

Working With Very Large Firebird Databases

Some Aspects To Consider

Paul Reeves
IBPhoenix

mail: preeves at ibphoenix.com

About the speaker

I work for IBPhoenix providing technical support.

I maintain the windows installer for Firebird and do the Windows builds.

Introduction

Background to this talk – Why?

Three questions to answer...

- Does transaction performance deteriorate and is it disproportionate to the increase in size?
- Is time taken for an admin task proportional to database size?
- Are admin tasks disproportionately affected by concurrent user activity as database size increases?

What are (not) looking at today?

- The focus of this talk is on administration issues rather than performance in general
- Performance tends to be application specific
- This talk is more interested in how :
 - Gstat
 - Gbak
 - NBackup
 - Sweep

behave as the db grows in size.

What is a very large database?

- How long is a piece of string?
- It changes with the technology.
- Larger than available RAM
- Perhaps insufficient disc space to manipulate copies and backups easily.
- Disk I/O inadequate for task of backing up/restoring large data volumes within an acceptable time frame.

Setting up a test environment

- **4 core AMD x64**
- **16 GB RAM**
- **Windows 2003 X64 server**
- **Firebird 2.5.2 Super Server, Vanilla config**
- **RAID 10 with H/W controller and write cache.**
- **SSD**

The test databases

- SDB – 'small' database
less than available RAM, 8 M rows
- MDB – 'medium' size
larger than available RAM, 24 M rows
- QLDB – 'Quite Large'
twice the size of medium db, 48 M rows
- VLDB – 'Very Large'
four times the size of medium db, 80 M rows.

Simulated Load Testing

- Based on TPC-C
- It is a simple product/order database
- Txns are randomised in an attempt to model real world.
- TPC-C performance is measured in new-order transactions per minute - tpmC

Caveats

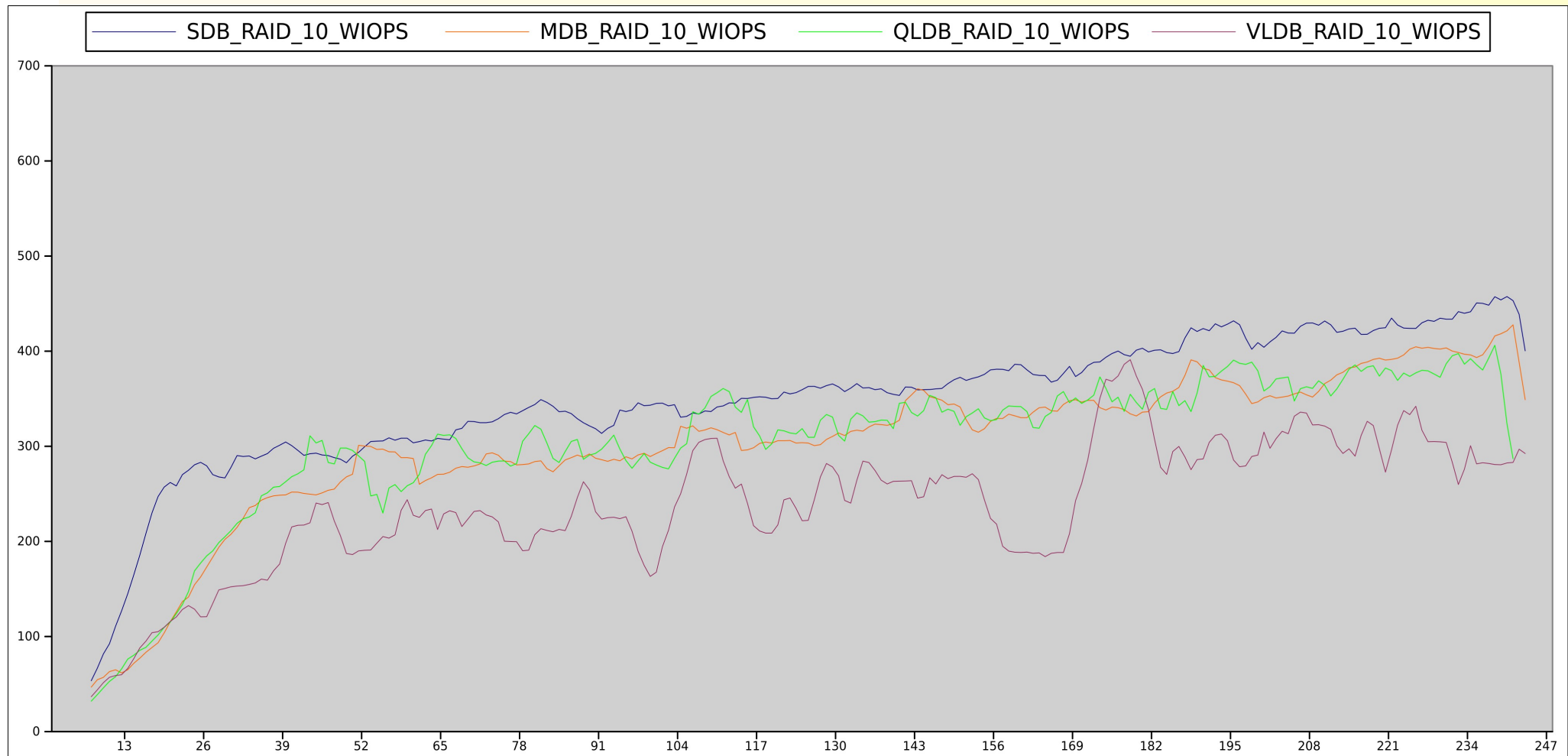
- **Not scientifically valid.**
- **Test suite takes ? days to run**
- **Results should be averaged over 5 test runs**
- **These tests relate to this test config.**
- **Other versions of Firebird may behave differently**
- **Other O/S may behave differently.**
- **Other databases or use cases may be different.**
- **Much more analysis is required.**

Baseline Performance Test - 1

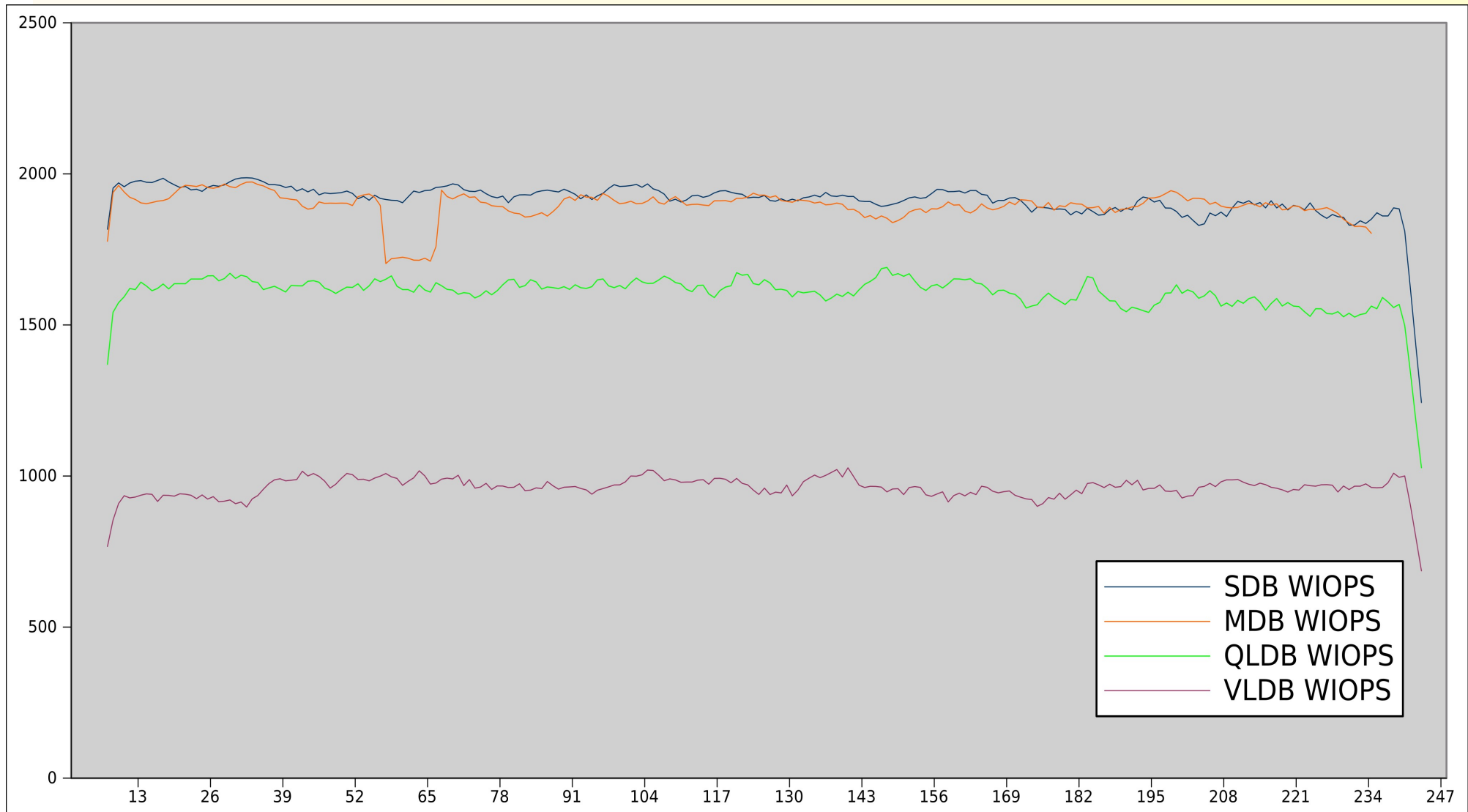
Run tpc-c benchmark test for one hour on each database simulating 25 concurrent users.

Database	Txn per min.
SDB_RAID10	410
MDB_RAID10	392
QLDB_RAID10	360
VLDB_RAID10	288
SDB_SSD	2232
MDB_SSD	2235
QLDB_SSD	1938
VLDB_SSD	1112

Baseline Performance WIOPS – RAID 10



Baseline Performance WIOPS – SSD

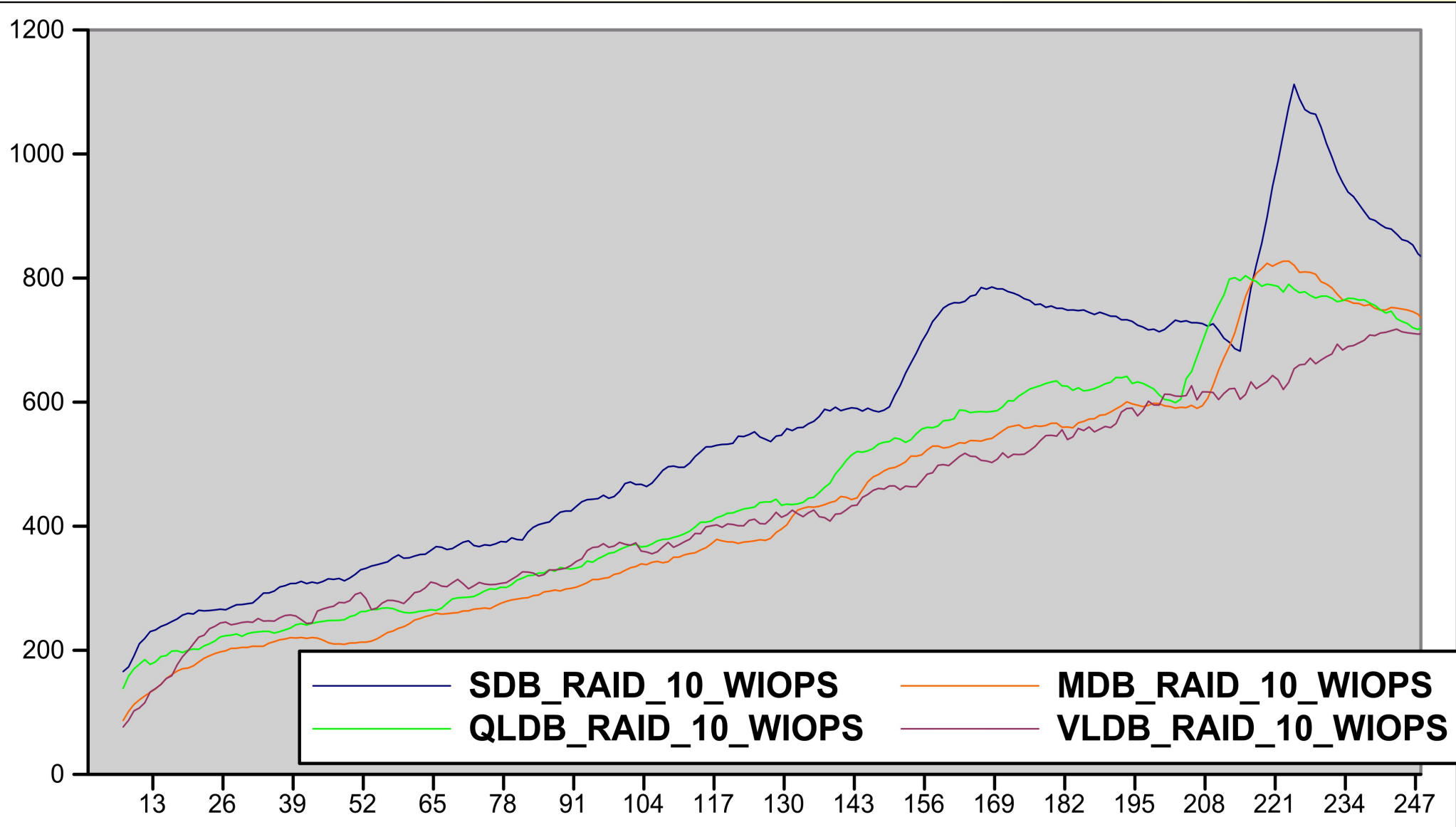


Baseline Performance Test - 3

Database	Txn per min (03)	Txn per min. (02)	Txn per min. (01)
SDB_RAID10	1295	907	410
MDB_RAID10	1007	932	392
QLDB_RAID10	1033	710	360
VLDB_RAID10	852	529	288
SDB_SSD	1455	1769	2232
MDB_SSD	1007	1810	2235
QLDB_SSD	1652	1771	1938
VLDB_SSD	1455	1301	1112

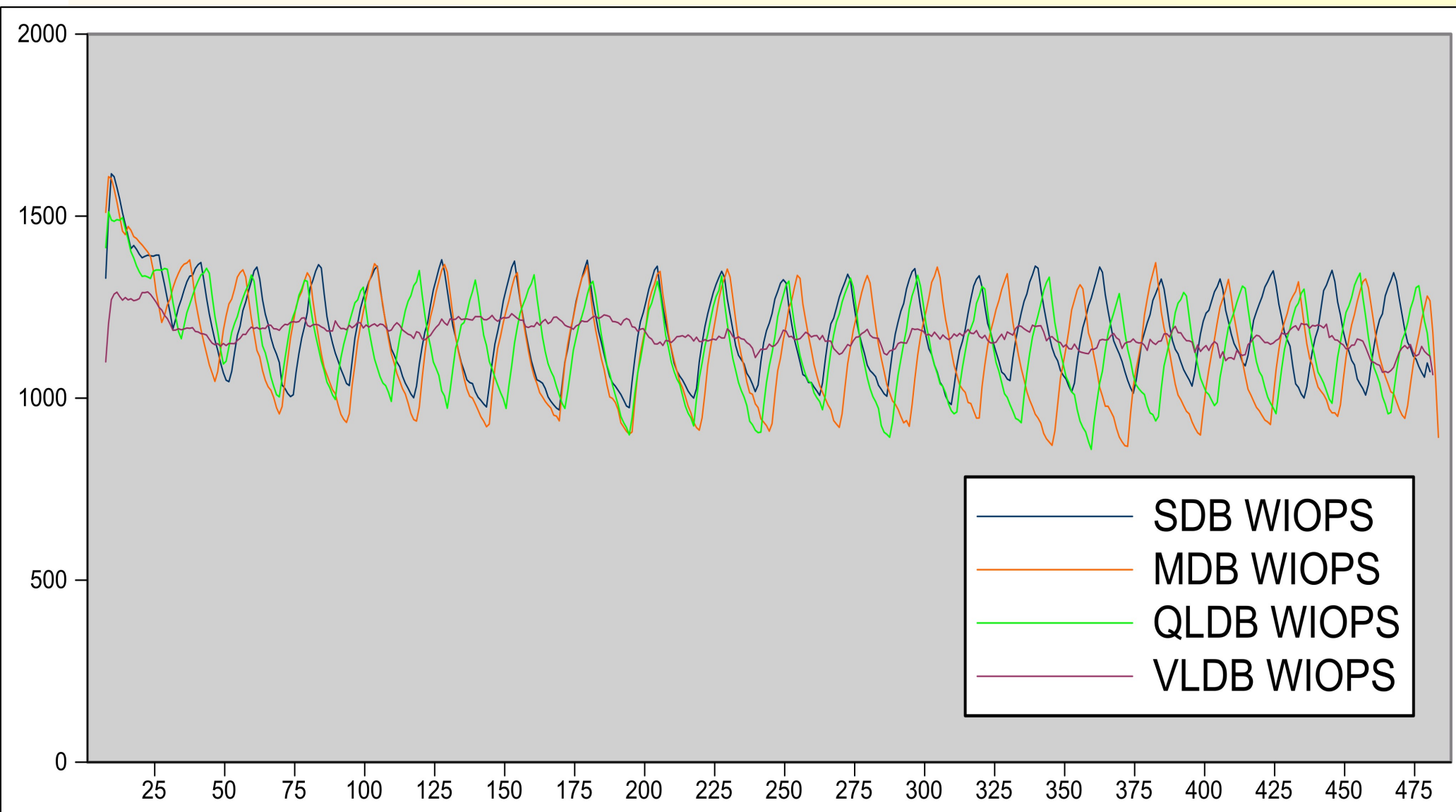
Baseline Performance (03)

WIOPS – RAID 10



Baseline Performance (03)

WIOPS – SSD



Impact of gstat -r fb inactive

Database	Size (GB)	Fb Inactive	MB/s		
SDB_RAID10	8.73	00:00:42	212.85		
MDB_RAID10	22.70	00:04:18	90.1		
QLDB_RAID10	44.00	00:07:44	97.1		
VLDB_RAID10	81.10	00:13:30	102.53		
SDB_SSD	8.73	00:00:30	297.98		
MDB_SSD	22.70	00:03:25	113.39		
QLDB_SSD	44.00	00:06:51	109.63		
VLDB_SSD*	81.10	00:19:06	72.47		

(* sweeper activated)

Impact of gstat -r when fb active

Database	Size (GB)	FB Active	MB/s
SDB_RAID10	8.73	00:04:34	32.63
MDB_RAID10	22.7	00:15:18	25.32
QLDB_RAID10	44	00:26:25	28.43
VLDB_RAID10	81.1	00:43:18	31.97
SDB_SSD	8.73	00:01:58	75.76
MDB_SSD	22.7	00:08:38	44.87
QLDB_SSD	44	00:17:44	42.35
VLDB_SSD (*)	81.1	00:26:52	51.52

(* uninvited SWEEPER)

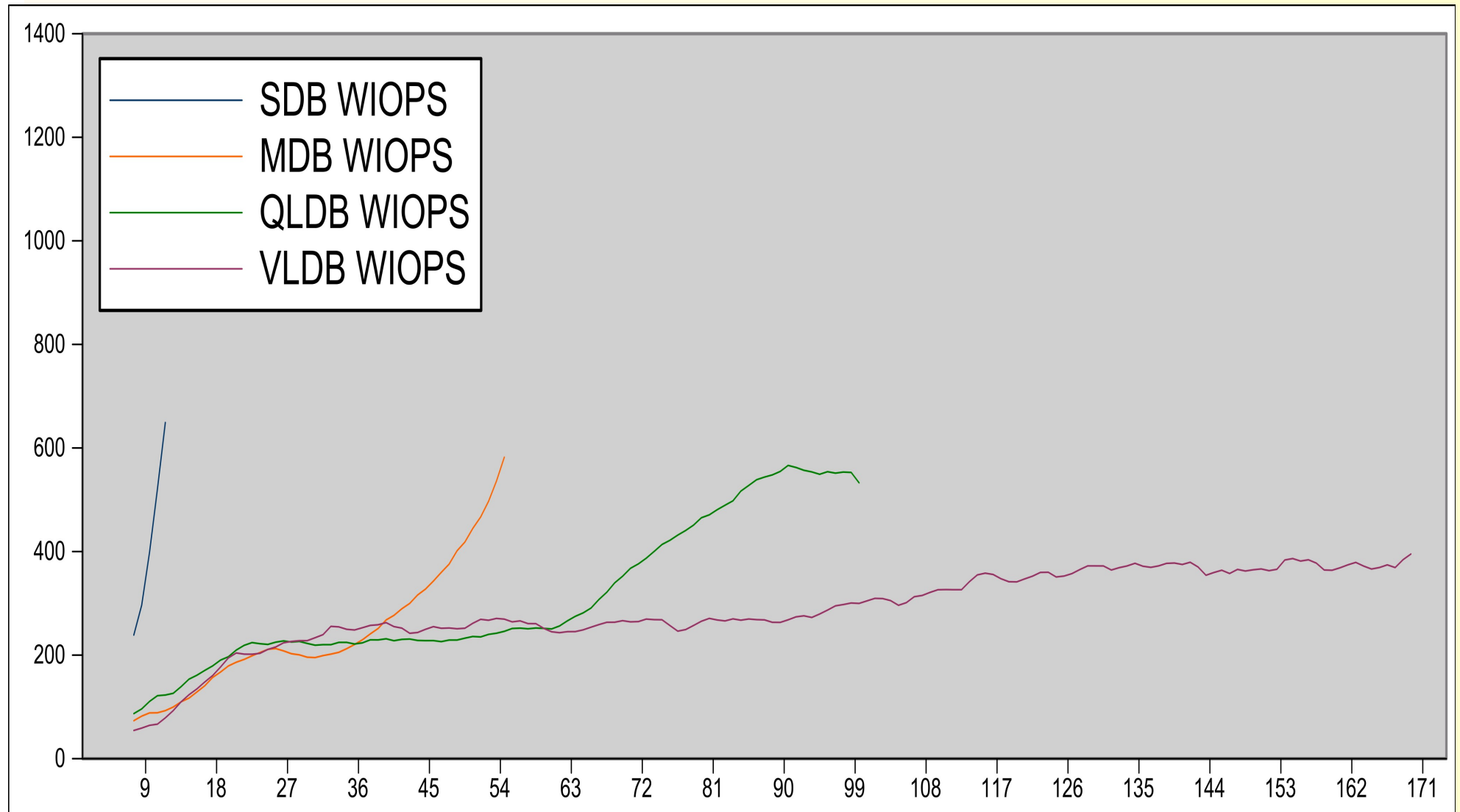
Impact of gstat -r - full comparison

Database	Size (GB)	Fb Inactive	MB/s	FB Active	MB/s
SDB_RAID10	8.73	00:00:42	212.85	00:04:34	32.63
MDB_RAID10	22.7	00:04:18	90.1	00:15:18	25.32
QLDB_RAID10	44	00:07:44	97.1	00:26:25	28.43
VLDB_RAID10	81.1	00:13:30	102.53	00:43:18	31.97
SDB_SSD	8.73	00:00:30	297.98	00:01:58	75.76
MDB_SSD	22.7	00:03:25	113.39	00:08:38	44.87
QLDB_SSD	44	00:06:51	109.63	00:17:44	42.35
VLDB_SSD (uninvited SWEEPER)	81.1	00:19:06	72.47	00:26:52	51.52

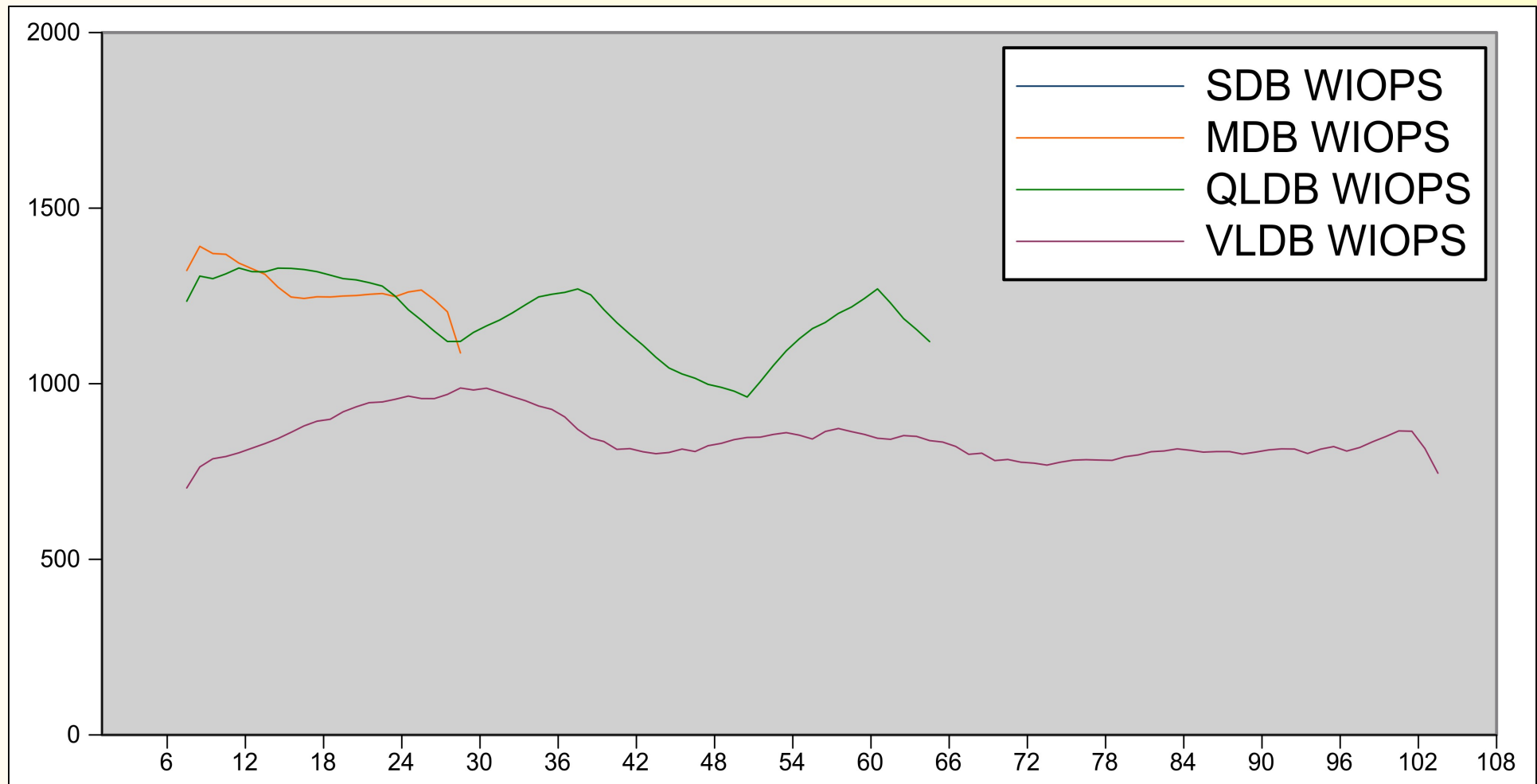
Impact of gstat -r – tpmC

Database	Txn per min (gstat)	Txn per min (Baseline)
SDB_RAID10	787	1295
MDB_RAID10	412	1007
QLDB_RAID10	487	1033
VLDB_RAID10	329	852
SDB_SSD	2365	1455
MDB_SSD	1951	1007
QLDB_SSD	1791	1652
VLDB_SSD	(*) 1040	1455

GStat and FB WIOPS RAID 10



GStat and FB WIOPS - SSD



Using GStat - Summary

- Duration approx 3 times longer when fire-bird is active.
- Impacts users when using RAID 10.
- SSD seems to cope with the performance hit.
- Safe to run at any time (with caveats).

Using GBak

- Too slow to be practical with large databases.
- Backing up an active database takes a long time and slows up users.
- Restoring is slow – over ten hours for VLDB on RAID 10 and even longer on SSD card.
- Only use is for changing page size.
- Rule of thumb :
 - restore requires 1 hour per 10 GB.

Nbackup

- Designed to backup VLDB's
- Works while firebird is active (just like gbak)
- What is the cost in time for the SYSDBA?
- What is the impact on users?

Nbackup with direct I/O (default for Win)

Database	FB Inactive	FB Active	Time Cost
SDB_RAID10	00:03:09	00:55:27	17.6
MDB_RAID10	00:09:57	01:37:58	9.85
QLDB_RAID10	00:25:53	No data	
VLDB_RAID10	No data	No data	
SDB_SSD	00:03:53	00:06:27	1.66
MDB_SSD	00:12:19	00:21:42	1.76
QLDB_SSD	00:31:23	No data	
VLDB_SSD	No data	No data	

NBackup without direct I/O

Database	FB Inactive	FB Active	Time Cost
SDB_RAID10	0:02:57	0:21:10	7.18
MDB_RAID10	0:10:07	0:56:57	5.63
QLDB_RAID10	0:26:40	No data	
VLDB_RAID10	No data	No data	
SDB_SSD	0:02:33	0:03:16	1.28
MDB_SSD	0:10:31	0:23:54	2.27
QLDB_SSD	0:33:19	No data	
VLDB_SSD	No data	No data	

NBackup - -D ON or OFF ?

Database	FB Inactive		FB Active	
	-D ON	-D OFF	-D ON	-D OFF
SDB_RAID10	00:03:09	0:02:57	00:55:27	0:21:10
MDB_RAID10	00:09:57	0:10:07	01:37:58	0:56:57
QLDB_RAID10	00:25:53	0:26:40	0	0
VLDB_RAID10	0	0	0	0
SDB_SSD	00:03:53	0:02:33	00:06:27	0:03:16
MDB_SSD	00:12:19	0:10:31	00:21:42	0:23:54
QLDB_SSD	00:31:23	0:33:19	0	0
VLDB_SSD	0	0	0	0

NBackup and xcopy

Database	FB Inactive	FB Active	Time Cost
SDB_RAID10	0:01:04	0:02:28	2.31
MDB_RAID10	0:05:30	0:28:37	5.2
QLDB_RAID10 (*)	0:12:08	1:16:36	6.31
VLDB_RAID10 (*)	0:24:14	1:12:22	2.99
SDB_SSD	0:01:39	0:01:57	1.18
MDB_SSD	0:06:06	0:08:18	1.36
QLDB_SSD (*)	0:13:23	0:18:06	1.35
VLDB_SSD (*)	0:24:36	0:29:12	1.19
(* Diff. Page size)			

NBackup vs. xcopy

Database	FB Inactive			FB Active		
	nbak	xcopy		nbak	xcopy	
SDB_RAID10	00:03:09	00:01:04	2.95	00:55:27	00:02:28	22.48
MDB_RAID10	00:09:57	00:05:30	1.81	01:37:58	00:28:37	3.42
QLDB_RAID10 (*)	00:33:04	00:12:08	2.73	04:51:55	01:16:36	3.81
VLDB_RAID10 (*)	01:10:45	00:24:14	2.92	03:37:58	01:12:22	3.01
SDB_SSD	00:03:53	00:01:39	2.35	00:06:27	00:01:57	3.31
MDB_SSD	00:12:19	00:06:06	2.02	00:21:42	00:08:18	2.61
QLDB_SSD (*)	00:49:13	00:13:23	3.68	01:30:28	00:18:06	5
VLDB_SSD (*)	01:36:22	00:24:36	3.92	01:55:31	00:29:12	3.96

Sweep

Database	FB Active
SDB_RAID10	01:44:30
MDB_RAID10	04:04:50
QLDB_RAID10	08:28:07
VLDB_RAID10 (*)	07:04:29
SDB_SSD	00:44:46
MDB_SSD	02:11:18
QLDB_SSD	04:15:16
VLDB_SSD (*)	03:19:41
4K page size except * - 8K	

What's up with sweep?

Before sweep:

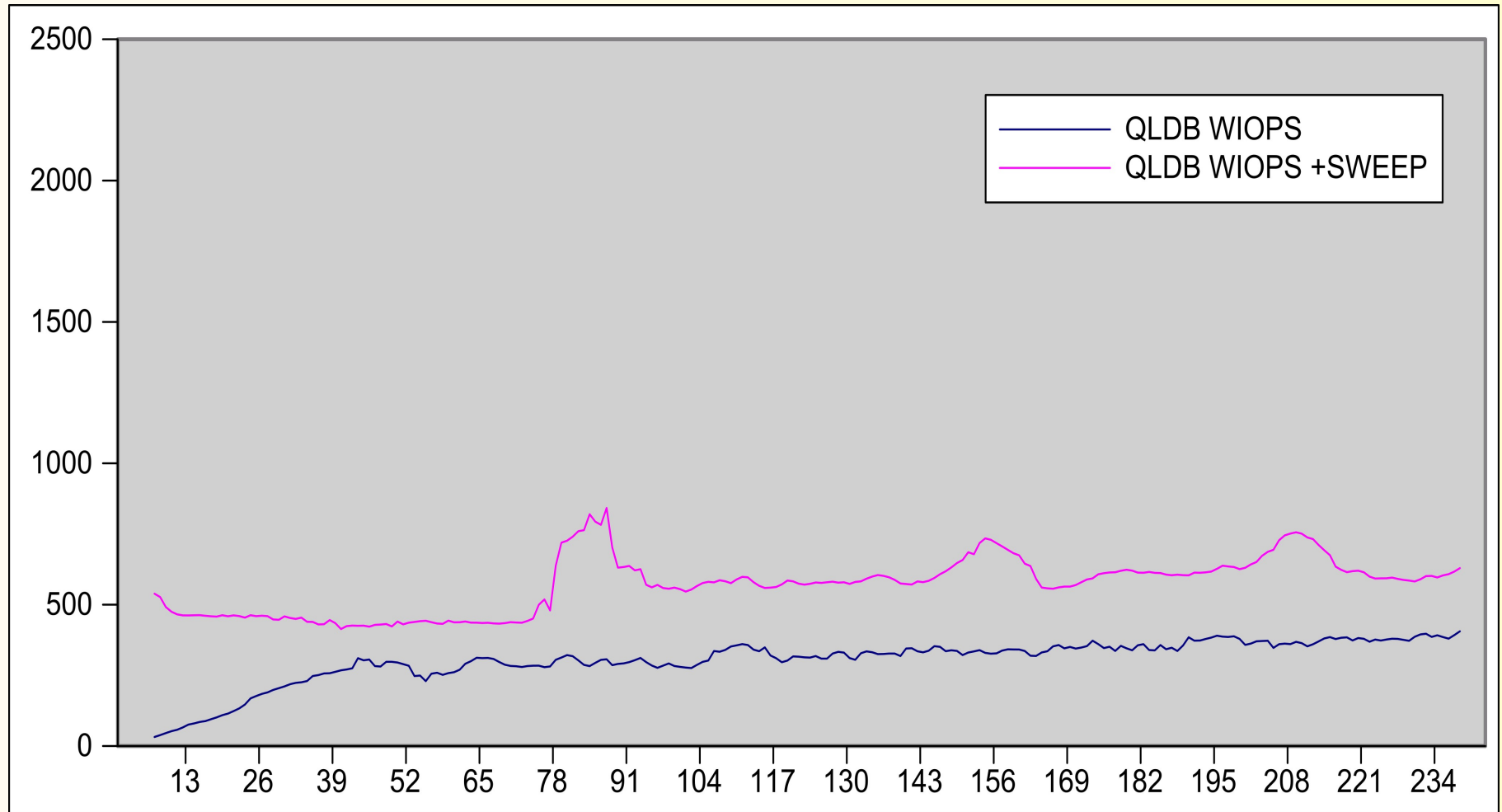
Oldest transaction	366,678
Oldest active	366,679

After sweep (8 hours later) :

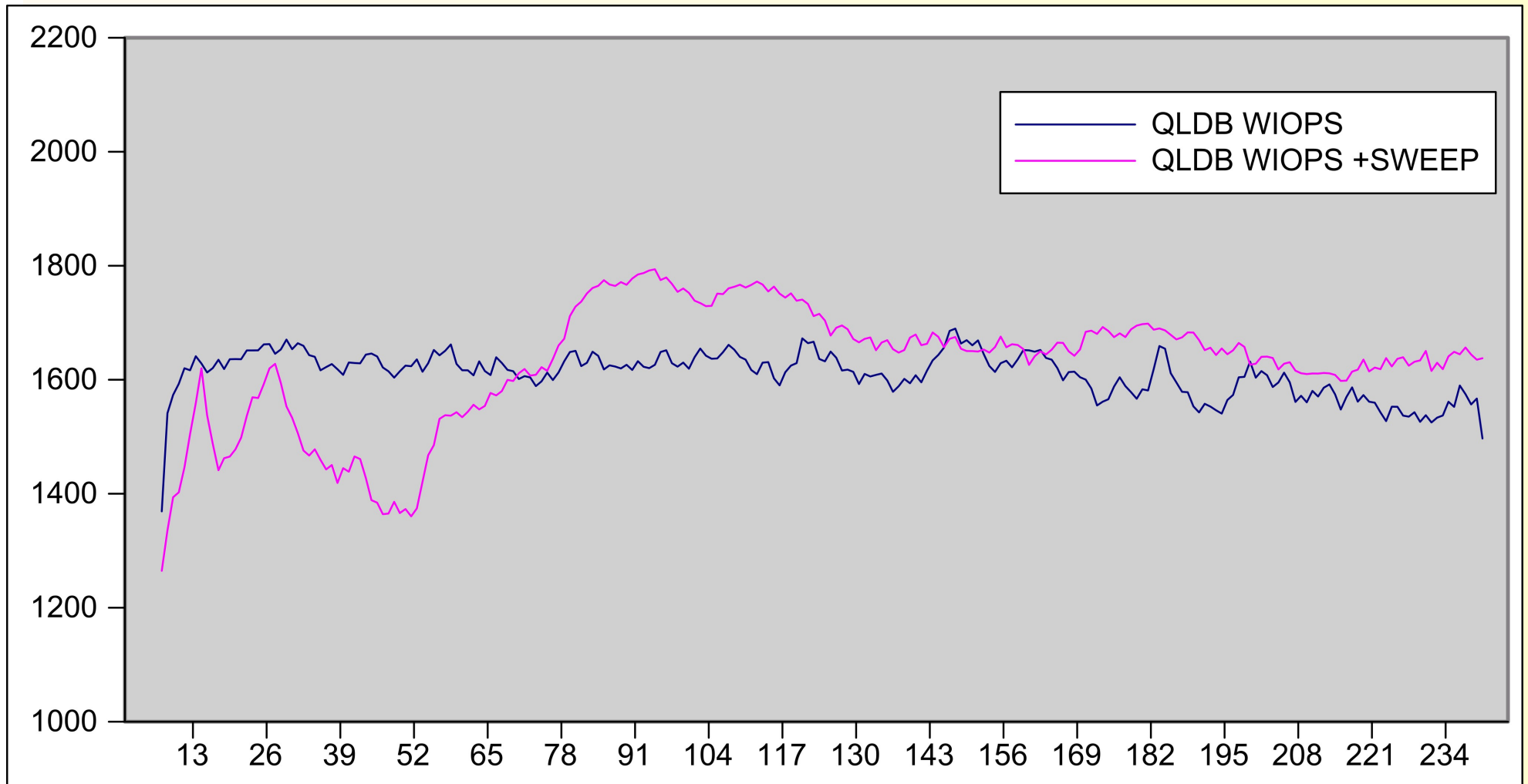
Oldest transaction	1,532,631
Oldest active	1,532,632

Does it matter?

Sweep and WIOPS (RAID 10)



Sweep, WIOPS and SSD



Things I haven't had time to look at

- Page size comparisons
- Gbak backup cost
- Free space on disc (SSD)

Questions?