Garbage collection mechanism and sweep in details

Firebird Conference 2019

Berlin, 17-19 October

IBSurgeon

Gstat -h

- Database header page information:
- Flags 0
- Checksum 12345
- Generation 112431494
- Page size 8192
- ODS version 11.1
- **Oldest transaction 100 x-1**
- **Oldest active 101 x**
- **Oldest snapshot 101 x**
- **Next transaction 102 x+1**
- Bumped transaction 1
- Sequence number 0
- Next attachment ID 0
- Implementation ID 16
- Shadow count 0
- Page buffers 256
- Next header page 0
- Database dialect 1
- Creation date Jun 5, 2011 10:02:19
- Attributes force write

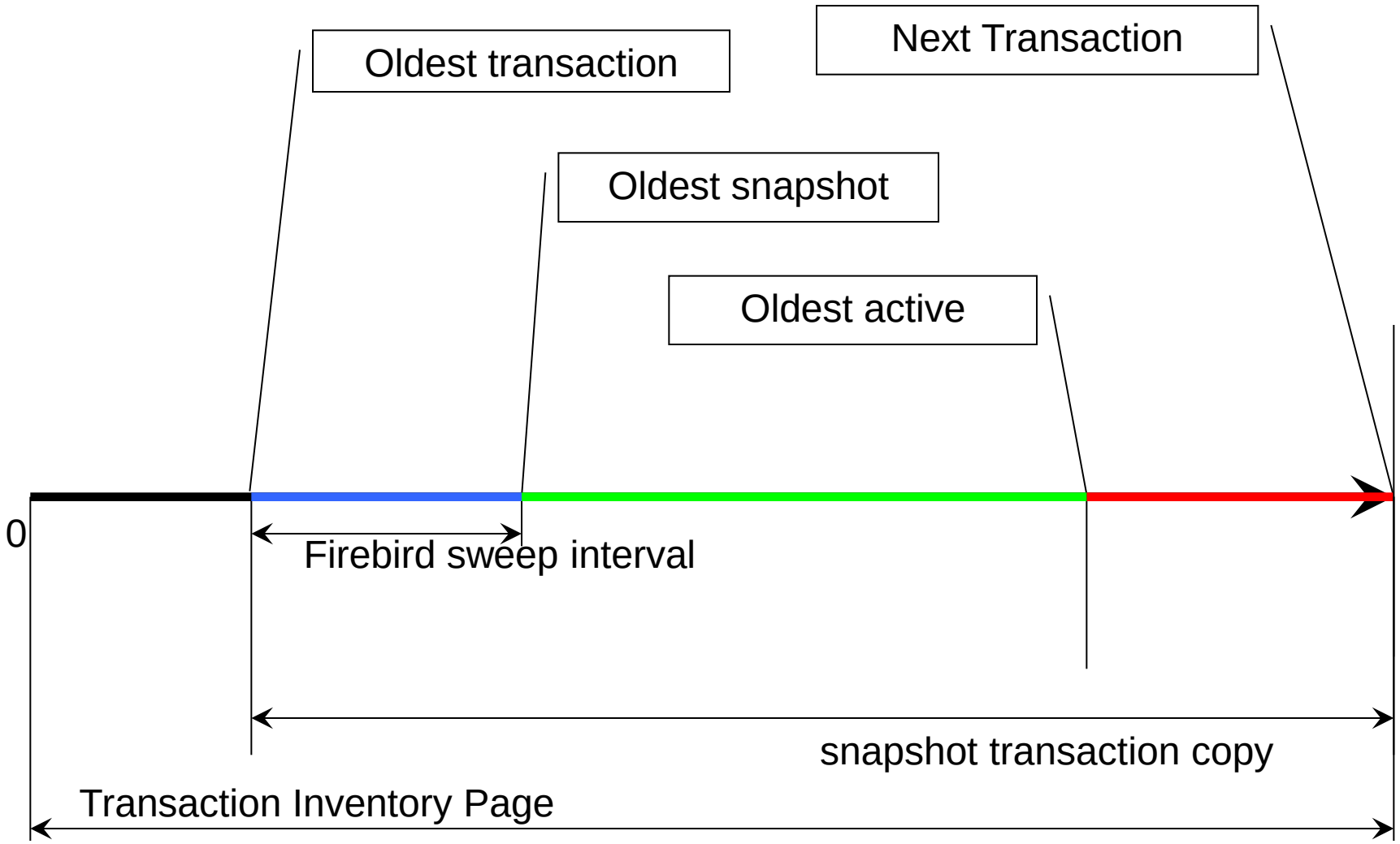
- Variable header data:
- Sweep interval: 20000
- *END*

OIT < OST <= OAT < Next

Gstat –h notes

- Sweep interval: 20000
  - Is not shown (at all) at the new database
  - If you do not see it, then sweep interval is on, and it is 20000
- Attributes force write
  - Shown only when FW is on. If you do not see it, then Forced Write is OFF (async writes, by OS)

TIP markers


Oldest transaction
Next Transaction
Oldest snapshot
Oldest active
0
Firebird sweep interval
snapshot transaction copy
Transaction Inventory Page

Tribal example

- Next transaction - newborn
- Oldest Active – the oldest alive
- Oldest Snapshot – the oldest dead, who is remembered by oldest alive (Oldest Active)
- Oldest interesting – died somewhere, need to be found and buried

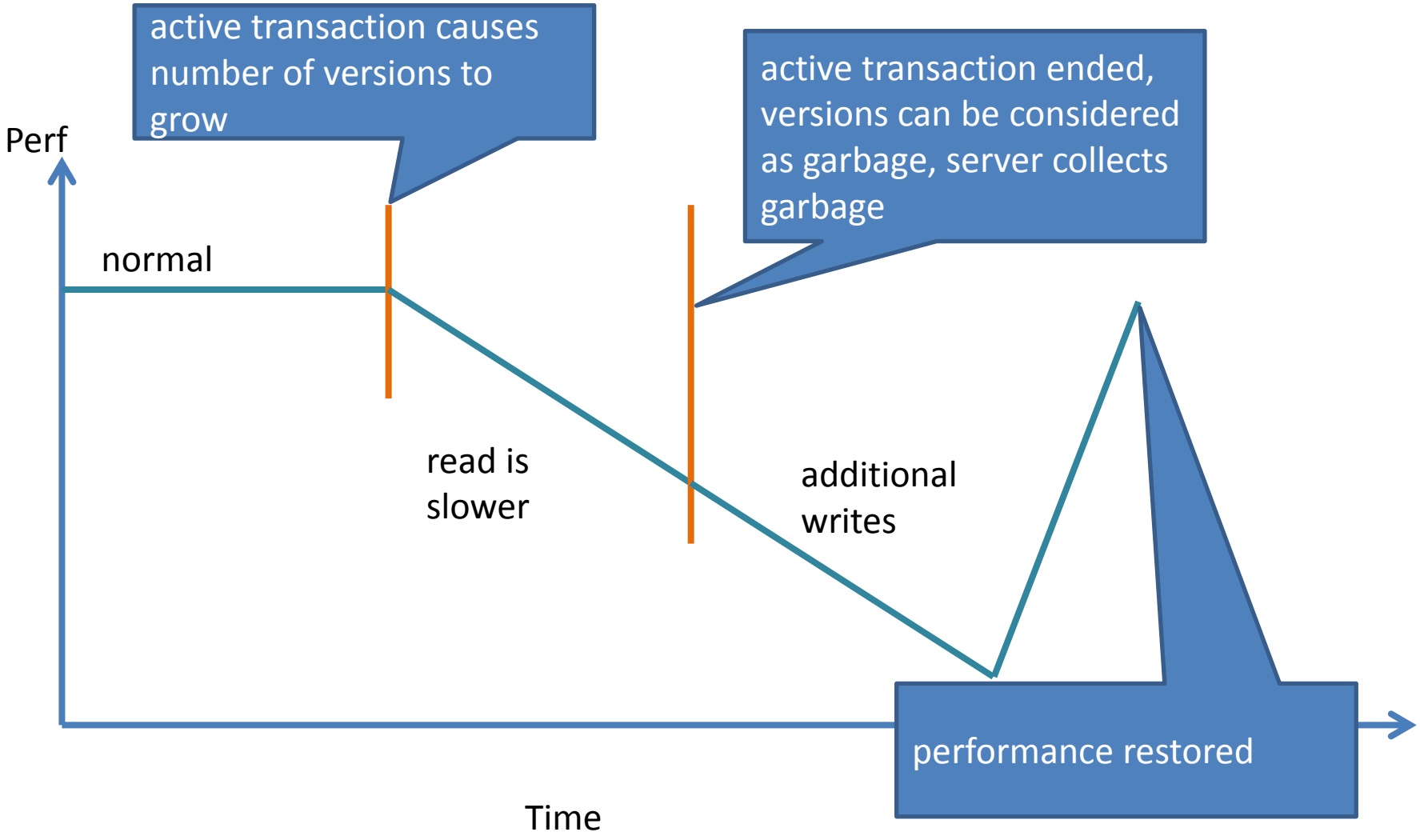
What to monitor

- Next – OAT
  - How long some active transaction lives, preventing versions to be garbage
  - Performance may degrade
- OIT freezes
  - after that OST goes forward, and OST – OIT may raise to the auto-sweep interval (default 20000), so sweep may suddenly start
  - Just an indicator of rolled back transaction

Garbage collection

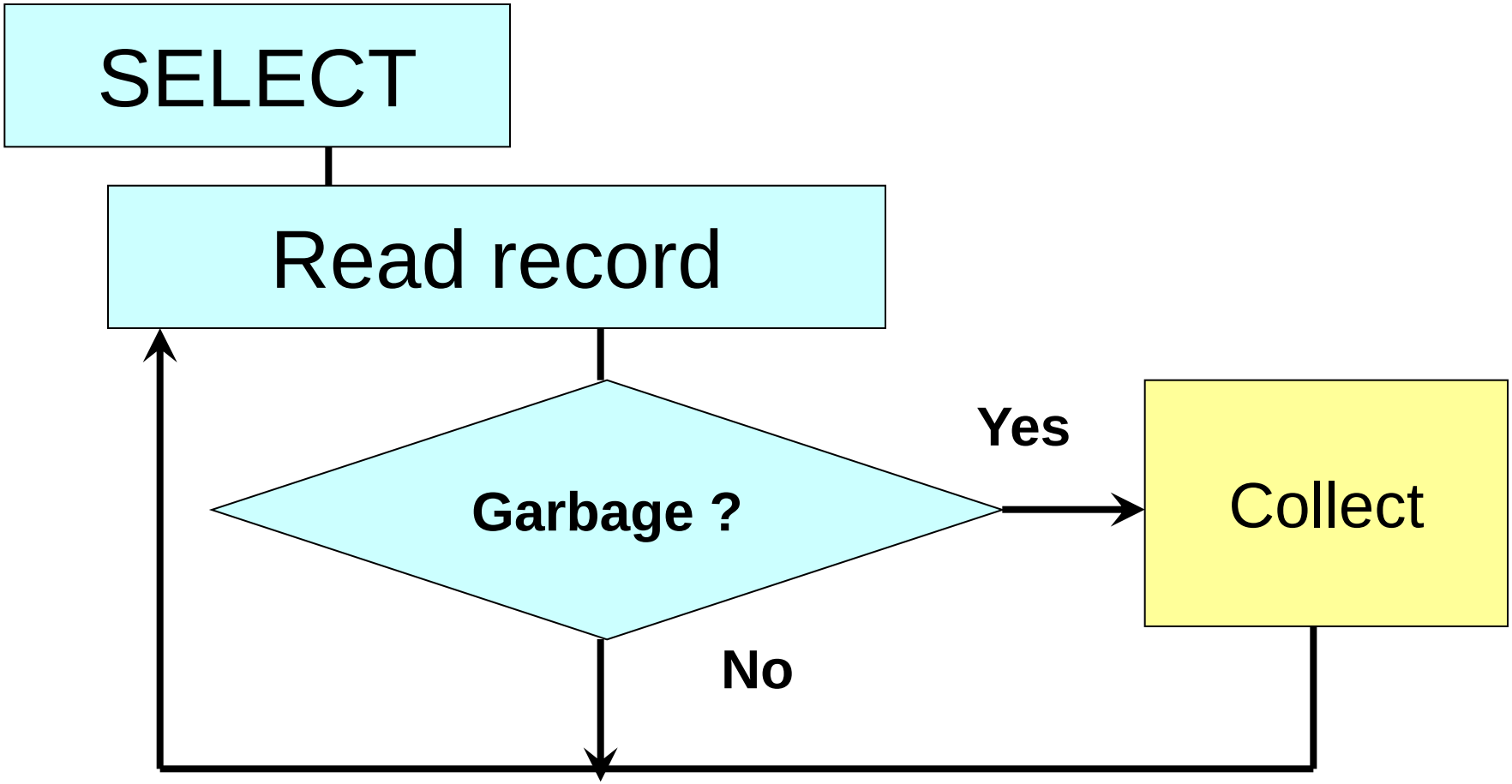
- Cooperative (gcpolicy in firebird.conf)
  - background, combined (SuperServer)
  - cooperative - (Classic, SuperClassic)
- Garbage collection works as
  - If you want to eat, clean plates first
- Auto sweep <> garbage collection
  - You can not turn off garbage collection (except isc_dpb_no_garbage_collect flag for connection)

Performance



garbage collection

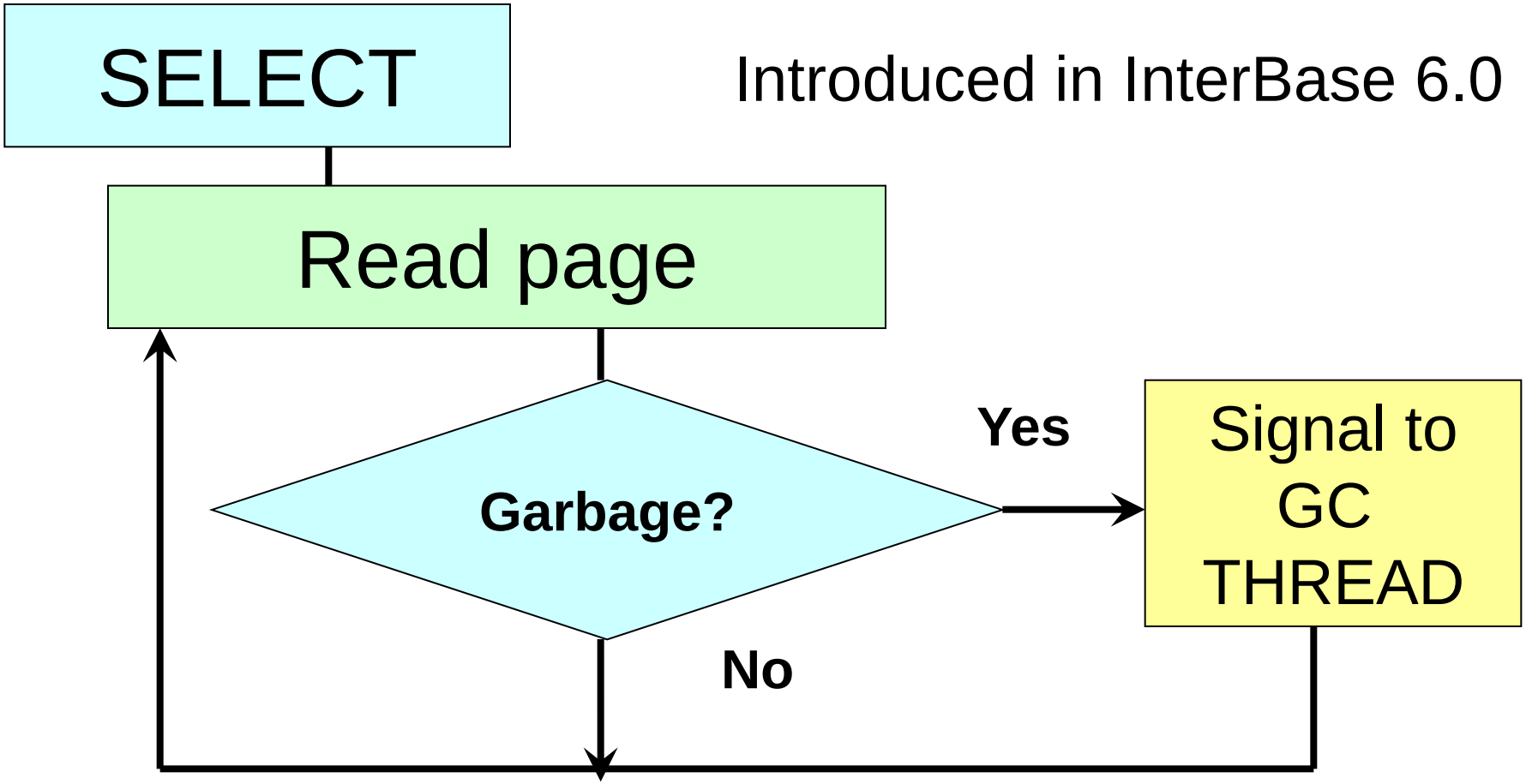
Cooperative (Classic and SuperClassic default)



Explicit garbage collection

- Classic (and SuperClassic)
- Slowdowns **select**
  - “classic" case – insert 100k records, delete, select, count - and 0 is returned after some time
- Slowdowns **update**

Background (SuperServer)



Combined (SuperServer default)

- Combined Cooperative (explicit) and Background GC
- Collect garbage as Cooperative
- Notifies GC thread about update/delete, as Background

- Garbage criteria for all GC types – version's transaction number is less than OST

Backup and garbage

- Backup is a process that makes backup copy
- It must make backup ASAP
- So, not need to collect garbage during backup
- Use gbak –b –g and gfix –sweep
- Garbage collection <> SWEEP

- always specify –g option when you tries to backup corrupted database

Test application

- By button press
  - StartTransaction
  - Update 10k records
  - Commit
- Press the button several times

What will happen?

- If pause between presses is ~4-5 seconds
  - Operation takes ~0.175 second
- Push more often
  - Operation takes 0.2 to 3.5 seconds
- Press, disconnect, connect, press
  - Operation takes 3.5 seconds

This is garbage collection…

- Cooperative – each select/update takes 3.5 seconds
- Combined – floating time, depends on frequency
- Background – always ~0.2 seconds
  - If we press button more often, GC will not succeed, and garbage will grow

Performance

- Non-proportional
- GC of 10k backversions – 3.5 seconds
- GC of 140k backversions – 10 seconds (not 35 or 50 seconds)

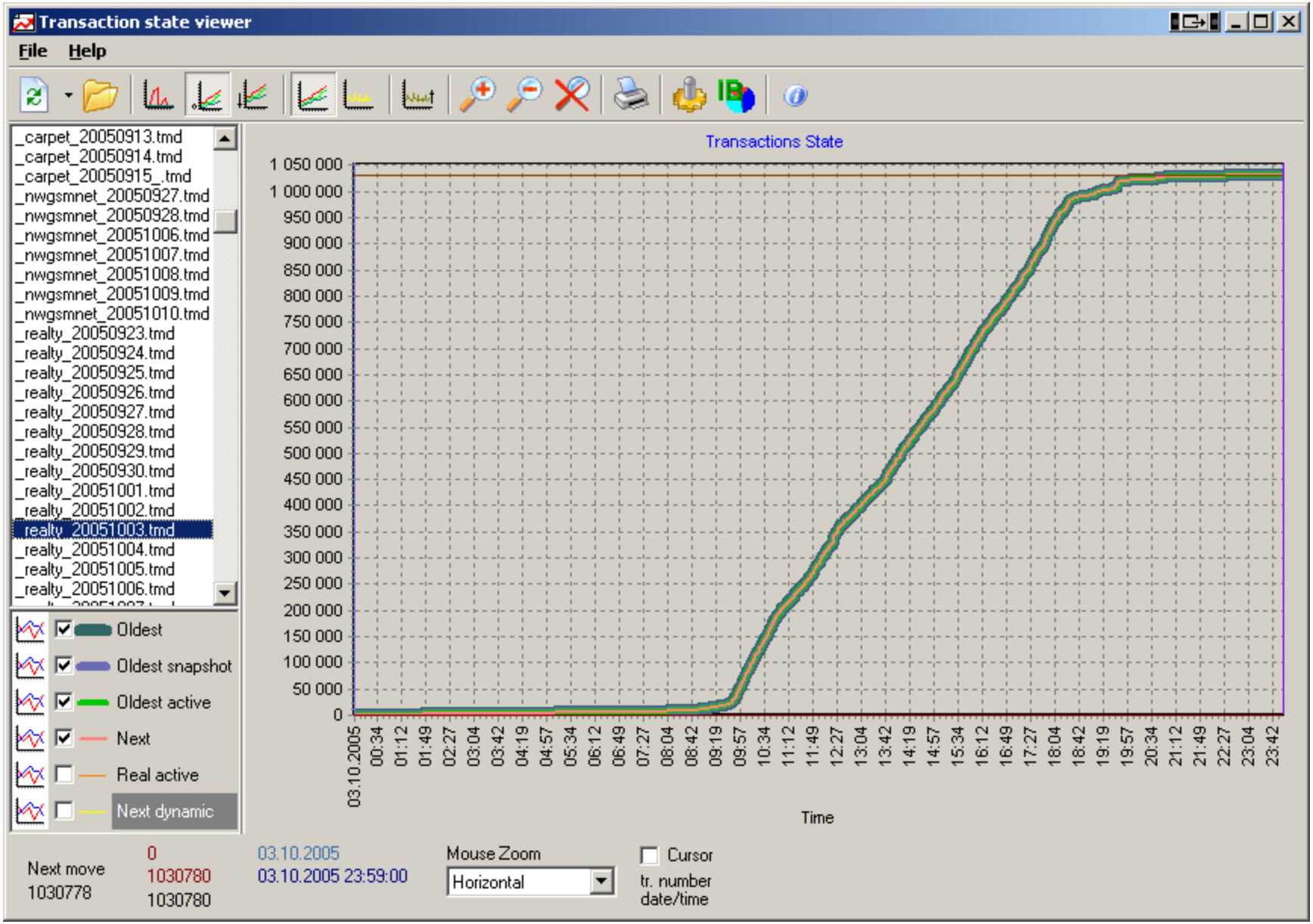
Disable GC

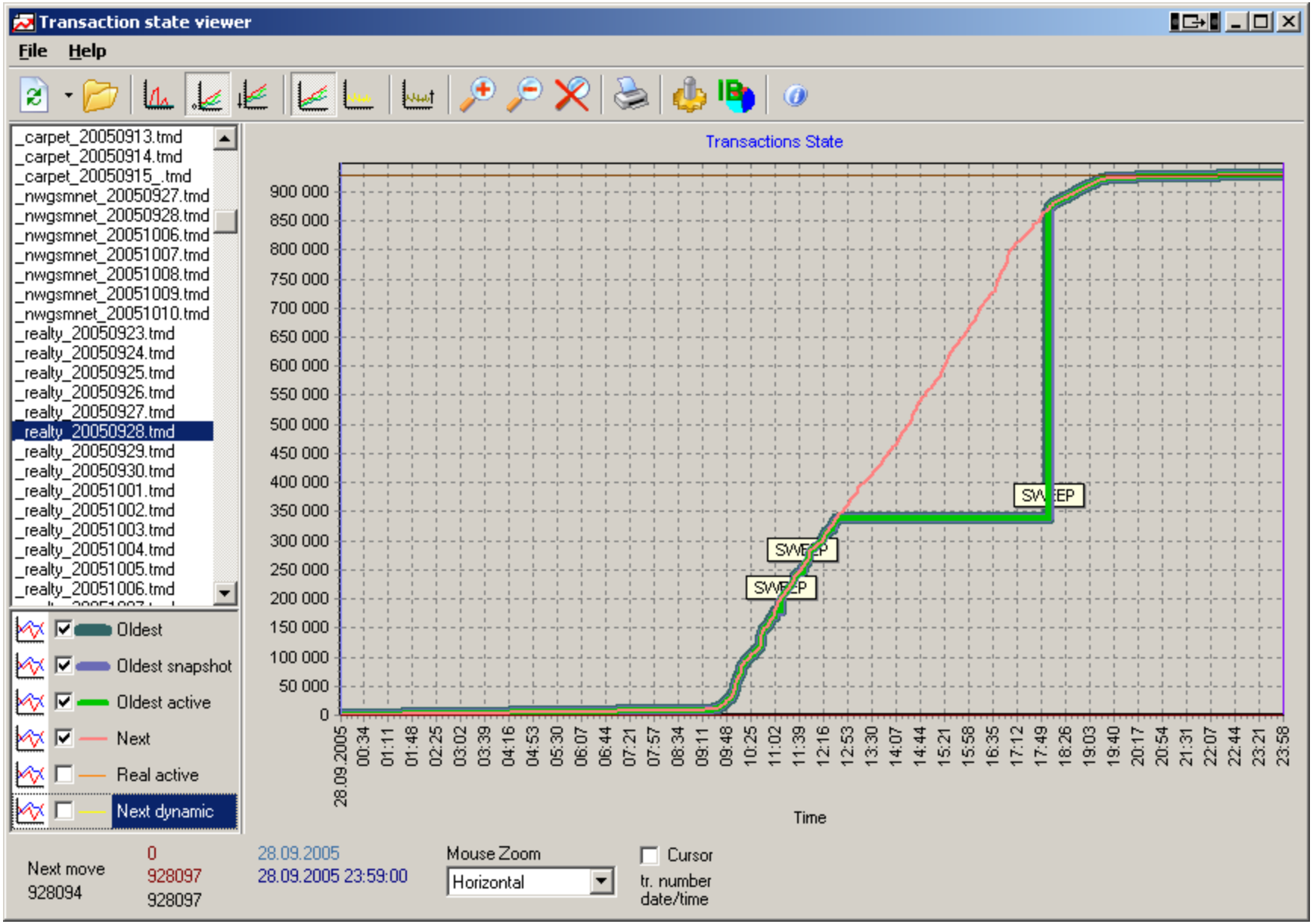
- gbak –b –g
  - Only for current gbak process
- isc_dpb_no_garbage_collect
  - Only for connection that uses this option

- All other connections will try to collect (at least check for) garbage

GC slowness

- The more indices on the table, the slower is GC

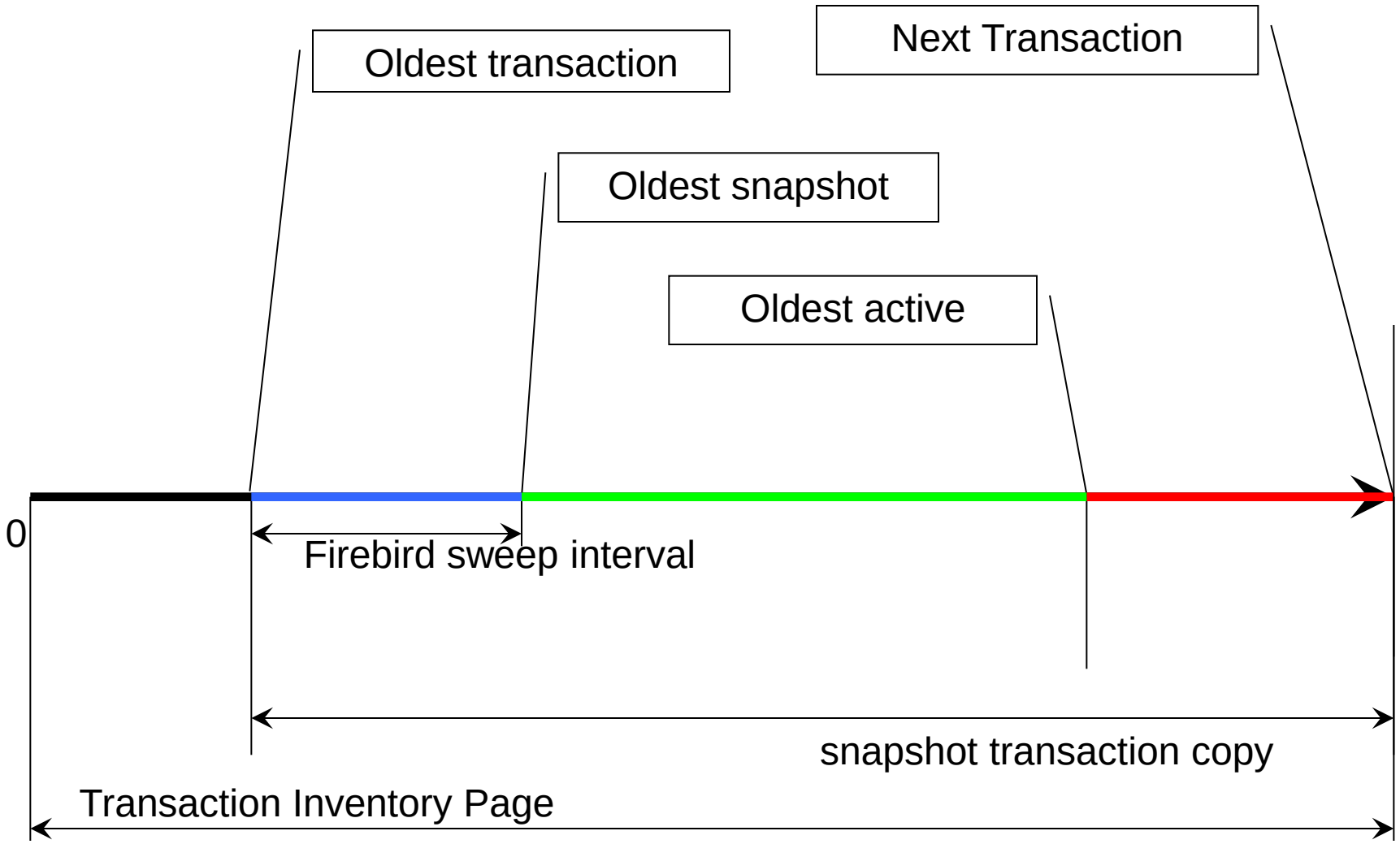

Transaction state viewer
File
Help
_carpet_20050913.tmd
_carpet_20050914.tmd
_carpet_20050915_.tmd
_nwgsmnet_20050927.tmd
_nwgsmnet_20050928.tmd
_nwgsmnet_20051006.tmd
_nwgsmnet_20051007.tmd
_nwgsmnet_20051008.tmd
_nwgsmnet_20051009.tmd
_nwgsmnet_20051010.tmd
_realty_20050923.tmd
_realty_20050924.tmd
_realty_20050925.tmd
_realty_20050926.tmd
_realty_20050927.tmd
_realty_20050928.tmd
_realty_20050929.tmd
_realty_20050930.tmd
_realty_20051001.tmd
_realty_20051002.tmd
realty_20051003.tmd
_realty_20051004.tmd
_realty_20051005.tmd
_realty_20051006.tmd
Oldest
Oldest snapshot
Oldest active
Next
Real active
Next dynamic
Transactions State
1 050 000
1 000 000
950 000
900 000
850 000
800 000
750 000
700 000
650 000
600 000
550 000
500 000
450 000
400 000
350 000
300 000
250 000
200 000
150 000
100 000
50 000
0
03.10.2005
00:34
01:12
01:49
02:27
03:04
03:42
04:19
04:57
05:34
06:12
06:49
07:27
08:04
08:42
09:19
09:57
10:34
11:12
11:49
12:27
13:04
13:42
14:19
14:57
15:34
16:12
16:49
17:27
18:04
18:42
19:19
19:57
20:34
21:12
21:49
22:27
23:04
23:42
Time
Next move
1030778
0
1030780
1030780
03.10.2005
03.10.2005 23:59:00
Mouse Zoom
Horizontal
Cursor
tr. number
date/time


Transaction state viewer
File
Help
_carpet_20050913.tmd
_carpet_20050914.tmd
_carpet_20050915_.tmd
_nwgsmnet_20050927.tmd
_nwgsmnet_20050928.tmd
_nwgsmnet_20051006.tmd
_nwgsmnet_20051007.tmd
_nwgsmnet_20051008.tmd
_nwgsmnet_20051009.tmd
_nwgsmnet_20051010.tmd
_realty_20050923.tmd
_realty_20050924.tmd
_realty_20050925.tmd
_realty_20050926.tmd
_realty_20050927.tmd
realty_20050928.tmd
_realty_20050929.tmd
_realty_20050930.tmd
_realty_20051001.tmd
_realty_20051002.tmd
_realty_20051003.tmd
_realty_20051004.tmd
_realty_20051005.tmd
_realty_20051006.tmd
Oldest
Oldest snapshot
Oldest active
Next
Real active
Next dynamic
Transactions State
900 000
850 000
800 000
750 000
700 000
650 000
600 000
550 000
500 000
450 000
400 000
350 000
300 000
250 000
200 000
150 000
100 000
50 000
0
28.09.2005
00:34
01:11
01:48
02:25
03:02
03:39
04:16
04:53
05:30
06:07
06:44
07:21
07:57
08:34
09:11
09:48
10:25
11:02
11:39
12:16
12:53
13:30
14:07
14:44
15:21
15:58
16:35
17:12
17:49
18:26
19:03
19:40
20:17
20:54
21:31
22:07
22:44
23:21
23:58
Time
SWEEP
SWEEP
SWEEP
Next move
928094
0
928097
928097
28.09.2005
28.09.2005 23:59:00
Mouse Zoom
Horizontal
Cursor
tr. number
date/time

SWEEP

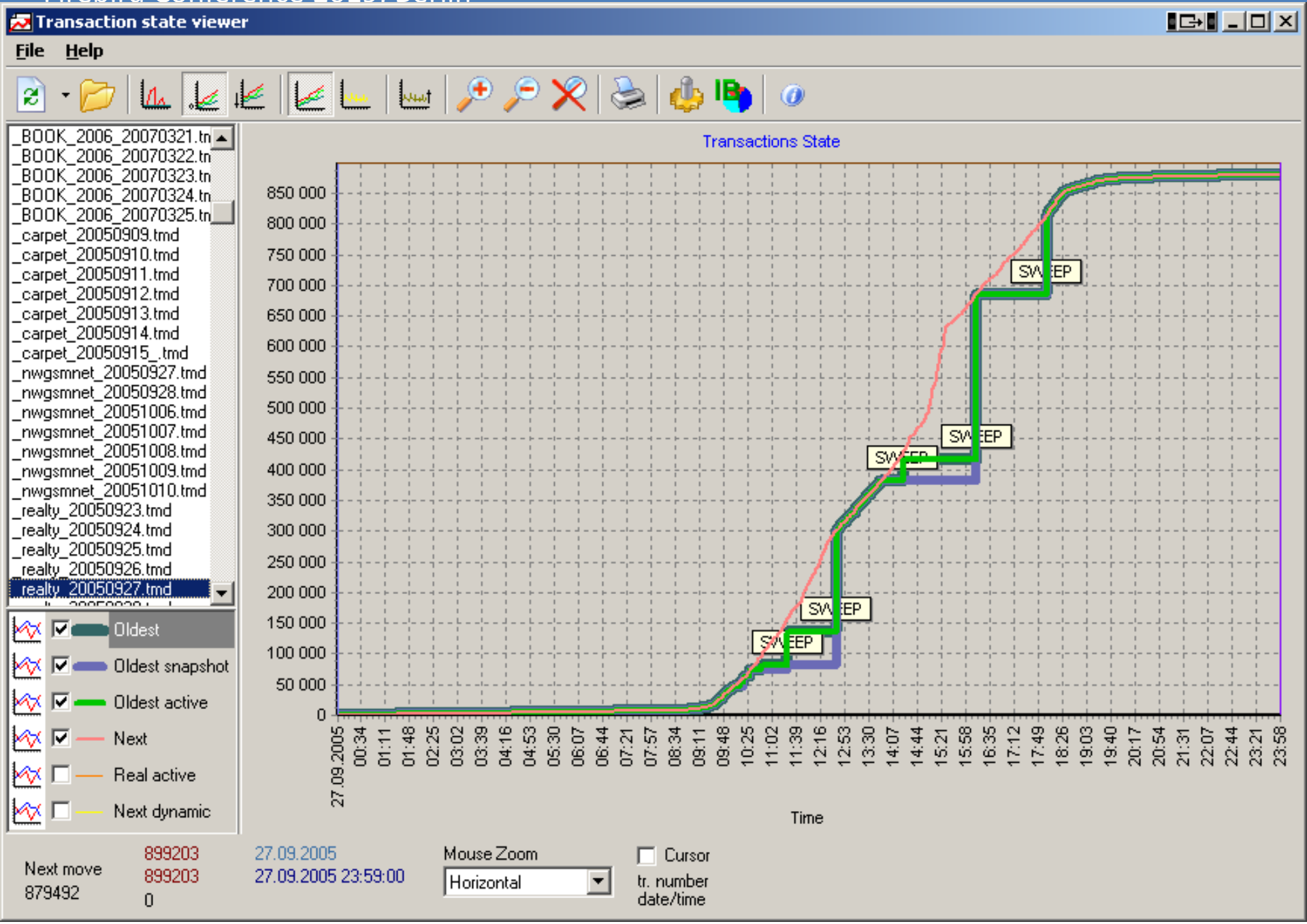
- SWEEP tasks:
  - Clean up garbage in the database
  - Move OIT forward

TIP markers

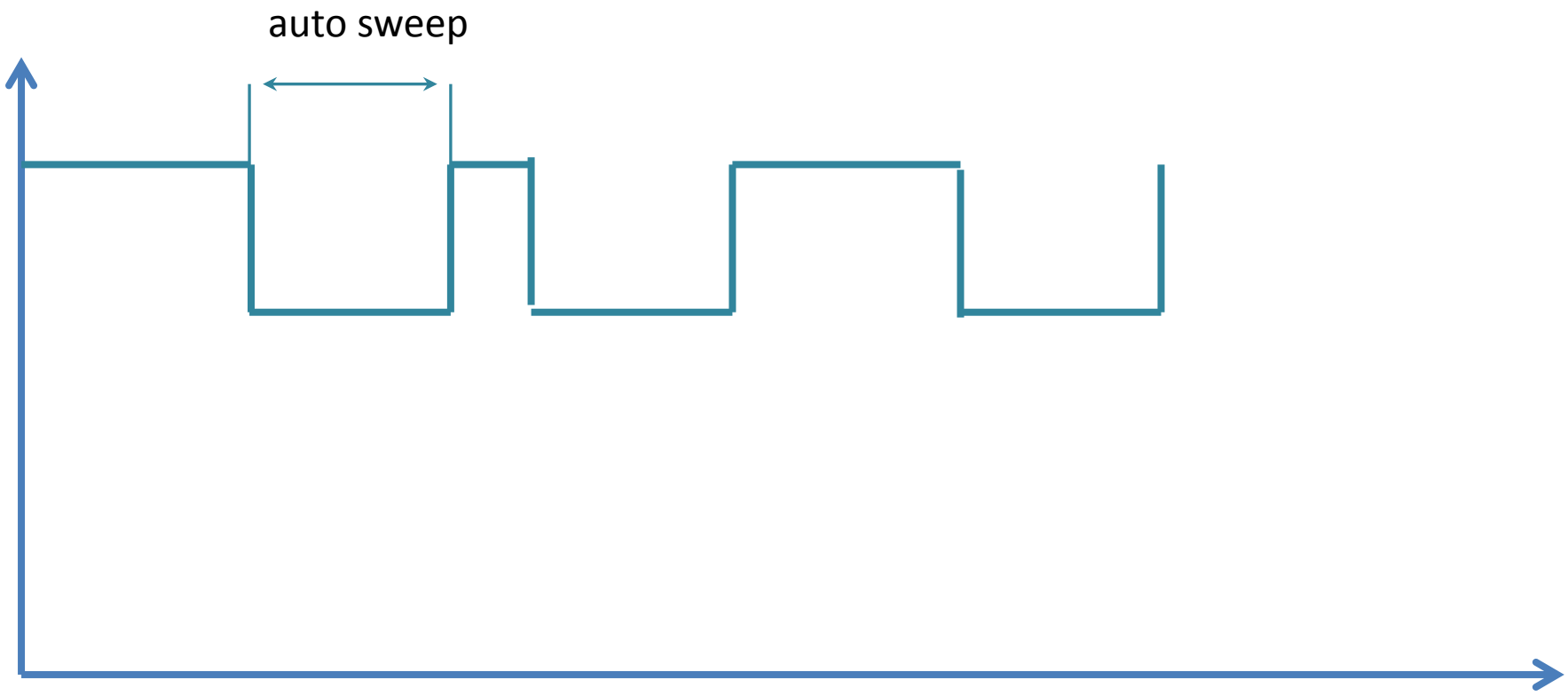

Oldest transaction
Next Transaction
Oldest snapshot
Oldest active
0
Firebird sweep interval
snapshot transaction copy
Transaction Inventory Page

Sweep disadvantages

- Means cleaning the whole db file
- Sweep can't fully work if there are long-running transactions
  (look at OIT-OST-OAT-Next numbers in the firebird.log on sweep start and end)
- Auto-sweep can start at any moment
  - Periodic slowdowns? Yes, it's auto-sweep.
- Firebird 3 introduces "swept flag" on data pages
  - only modified pages are scanned during sweep
  - first sweep after restore scans all pages and modifies them


Transaction state viewer
File
Help
_BOOK_2006_20070321.tn
_BOOK_2006_20070322.tn
_BOOK_2006_20070323.tn
_BOOK_2006_20070324.tn
_BOOK_2006_20070325.tn
_carpet_20050909.tmd
_carpet_20050910.tmd
_carpet_20050911.tmd
_carpet_20050912.tmd
_carpet_20050913.tmd
_carpet_20050914.tmd
_carpet_20050915_.tmd
_nwgsmnet_20050927.tmd
_nwgsmnet_20050928.tmd
_nwgsmnet_20051006.tmd
_nwgsmnet_20051007.tmd
_nwgsmnet_20051008.tmd
_nwgsmnet_20051009.tmd
_nwgsmnet_20051010.tmd
_realty_20050923.tmd
_realty_20050924.tmd
_realty_20050925.tmd
_realty_20050926.tmd
realty_20050927.tmd
Oldest
Oldest snapshot
Oldest active
Next
Real active
Next dynamic
Transactions State
850 000
800 000
750 000
700 000
650 000
600 000
550 000
500 000
450 000
400 000
350 000
300 000
250 000
200 000
150 000
100 000
50 000
0
27.09.2005
00:34
01:11
01:48
02:25
03:02
03:39
04:16
04:53
05:30
06:07
06:44
07:21
07:57
08:34
09:11
09:48
10:25
11:02
11:39
12:16
12:53
13:30
14:07
14:44
15:21
15:58
16:35
17:12
17:49
18:26
19:03
19:40
20:17
20:54
21:31
22:07
22:44
23:21
23:58
Time
Next move
879492
899203
899203
0
27.09.2005
27.09.2005 23:59:00
Mouse Zoom
Horizontal
Cursor
tr. number
date/time

Performance with auto sweep

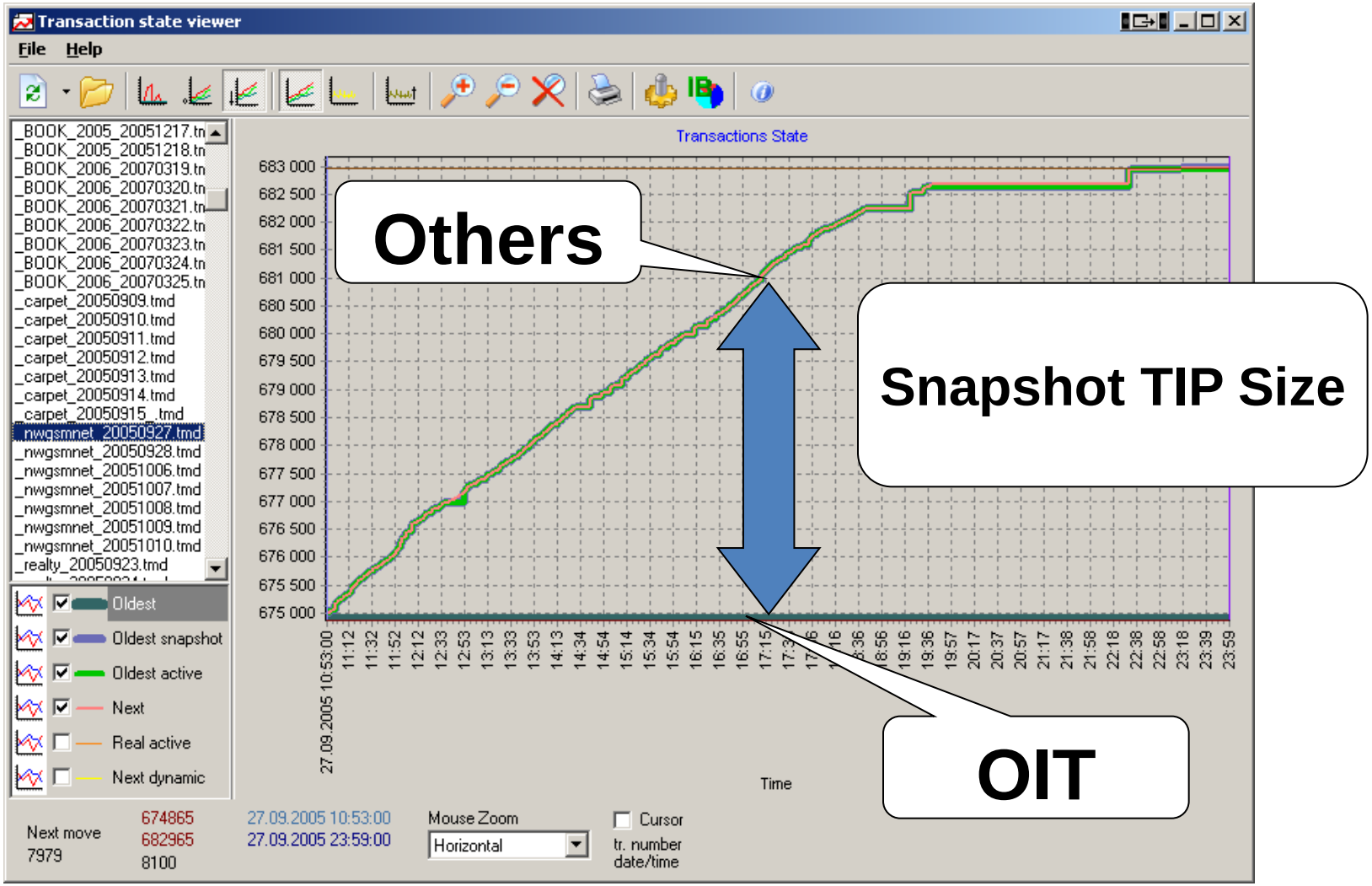


Turning off auto-sweep

- you do not want Firebird to scan all or part of the database at unpredictable time

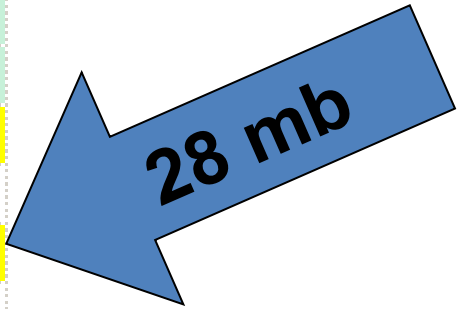
- gfix database.fdb –housekeeping 0

Turning off auto-sweep



Snapshot TIP Size

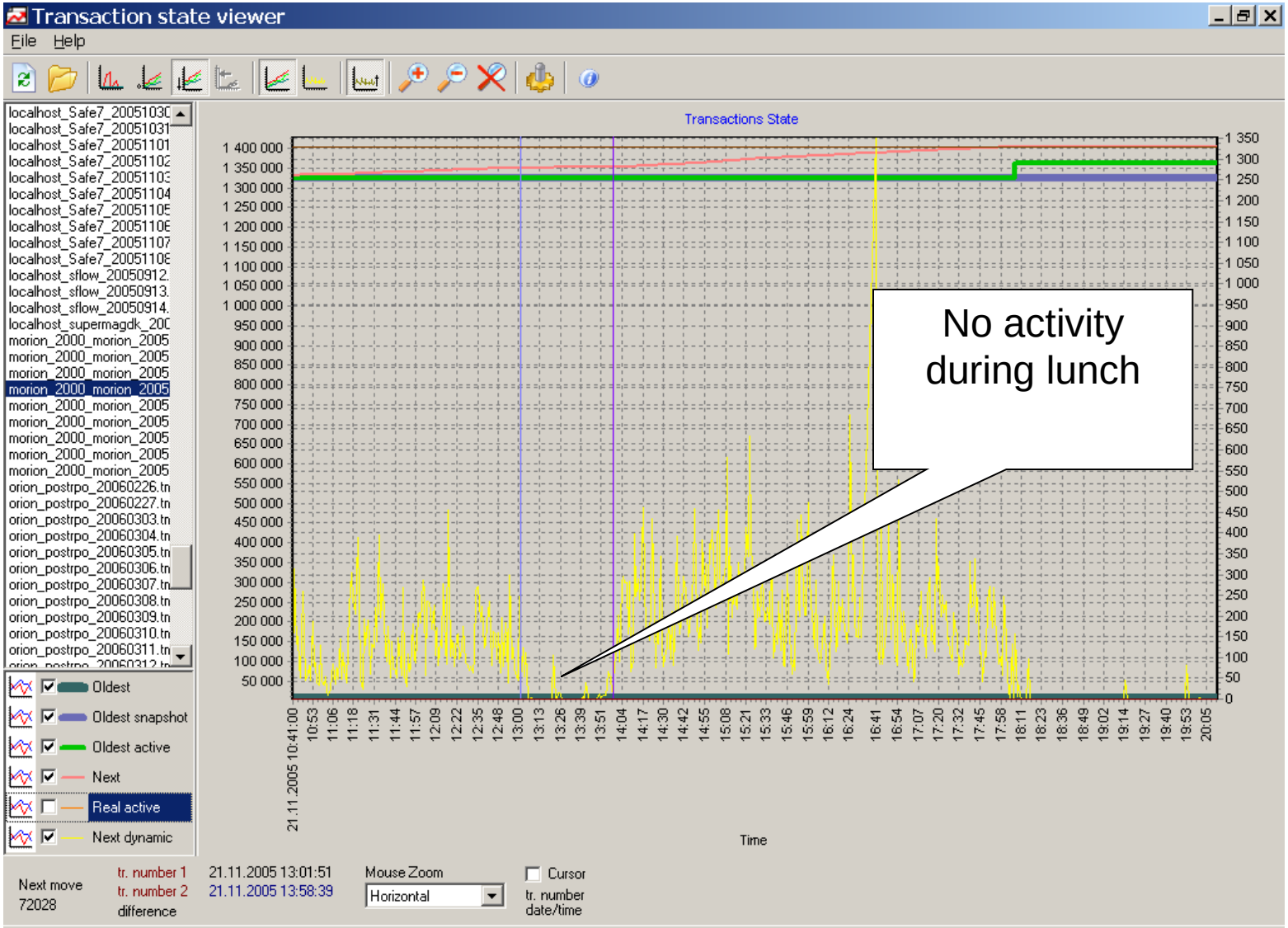
Creation date	05.06.2003 10:02:19
Statistics date	22.06.2004 20:55:46
Page size	8192
Forced Write	ON
Dialect	1
OnDiskStructure	10.0
Sweep interval	0
Oldest transaction	839568
Oldest snapshot	112430561
Oldest active	112430625
Next transaction	112431441
Sweep gap (snapshot - oldest)	111590993
TIP size	3432 pages, 27457 kilobytes
Snapshot TIP size	111591873 transactions, 27252 kilobytes
Active transactions	816, 0% of daily average
Transactions per day	292790, for 384 days

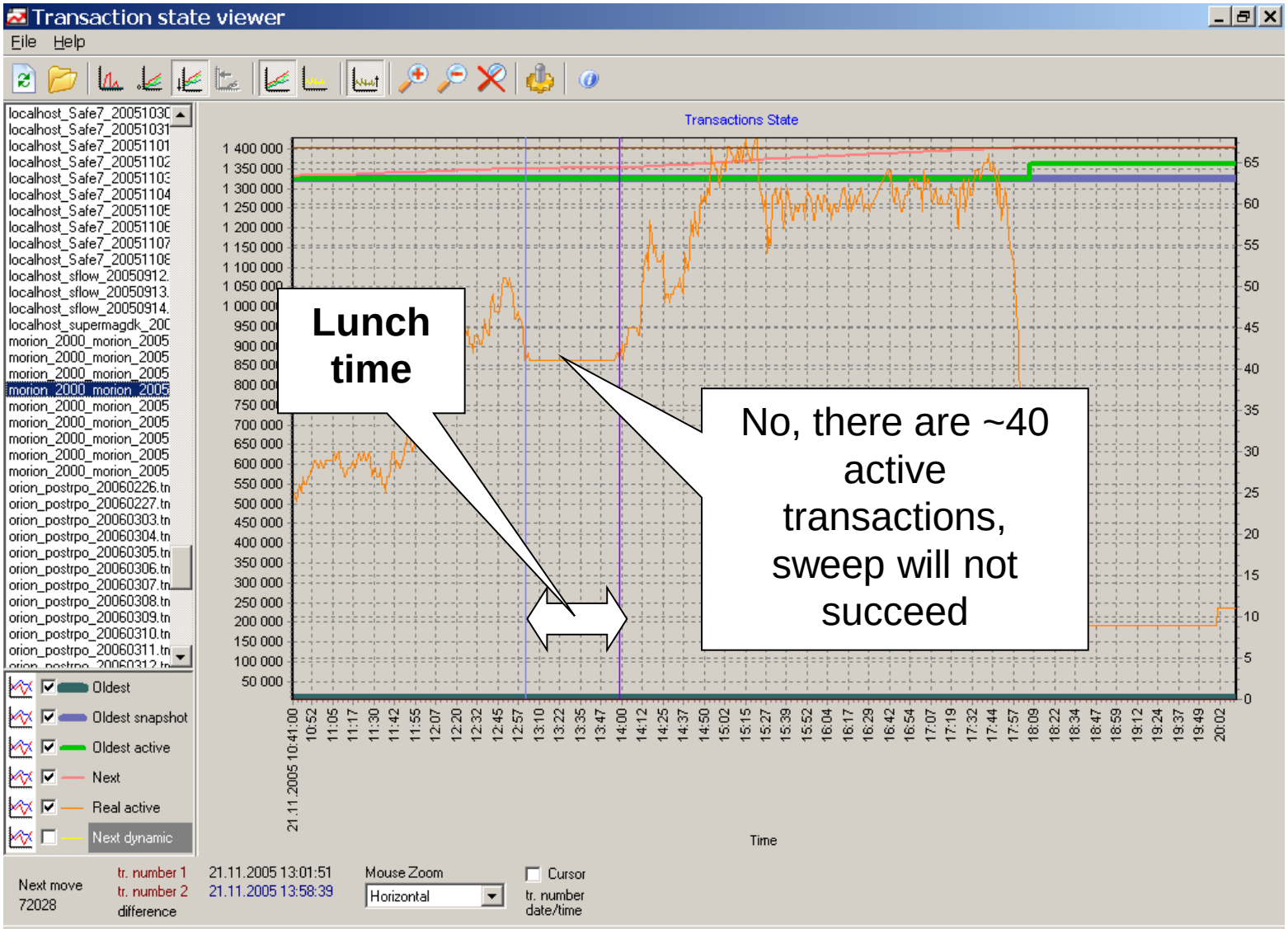


OIT freezes when

- Connection lost
- Rollback of lot of changes
  - StartTransaction
    …. update/delete/insert
    Rollback
    - If manyChanges then
      Rollback (let others collect garbage)
    else
      Rollback through savepoints and do **Commit**

When to run sweep




Transaction state viewer
File Help
Transactions State
Lunch time
No, there are ~40 active transactions, sweep will not succeed
Oldest
Oldest snapshot
Oldest active
Next
Real active
Next dynamic
Time
Next move
72028
tr. number 1
tr. number 2
difference
21.11.2005 13:01:51
21.11.2005 13:58:39
Mouse Zoom
Horizontal
Cursor
tr. number
date/time

- Find appropriate time to run sweep manually
  - At the lowest activity of your system
    - Usually at night

NEXT…