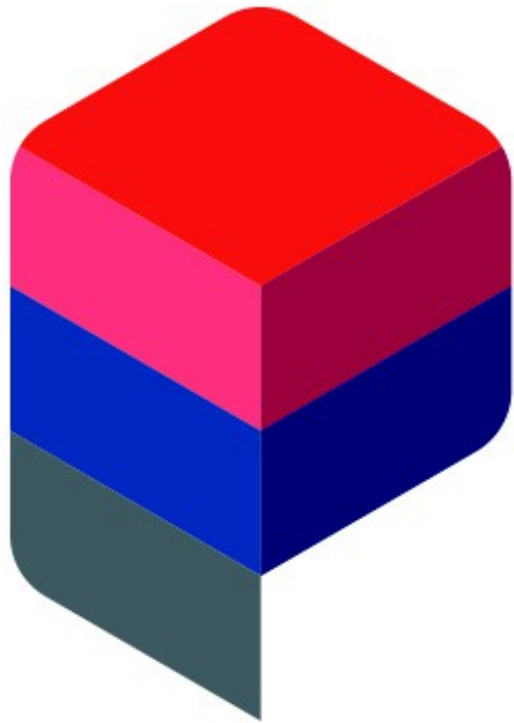




**performing
databases**

Oracle Core - Database I/O

Martin Klier

Performing Databases GmbH
Mitterteich



#FiveWordTechHorrors „Storage comes from another department“



performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

Today ...

- How the RDBMS works (Overview)
- Using storage
- Automatic Storage Management (Overview)
- Different IO-Types
- Storage Technology // Sizing Basics
- Virtualization

Oracle DBMS Basics



#FiveWordTechHorrors „John knows everything about databases.“

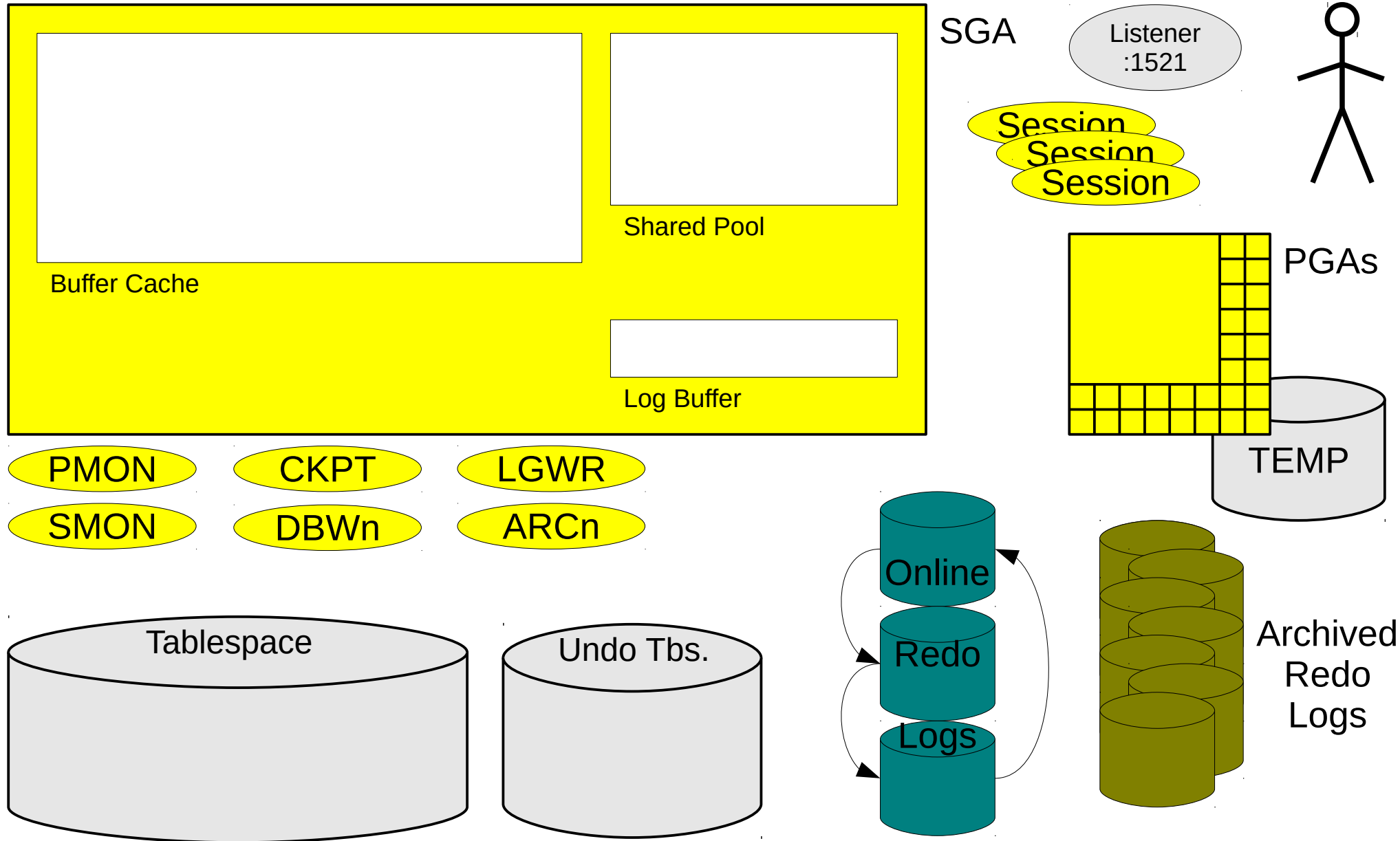


performing
databases

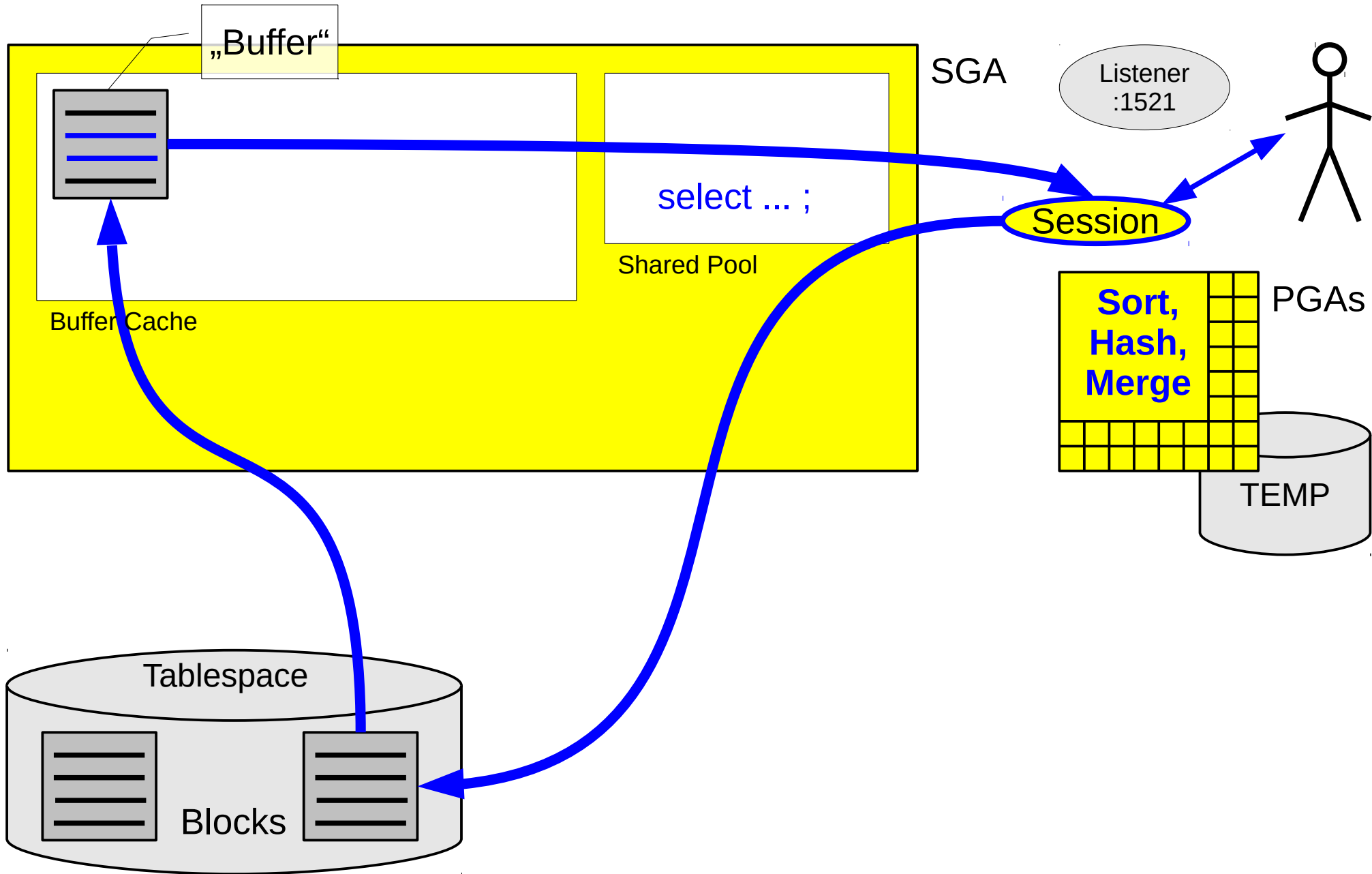
Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

Oracle Architecture

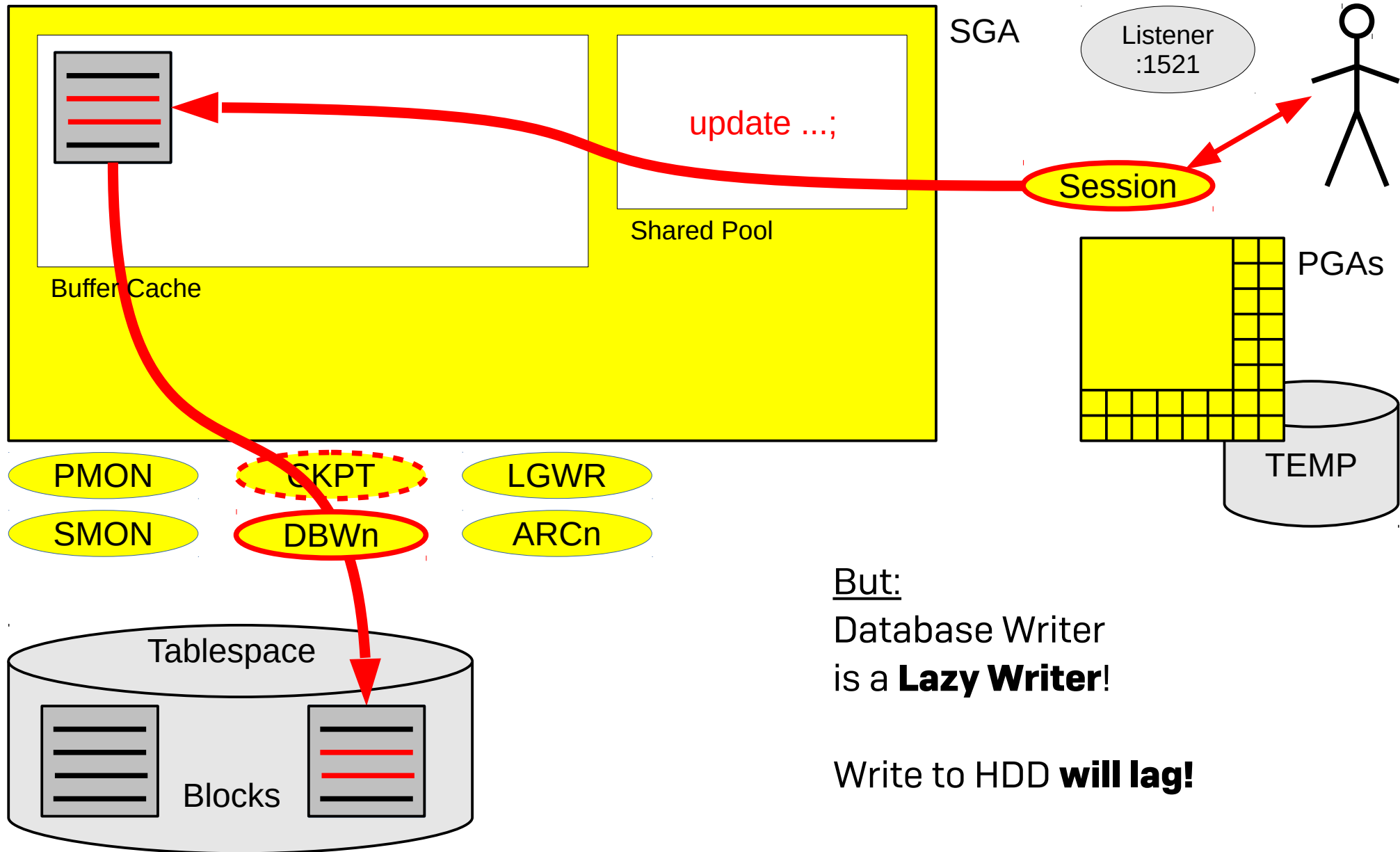
(much simplified)



Reading



Writing // DBWn

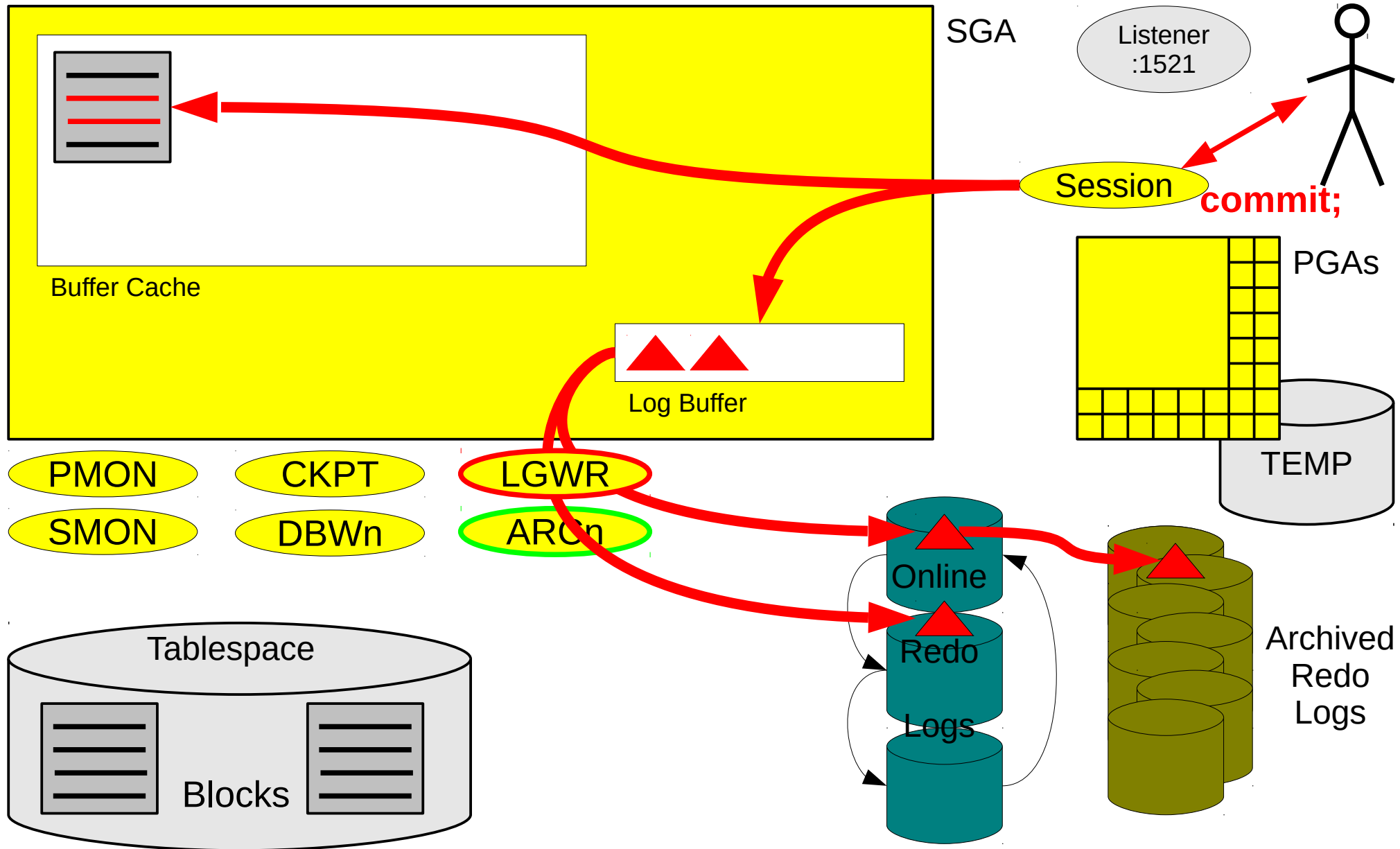


But:
Database Writer
is a **Lazy Writer**!

Write to HDD **will lag**!

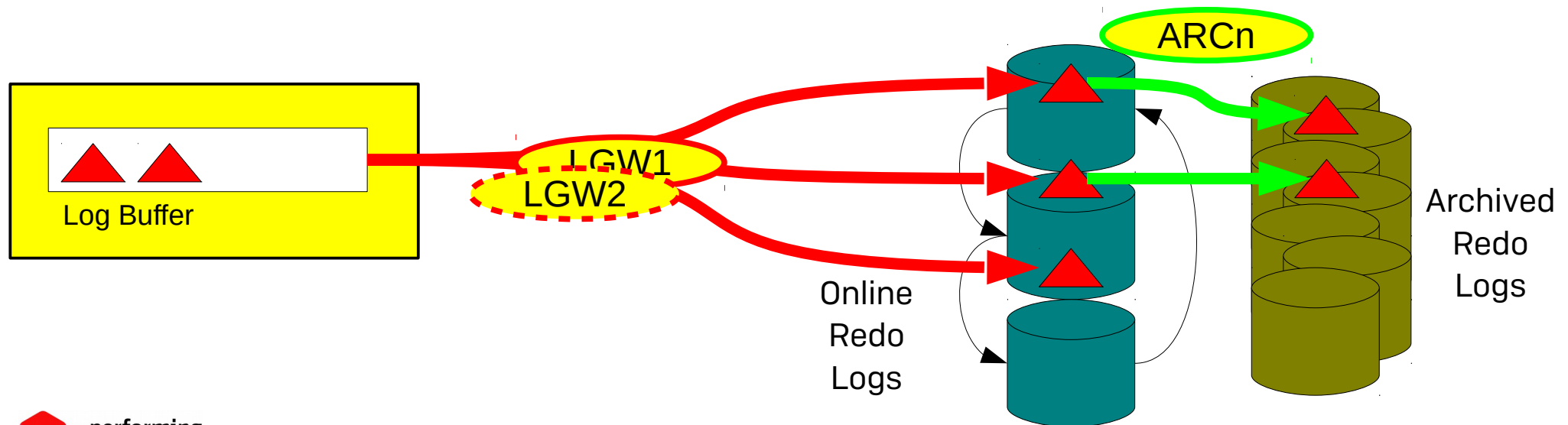
Changing Data

[Redo part]



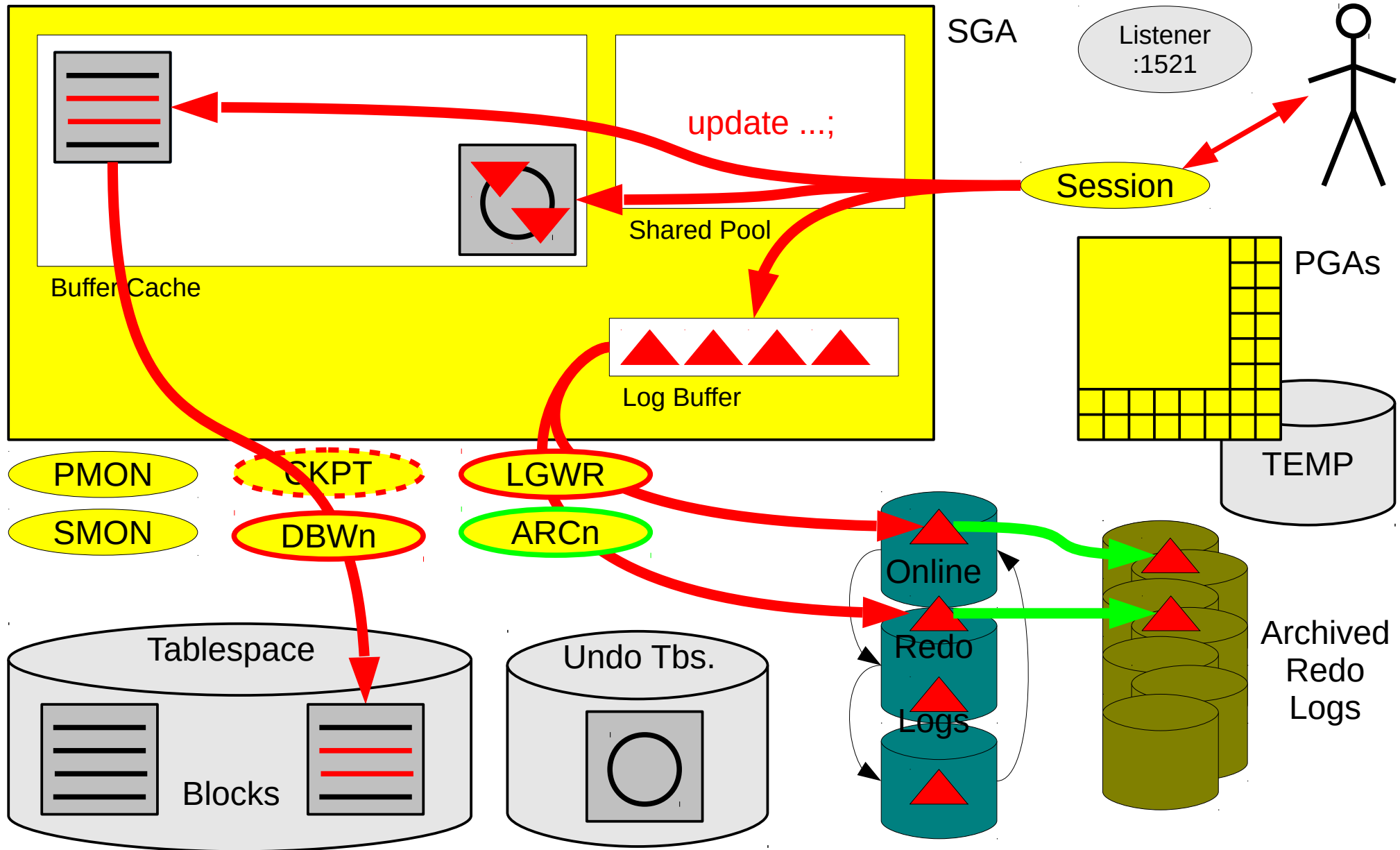
Writing Redo Logs

- At Commit
- Log Buffer Threshold
- latest after 3 sec
- [private strand flushed]



Changing Data

[Redo + Undo]



Some IO-Categories



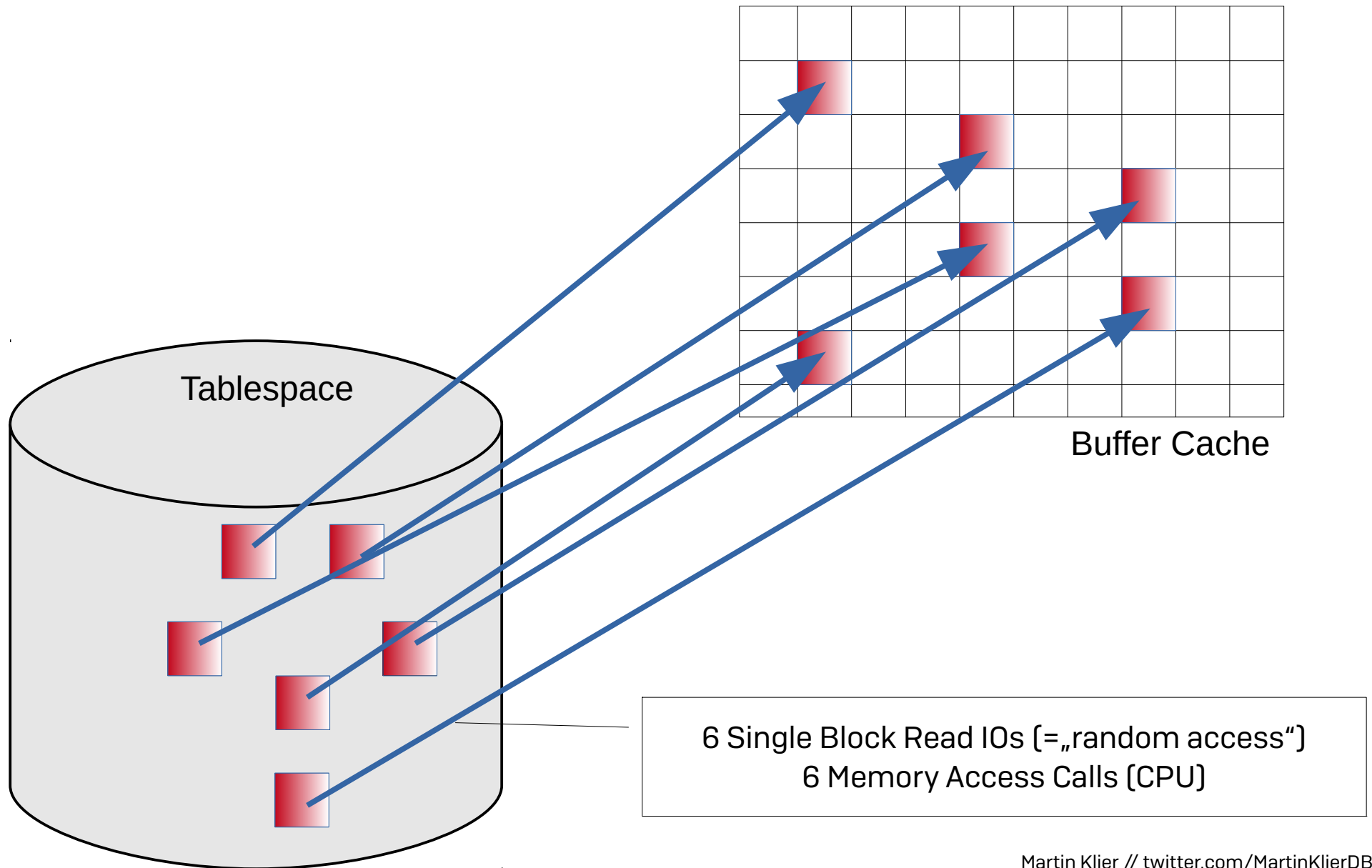
#FiveWordTechHorrors „Storage is fast. What's IOPS?“



performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

Wait: DB file sequential read



Wait: DB file sequential read

- Used for Single Block Access (Index Block)
- Also used for Full Access, but in very low quantities
- 1 I/O call reads **1 block** from TBS
- Writes block content into the optimal buffer in the buffer cache

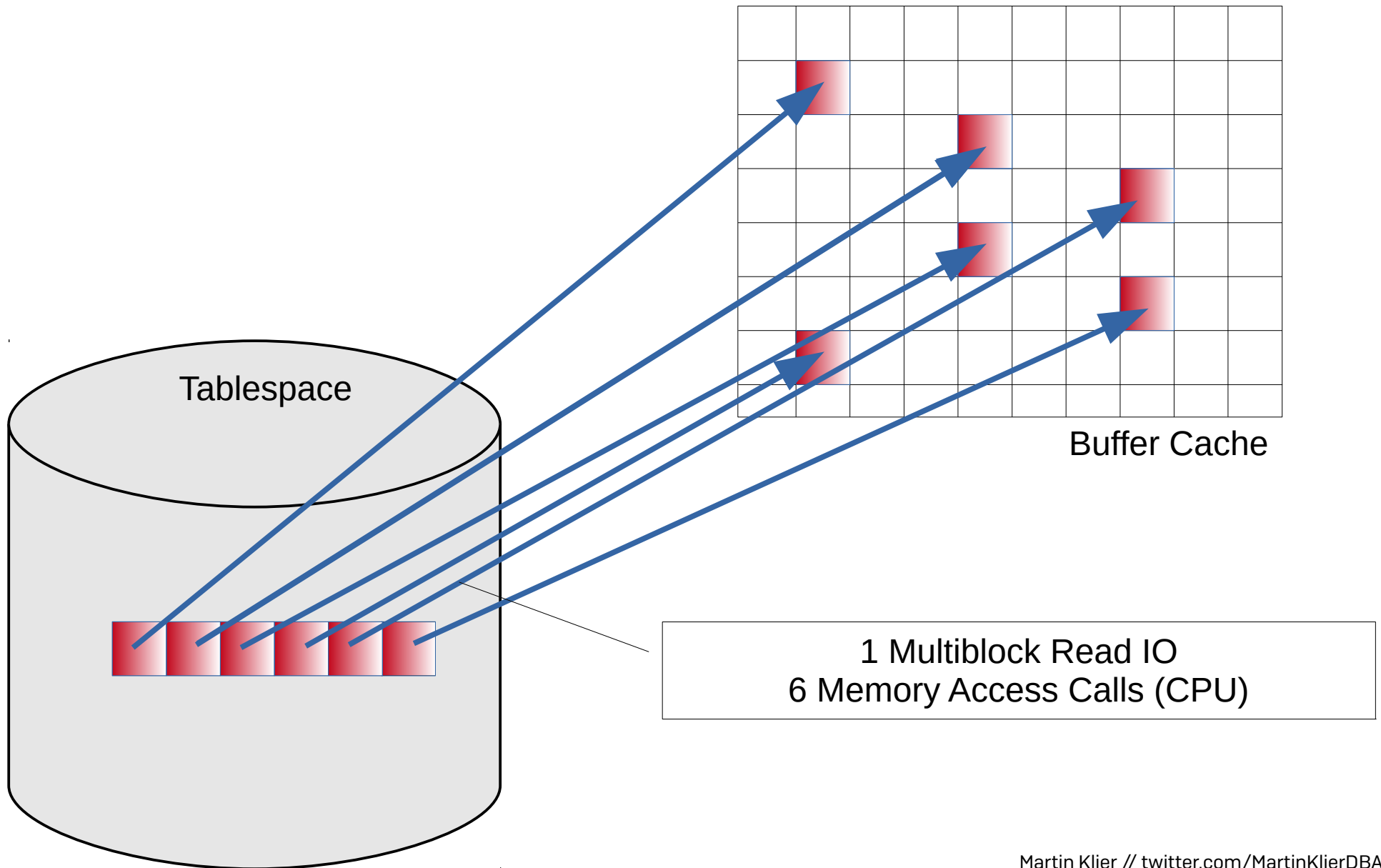
Quickie: TABLE ACCESS FULL

Special Thanks to Roger MacNicol :)

	1	SORT AGGREGATE
*	2	HASH JOIN
*	3	HASH JOIN
*	4	TABLE ACCESS FULL
	5	TABLE ACCESS FULL
	6	TABLE ACCESS FULL

- Two Engines
 - Row-at-a-time driver
 - By row SCN
 - Not pushing down predicates
 - No aggregation
 - Block at a time driver on RDBMS (Turboscan)
 - Requesting blocks from TX layer
 - Can process predicates
 - ...
- Can be „into Buffer Cache“ or „Direct Read“

Wait: DB file scattered read

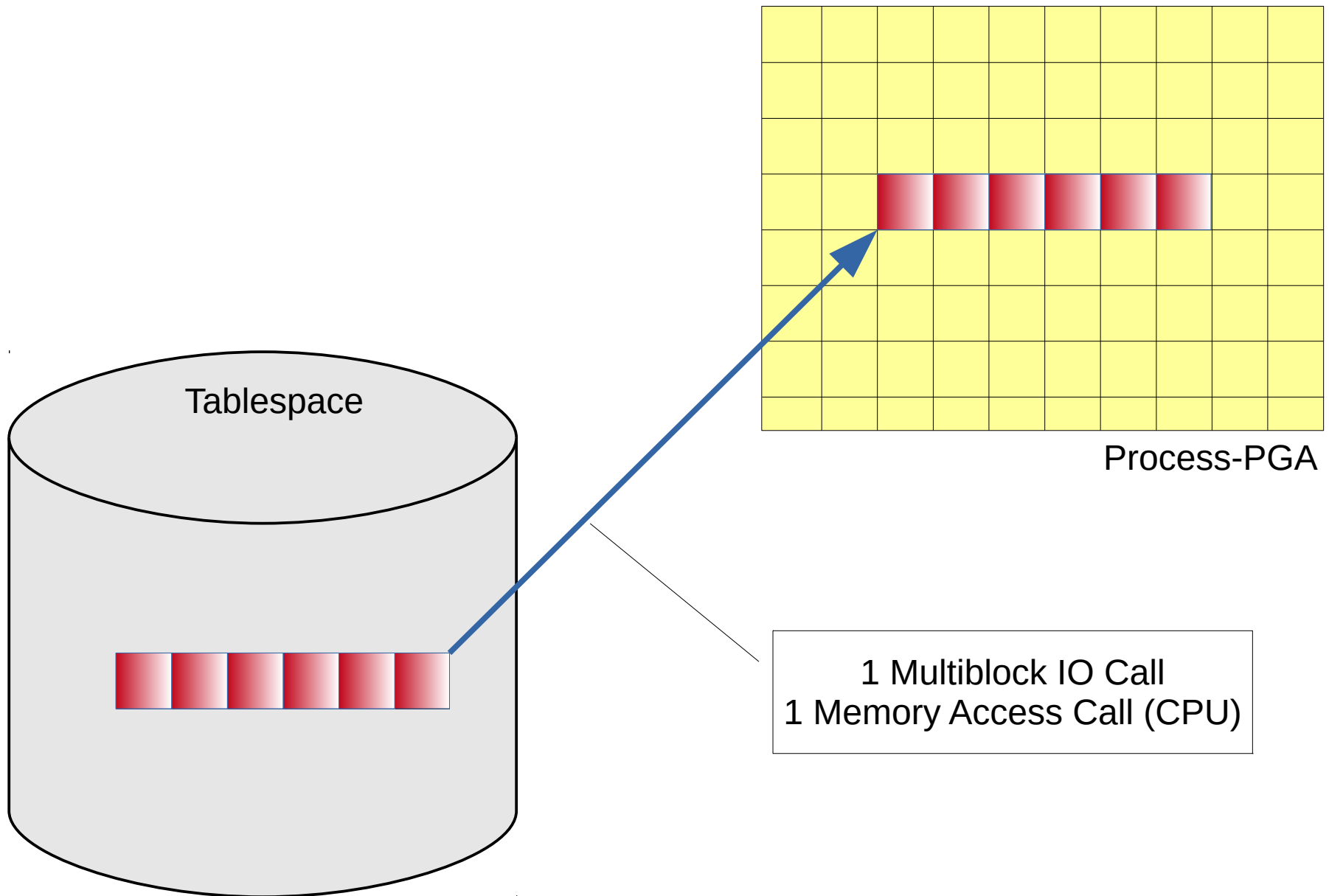


Wait: DB file scattered read

- Used for FULL SCANS (Table, Index ...)
- 1 I/O call reads **n contiguous blocks** from TBS
- $\max(n) = \text{DB_FILE_MULTIBLOCK_READ_COUNT}$
- Writes in a „scattered“ way into the buffer cache

Purpose: Saving IO time

Wait: Direct Path Read



Wait: Direct Path Read

- Used for massive reads (Parallel Exec, Export...)
- 1 I/O call reads **n contiguous blocks** from TBS
- $\max(n) = \text{DB_FILE_MULTIBLOCK_READ_COUNT}$
- Writes contiguously into the PGA of the calling process

Purpose: Saving IO- and CPU time (less memory operations)

Oracle ASM



#FiveWordTechHorrors „No ASM diskgroups are online“



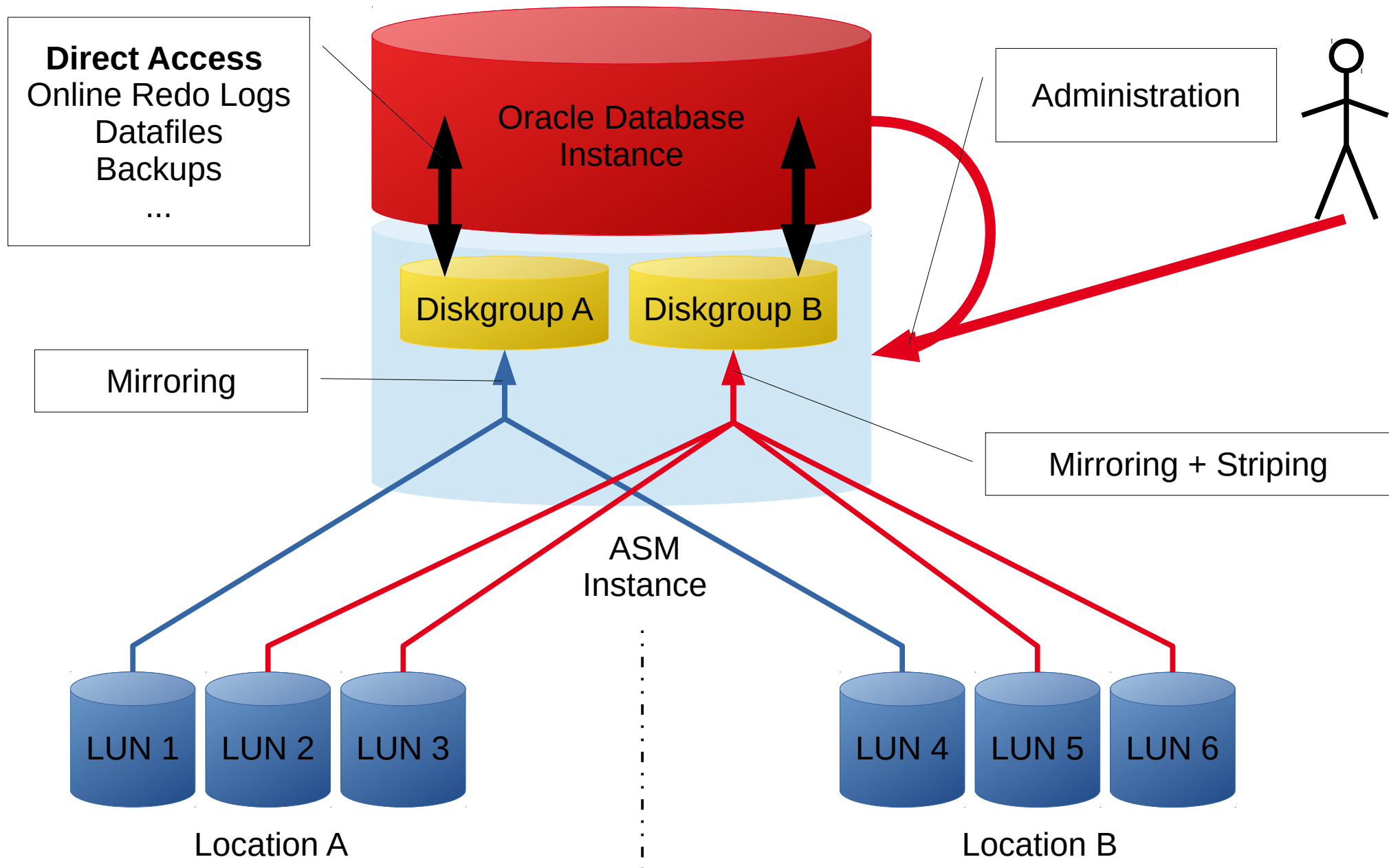
performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

ASM

- „Automatic Storage Management“
[Automagic?]
- Logical Volume Manager
- Bypassing file access layer („File System“)
[like a Raw Device]
- Multi-Master possible
- Human interface / ops by SQL or CLI

Automagic SM



Storage Sizing Basics



#FiveWordTechHorrors „Database fits into cache anyway“



performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

#StorageTechHorrors

- The database fits into the cache anyway!
- We avoid IO anyway if possible!
- Mirroring? Why? We have hourly Archivelog Backups!
- Ten years ago, this have been Megabytes!
- I've been doing this for 20 years.
Our storage IS fast enough for!

5-Fingers IOPS calculator

1. Data Block - one
2. Undo Block - two
3. Online Redo Log - three
4. Archived Redo Log - four
5. Control File - five

Wrong -

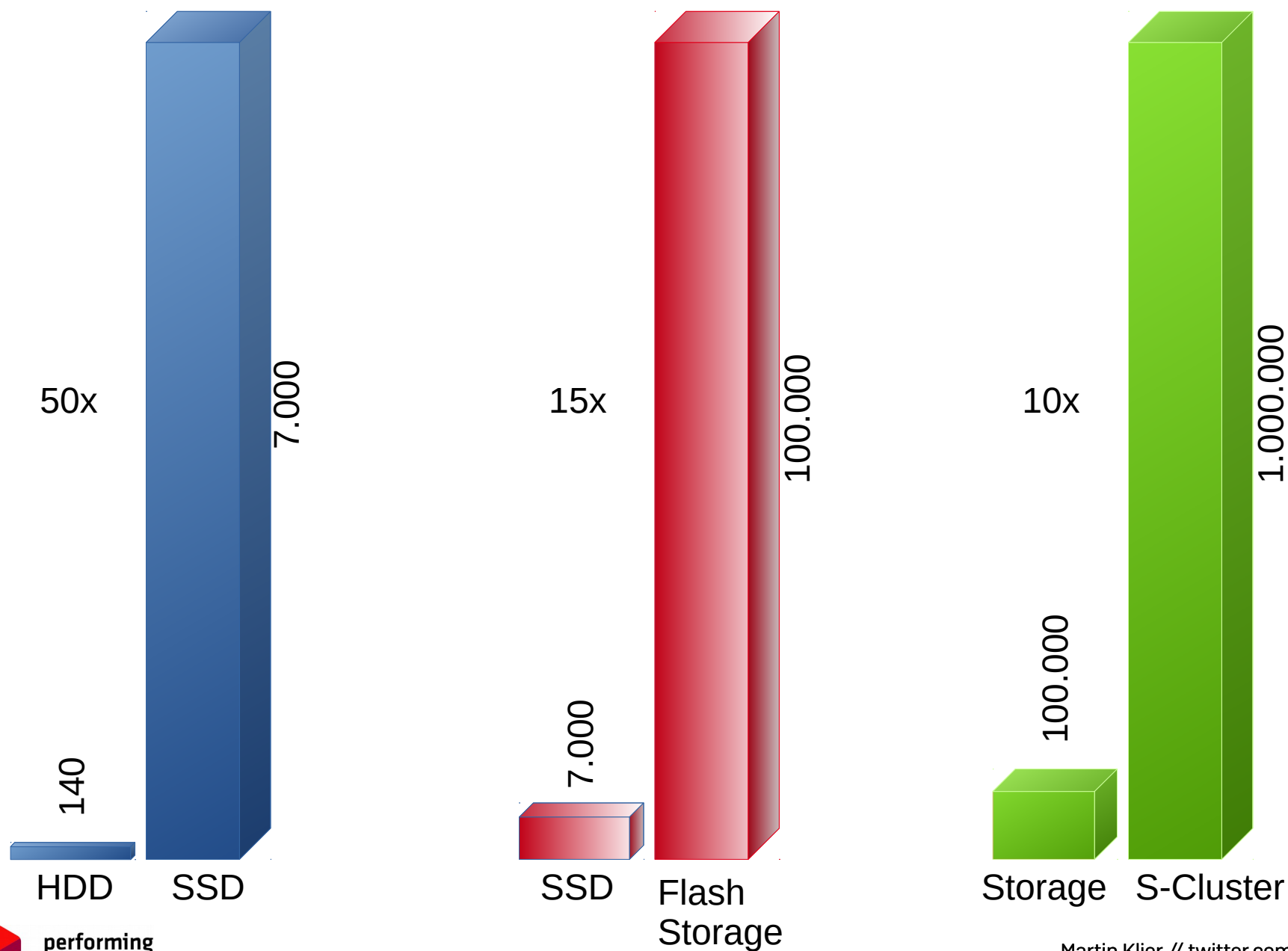
- but OK -

+1 for any further block

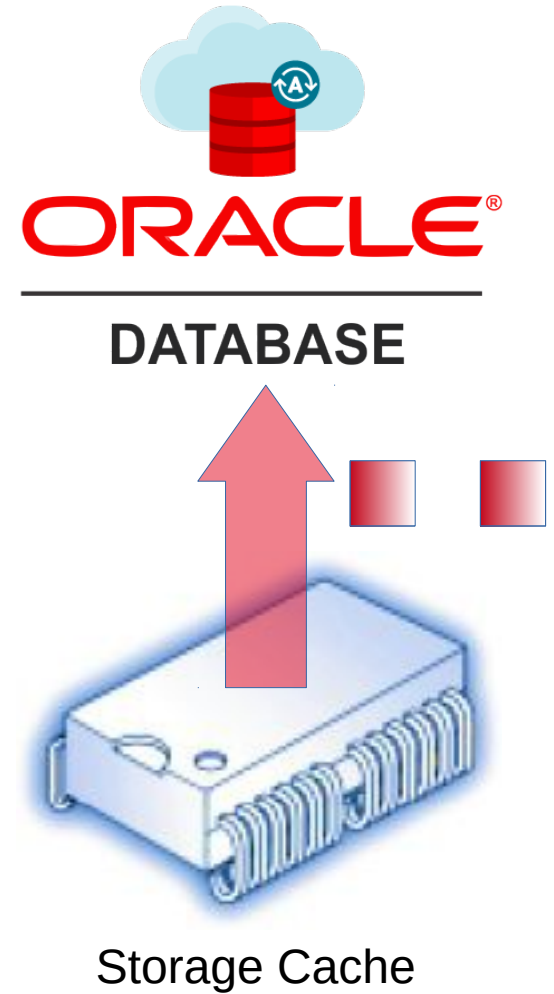
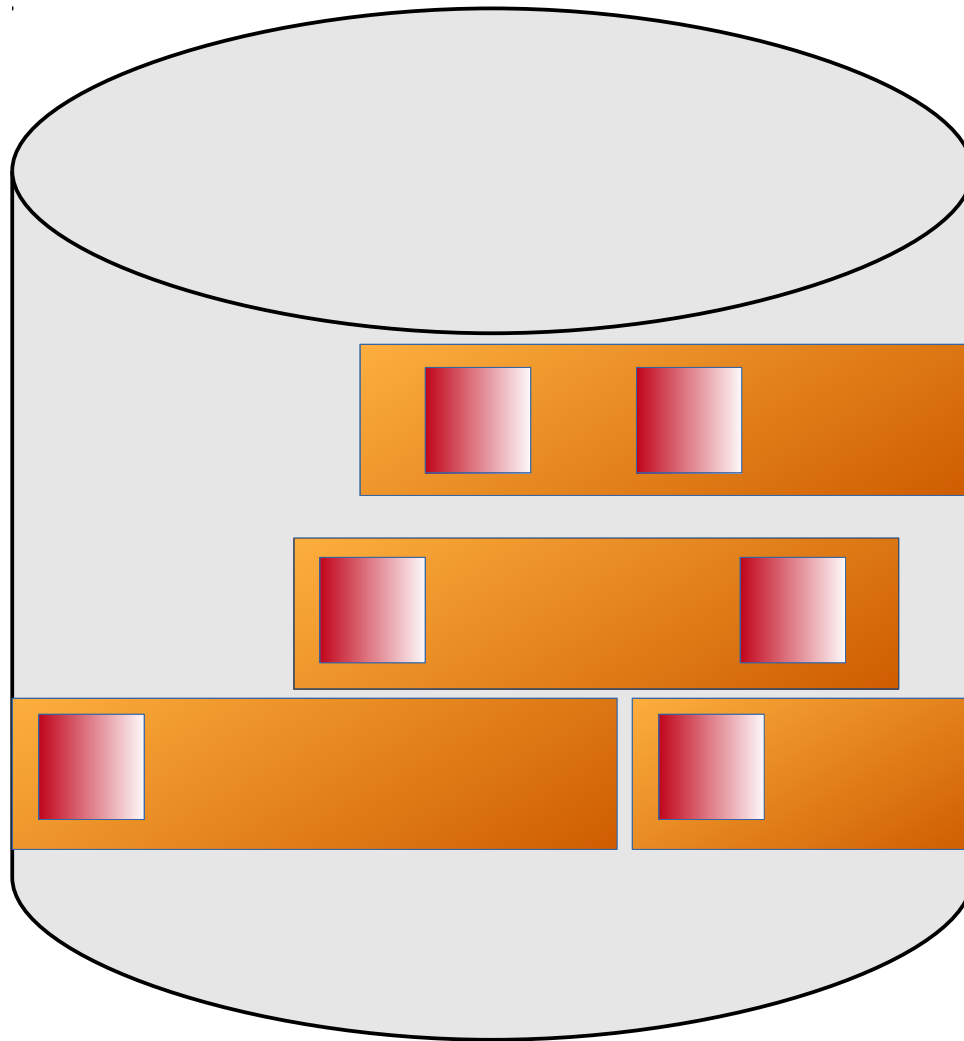
+1 for Flashback Log



IOPS compared

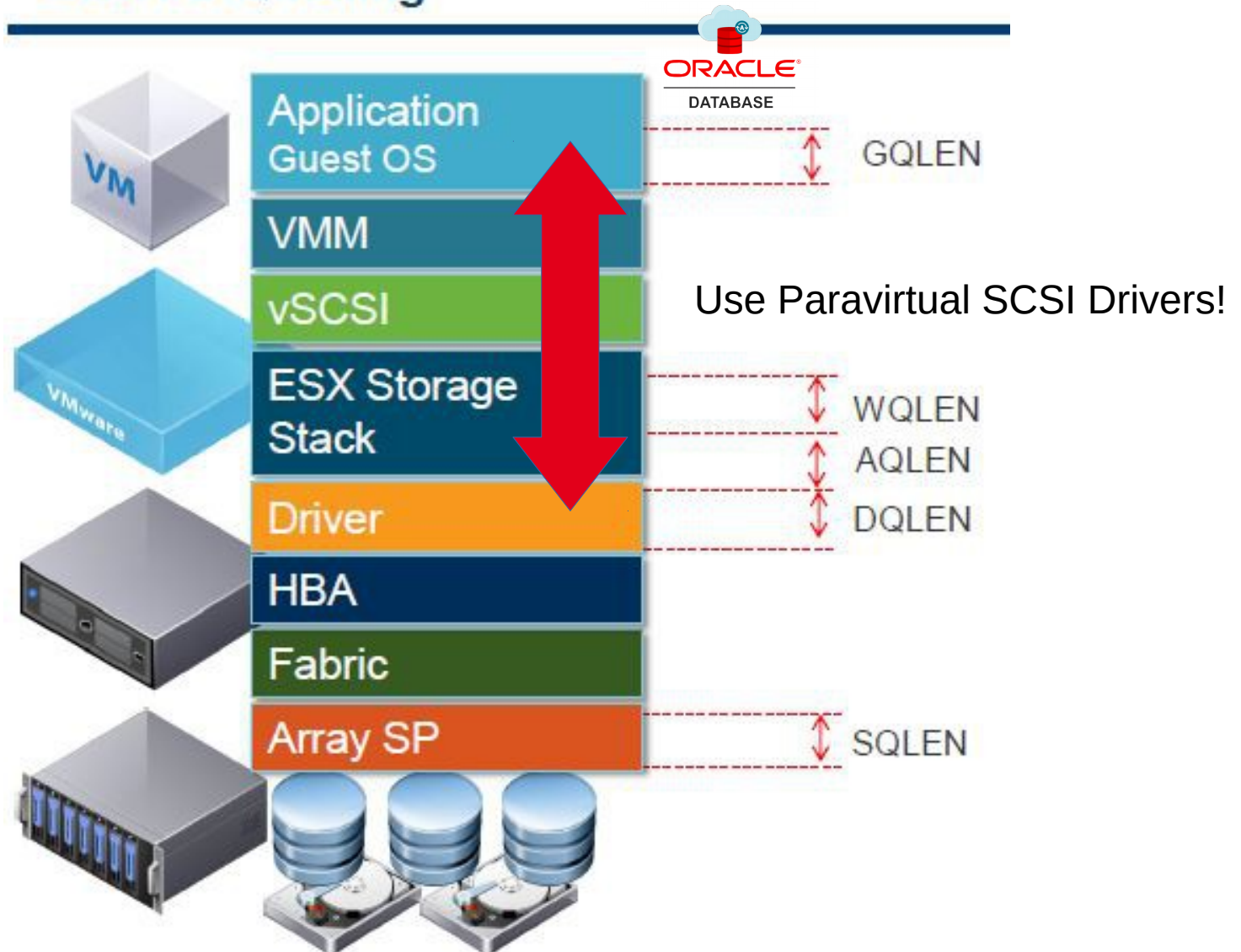


Real-world Storage Read (Ahead)



VMware IO

Disk I/O Queuing



Today ...

- How the RDBMS works (Overview)
- Using storage
- Automatic Storage Management (Overview)
- Different IO-Types
- Storage Technology // Sizing Basics
- Virtualization

#NoHorror



Speaker

- Martin Klier
- Solution Architect and Database Expert
- My focus:
 - Performance + Tuning
 - highly available Systems
 - Cluster and Replication
- Linux since 1997
- Oracle Database since 2003



ORACLE®
ACE Director



performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com

Performing Databases

- Experts for Database Technology
 - Concepts and Project Competence
 - Architektur- und System planning
 - Licensing
 - Implementation and Troubleshooting
- Contact
 - Performing Databases GmbH
Wiesauer Straße 27
95666 Mitterteich // Germany
 - Web: <http://www.performing-databases.com>
 - Twitter: @PerformingDB



Join PERFORMING DATABASES!

What is
efficiency
for you?

Verstärke unser Team!

DB CONSULTANT

Technische Anforderungen

- mindestens 3 Jahre Erfahrung als DBA oder DB-Entwickler
- Oracle xor SQL Server xor PostgreSQL
- Linux // Windows // Mac OS // AIX // Solaris
- bash // PL/SQL // SQL // RMAN // APEX

Willkommen in
Deiner neuen
Arbeitswelt:

Details at

<http://www.performing-databases.com>



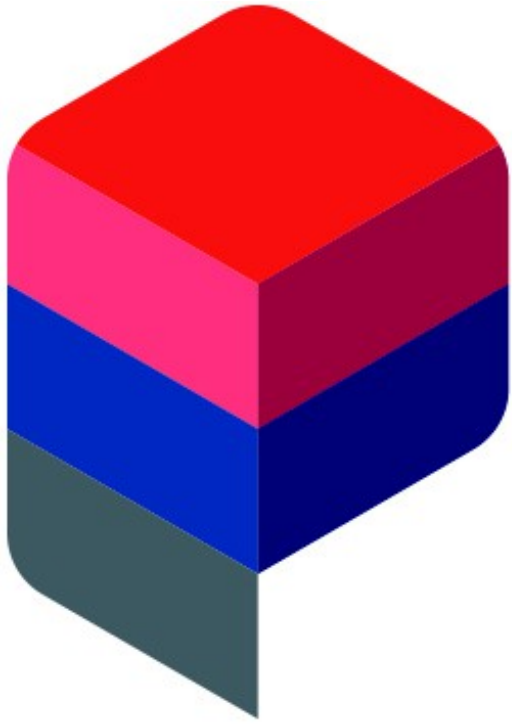
performing
databases

Martin Klier // twitter.com/MartinKlierDBA
www.performing-databases.com



Download Presentation

<http://www.performing-databases.com>



**performing
databases**