstk	status
32883	sqlexec	1a532e618	---PX--	11151	running



Stack for main_loop()

```
0x100851d44 oninit :: yield_processor_mvp + 0x678
    sp=0x1a42c2170(0xf0a2d8,
0xefdce8, 0xedf9d8, 0xa3022b50, 0xa47d8028, 0xa3022ad8)
0x1005a9640 oninit :: wait4critex + 0x158 sp=0x1a42c22a0
    delta_sp=304(0xf01fd8,
0xa100180, 0x1, 0x41400, 0xa535c428, 0xa532e618)
0x10054c444 oninit :: checkpoint + 0x184 sp=0x1a42c2350
    delta_sp=176(0xed9950,
0x2000, 0x8000, 0xa07b800, 0x41912e35, 0x86f18)
0x1000cb064 oninit :: main_loop + 0x308c sp=0x1a42c2420
    delta_sp=208(0xee7d78,
0x71, 0x59ed5, 0xa13b420, 0x86200, 0xa101a00)
0x1008524d0 oninit :: startup + 0xfc sp=0x1a42c2e50
    delta_sp=2608(0xf0a2d8, 0x0,
0xedf9d8, 0xa3022b50, 0xa41b4028, 0xa3022ad8)
```

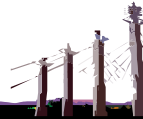


Stack for sqlexec

```
IBM Informix Dynamic Server Version 9.40.FC4X4  -- On-Line (CKPT  
REQ) -- Up 4 days 11:24:06 - 4316160 Kbytes Blocked:CKPT
```

```
Stack for thread: 532883 sqlexec  
base: 0x00000001aec30000  
len:   69632  
pc: 0x0000000100851d44  
tos: 0x00000001aec3e471  
state: running  
vp: 10
```

```
0x100851d44 oninit :: yield_processor_mvp + 0x678  
sp=0x1aec3ec70(0x0, 0x0, 0x0,  
0x0, 0x0, 0x0)
```



What is the sqlexec thread doing?

onstat -g ses:

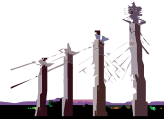
Sess Id	SQL Stmt type Explain	Current Database	Iso Lvl	Lock Mode	SQL ERR	ISAM ERR	F.E. Vers
57096	DELETE Off	dataware	DR	Wait 300	0	0	9.03

Current SQL statement :
delete from tmpmain where vdralloc is null or vdralloc = 0

Last parsed SQL statement :
delete from tmpmain where vdralloc is null or vdralloc = 0

User-created Temp tables :

partnum	tabname	rowsize
f000f8	tmpmain	267
f000f8	tmpmain	267



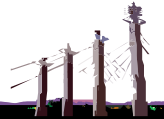
onmode -X s

- onmode -X s <number of cpu vp on which this thread is running>
 - onmode -X s 10

```
btusearch()  
btdelitem()  
delrecord()
```

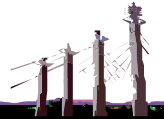
```
... *
```

```
... *
```



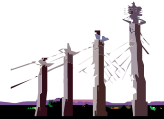
Symptoms

- The sqlexec thread which is trying to do a delete is in critical section and is looping in btusearch() – reading index keys
- The main_loop() thread is about to do a checkpoint and is waiting for the sqlexec thread to come out of critical section
- Other threads are waiting for the condition checkpoint to occur



Conclusion

- Bad Page – Index Corruption
- Index related ONCONFIG
 - FILLFACTOR 100
- Using FILLFACTOR 100 was able to reproduce the problem in-house



Long Checkpoint Duration

onstat -F

```
Informix Dynamic Server Version 7.31.UD6W3  --
On-Line (CKPT REQ) -- Up 19:45:31 --
2914496 Kbytes Blocked:CKPT
```

Fg Writes	LRU Writes	Chunk Writes
0	617	3795

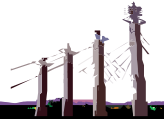
address	flusher	state	data
5e05c514	0	I	0 = 0X0
5e05ca10	1	I	0 = 0X0
5e05cf0c	2	I	0 = 0X0
5e05d408	3	I	0 = 0X0
...			
5e0623c8	19	I	0 = 0X0
5e0632bc	22	C	27 = 0X1b
5e0637b8	23	I	0 = 0X0
...			
5e06ff18	63	I	0 = 0X0

states: Exit Idle Chunk Lru

**What is this onstat -F
output telling us?**

**State of "C" indicates a
chunk write, and the
0x1b (27) is the chunk
number.**

**What onstat should we
run next?**



Long Checkpoint

onstat -D

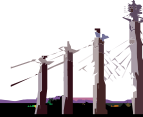
Informix Dynamic Server Version 7.31.UD6W3 -- On-Line -- Up 19:22:43 --
2914496 Kbytes

Dbspaces

address	number	flags	fchunk	nchunks	flags	owner	name
5eb54dd8	4	1	7	2	N	informix	cmis_dbs2
...							
5eb55b58	23	2001	27	1	N T	informix	tempdbs1
22 active, 2047 maximum							

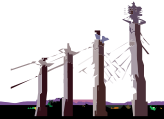
Chunks

address	chk/dbs	offset	page Rd	page Wr	pathname
...					
5eb53018	27 23	0	27885	36019	/opt/informix/dblinks/a3



Other Useful onstats

- onstat -g stk
 - Stack info
- onstat -g sts
 - Maximum and current stack sizes for each thread
- onstat -g stm [sesid]
 - Prepared statements for the session(s) and memory used
- onstat -g env [sesid]
 - Environment variables for the server or the session
- onstat -g opn
 - List of open partitions
- onstat -g tpf
 - Thread Profiles

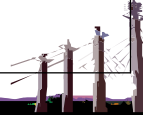


IDS 11 – Query Drill Down Feature

- Allow tracing of SQL statements
- Provide detailed information about SQL statements
- Information available through onstat or sysmaster database
- Dynamically configurable
- By default disabled
- Global and User Tracing Modes

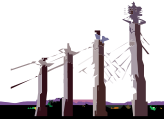


Name	Description
User ID	User id
Session ID	Database Session ID
Database Name	Current Database Name
Statement Type	Type of SQL statement being executed
Statement Execution time	Duration of the SQL statement
Time of Execution	Time this statement completed
Statement Text	SQL statement text or the procedure stack trace with statement type
RSAM statistics	• Buffer Reads & Writes • Page reads & Writes • Memory Sorts, disk Sorts, • Lock requests, waits • Logical Log Records
SQL statistics	• Estimated # of rows • Estimated Cost • # of rows returned • Statement Type • Database Isolation Level



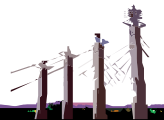
Modes of Tracing

- Global Mode
 - Enable tracing for all sessions on the system
 - To do a quick comparative analysis on resource usage of all the sessions that are running
 - To identify the sessions to trace, when you are unsure of which specific user or users to trace
- User Mode
 - Enable tracing for all users
 - Actual tracing doesn't start unless a SQL API command is executed
 - To narrow down tracing to specific set of users or just one user
 - When you know what user or user sessions to trace



SQLTRACE parameter

- onconfig parameter
- To enable default tracing at system startup
- Syntax:
 - SQLTRACE level=[Low|Medium|High|Off],
ntraces=[number of traces],
size=[size of each trace buffer],
mode=[global|user]



Setting SQLTRACE for global tracing

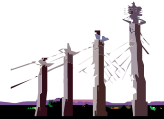
```
SQLTRACE level=LOW,ntraces=2000,size=2,mode=global
```

- The database server will gather low level information for SQL statements
- Tracks information on the most recent 2000 SQL statements
- Allocates approximately 4 MB (2000 traces * 2 KB)



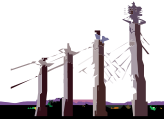
Admin API command

- Use function “*task*” or “*admin*” of sysadmin database
- Order of arguments important
- Syntax:
 - execute function task(“set sql tracing on”,
2000, 2, “low”, “global”);



Admin API command

- Use function “*task*” or “*admin*” of sysadmin database
- Order of arguments important
- Syntax:
 - execute function task(“set sql tracing on”,
2000, “4k”, “med”, “user”);



onstat -g his output (truncated)

Iterator/Explain

=====

ID	Left	Right	Est Cost	Est Rows	Num Rows	Type
2	0	0	5	1	10	Seq Scan
1	2	0	1	1	1	Sort

Statement information:

Sess_id	User_id	Stmt Type	Finish Time	Run Time
16	6	SELECT	15:22:00	0.0004

Statement Statistics:

Page	Buffer	Read	Buffer	Page	Buffer	Write
Read	Read	% Cache	IDX Read	Write	Write	% Cache
0	22	100.00	0	0	0	0.00
Lock	Lock	LK Wait	Log	Num	Disk	Memory
Requests	Waits	Time (S)	Space	Sorts	Sorts	Sorts
0	0	0.0000	0.000 B	0	0	0
Total	Total	Avg	Max	Avg	I/O Wait	Avg Rows
Executions	Time (S)	Time (S)	Time (S)	IO Wait	Time (S)	Per Sec
11	0.0109	0.0010	0.0004	0.000000	0.000000	2830.6297
Estimated	Estimated	Actual	SQL	ISAM	Isolation	SQL
Cost	Rows	Rows	Error	Error	Level	Memory
1	1	1	0	0	DR	17192



www.iiug.org

Checkpoint Statistics - onstat -g ckp

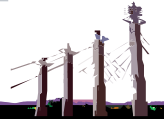
Auto Checkpoints=On RTO_SERVER_RESTART=60 seconds Estimated recovery time 7 seconds

Critical Sections											Physical Log		Logical Log				
Interval	Clock Time	Trigger	LSN	Total	Flush	Block	#	Ckpt	Wait	Long	# Dirty	Dskfl	Total	Avg	Total	Avg	
				Time	Time	Time	Waits	Time	Time	Buffers	/Sec	pages	/Sec	Pages	/Sec		
1	18:41:36	Startup	1:18	0.0	0.0	0.0	0	0.0	0.0	0.0	4	4	3	0	1	0	
2	18:41:49	Admin	1:11c12cc	0.3	0.2	0.0	1	0.0	0.0	0.0	2884	2884	1966	162	4549	379	
3	18:42:21	Log	8:188	2.3	2.0	2.0	1	0.0	2.0	2.0	14438	7388	318	10	65442	2181	
4	18:42:44	User	10:19c018	0.0	0.0	0.0	1	0.0	0.0	0.0	39	39	536	21	20412	816	
5	18:46:21	RTO	12:188	54.8	54.2	0.0	30	0.6	0.4	0.6	68232	1259	210757	1033	150118	735	
Max Plog pages/sec				Max Llog pages/sec		Max Dskflush Time		Avg Dskflush pages/sec		Avg Dirty pages/sec		Blocked Time					
8796				6581		54		43975		2314		0					

Time between
checkpoints

What caused
the
checkpoint?

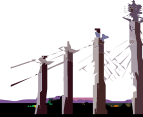
Was the
checkpoint
blocking?



IDS 11 – Improved onstat –g ath

Threads

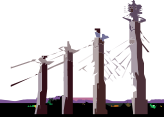
tid	tcb	rstcb	prty	status	vp-class	name
2	10bbf36a8	0	2	IO Idle	3lio	lio vp 0
3	10bc12218	0	2	IO Idle	4pio	pio vp 0
4	10bc31218	0	2	running	5aio	aio vp 0
5	10bc50218	0	2	IO Idle	6msc	msc vp 0
6	10bc7f218	0	2	running	7aio	aio vp 1
7	10bc9e540	10b231028	4	sleeping secs: 1	1cpu	main_loop()
8	10bc12548	0	2	running	1cpu	tlitcpoll
9	10bc317f0	0	3	sleeping forever	1cpu	tlitcplst
10	10bc50438	10b231780	2	IO Wait	1cpu	flush_sub(0)
11	10bc7f740	0	2	IO Idle	8aio	aio vp 2
12	10bc7fa00	0	2	IO Idle	9aio	aio vp 3
13	10bd56218	0	2	IO Idle	10aio	aio vp 4
14	10bd75218	0	2	IO Idle	11aio	aio vp 5
15	10bd94548	10b231ed8	3	sleeping forever	1cpu	aslogflush
16	10bc7fd00	10b232630	1	sleeping secs: 34	1cpu	btscanner 0
32	10c738ad8	10b233c38	4	sleeping secs: 1	1cpu	onmode_mon
50	10c0db710	10b232d88	2	IO Wait	1cpu	sqlxec



onstat -g cpu

Thread CPU Info:

tid	name	vp	Last Run	CPU Time	#scheds	status
*2	lio vp 0	3lio	06/27 13:26:39	28.6397	3749	IO Idle
*3	pio vp 0	4pio	06/27 13:25:09	5.0609	517	IO Idle
*4	aio vp 0	5aio	06/27 13:29:23	31.1610	112645	IO Idle
*5	msc vp 0	6msc	06/27 13:27:57	0.1137	50	IO Idle
*6	aio vp 1	7aio	06/27 13:29:23	19.1152	5524	IO Idle
7	main_loop()	1cpu	06/27 13:31:55	7.1407	678090	sleeping secs: 1
*8	sm_poll	1cpu	06/27 13:31:55	677245.0333	940398	running
9	sm_listen	1cpu	06/27 13:27:57	0.0057	32	sleeping forever
10	sm_discon	1cpu	06/27 13:31:55	2.5516	676641	sleeping secs: 1
11	flush_sub(0)	1cpu	06/27 13:31:55	1.7716	677707	sleeping secs: 1
*12	aio vp 2	8aio	06/27 13:29:23	21.7697	727	IO Idle
*13	aio vp 3	9aio	06/27 13:25:09	23.7650	677	IO Idle
*14	aio vp 4	10aio	06/27 13:25:09	18.0777	1118	IO Idle
*15	aio vp 5	11aio	06/27 13:25:09	17.0063	350	IO Idle
16	aslogflush	1cpu	06/27 13:31:55	2.0833	676638	sleeping secs: 1
17	btscanner_0	1cpu	06/27 13:31:35	1.7299	22352	sleeping secs: 31
*18	onmode_mon	1cpu	06/27 13:31:55	2.9390	676641	sleeping secs: 1
*40	dbScheduler	1cpu	06/27 13:29:23	1.5202	3444	sleeping secs: 148
*41	dbWorker1	1cpu	06/27 13:24:22	0.9907	2655	sleeping forever



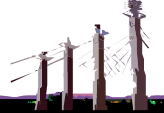
IDS 11.50 - Improved onstat -g glo

Individual virtual processors:

vp	pid	class	usercpu	syscpu	total	Thread	Eff
1	1245430	cpu	23.59	7.69	31.28	31.28	100%
2	1347720	adm	3.10	3.94	7.04	0.00	0%
3	888872	cpu	14.68	2.66	17.34	17.34	100%
4	782406	lio	0.60	1.62	2.22	43.03	5%
5	364756	pio	0.39	1.05	1.44	1.44	100%
6	819436	aio	0.72	3.71	4.43	30.26	14%
7	1093828	msc	0.01	0.00	0.01	1.13	0%
8	442598	aio	0.43	1.06	1.49	5.00	29%
9	1024150	aio	0.39	1.02	1.41	4.81	29%
10	938210	aio	0.36	0.98	1.34	5.12	26%
12	1511600	aio	0.37	0.99	1.36	3.88	35%
11	594110	aio	0.35	0.94	1.29	5.08	25%
		tot	44.99	25.66	70.65		

Number of seconds the process ran on
the physical processorPercentage of effective processor
utilization

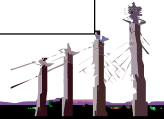
of seconds the thread ran on this VP



IDS 11.50 - onstat -g osi

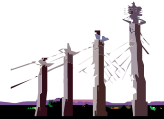
- Displays information about the machine
- Information available in sysmaster:sysmachineinfo

```
Machine Configuration....
OS Name                      AIX
OS Release                   3
OS Node Name                 rat-06
OS Version                   5
OS Machine                   000BC44D4C00
Number of processors         2
Number of online processors  2
System memory page size     4096 bytes
System memory                0 MB
System free memory          0 MB
Number of open files per process 65534
...
...
```



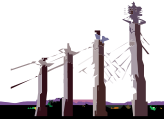
DBScheduler for Information Collection

- Enable/Disable tracing at predefined times
 - At system start up time
 - At specific times
- Generate reports based on tracing for later analysis
- Monitor session statements and take actions as needed
- Collect information at predefined times or at periodic intervals



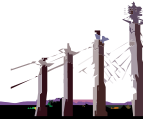
How to use DBScheduler?

- Connect as informix to sysadmin database
 - ph_alert, ph_run, ph_task, ph_group, ph_threshold
- A set of predefined tasks are available
- All the executed tasks are logged into ph_run table
- Create a task for your needs by inserting a row into ph_task table



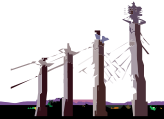
DBScheduler – Example

```
INSERT INTO ph_task (tk_name, tk_type,tk_group,tk_description,  
tk_result_table,tk_create,tk_execute,tk_start_time,tk_stop_time,  
tk_delete,tk_frequency,tk_next_execution)  
VALUES  
('Profile Information',  
'SENSOR',  
'SERVER',  
'Profile Information',  
'prof_info',  
"create table prof_info (dskreads integer, bufreads int, dskwrites  
int, bufwrites ); ",  
"insert into prof_tab select dskreads, bufreads, dskwrites, bufwrites  
from sysmaster:sysprofile",  
"04:00:00", "04:05:00",NULL,  
INTERVAL ( 1 ) minute TO minute, DATETIME(04:00:00) HOUR TO SECOND);
```



DBScheduler References

- **How To Terminate Idle Users with the Database Admin System**
 - http://www-128.ibm.com/developerworks/blogs/page/idsteam?entry=terminate_idle_users_with_the
- More about DBScheduler
 - http://www.ibm.com/developerworks/blogs/page/idsteam?entry=the_db_scheduler



Session D04

Using onstats for Performance Tuning and Troubleshooting

Suma Vinod

IBM

sumam@us.ibm.com

