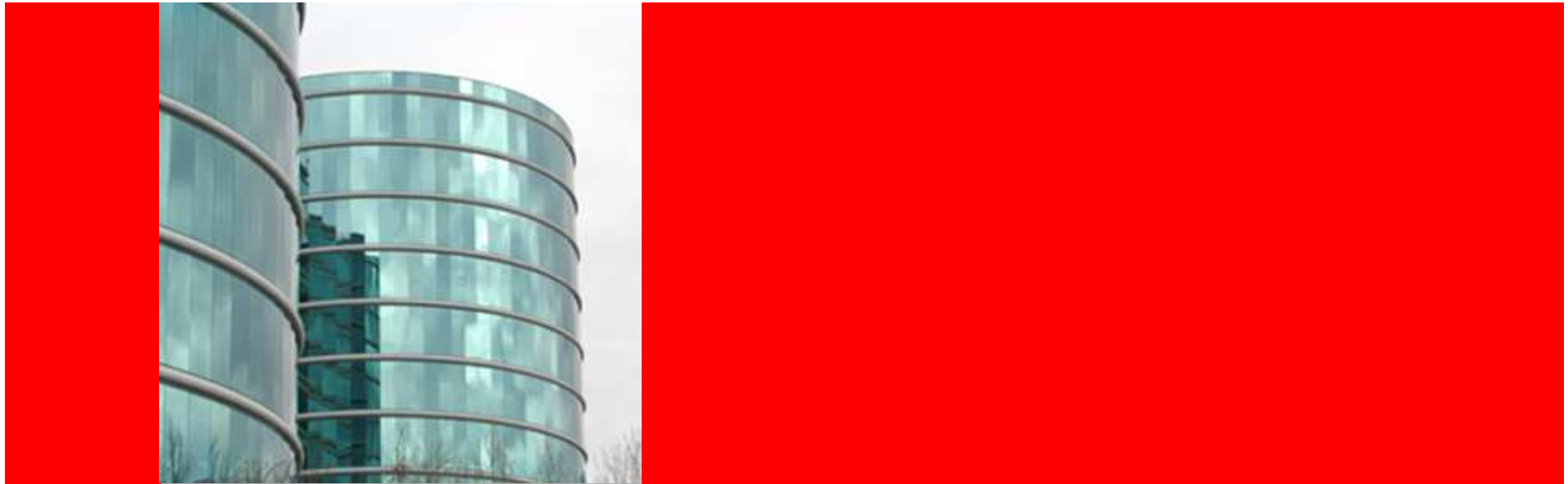




ORACLE®



**ORACLE®**



## **Data Caches and Buffers Review**

Norman Lastovica  
Oracle OpenVMS Development Team

13 October 2009

# Agenda

- A walk down memory lane
- Disks
- Caches
- Databases



# Remember When...



**Remember When...**



ORACLE

# Remember When...



# Remember When...





# Remember When...







**Where is my data?**

**Why is it there?**

**How long will it be there?**



# The Big Picture

- IO is slow
- Memory is fast





# Modern Disk Device

- On-board read cache
- Write-back cache
  - Great for test & development systems with local SCSI disks
  - Significant performance gain for some applications



## Enable Local SCSI WBC

```
$ mount/system $1$dka800: foobar
$ mcr sys$etc:scsi_mode -
    -devname dka800 -
    -devtype dg036a8b53 -
    -page 8 -
    -offset 0e 14 -
    -mount -
    -noconfirm
```



# SSD

- 256 GB for Laptop
  - \$650
  - Sequential 220MB/s
- PCIe for HPC in 2U
  - 504GB
  - 3 microseconds latency
  - Over 3 million random IOPS
  - Throughput up to 1,700MB/s



ORACLE





# Controller

- Read cache
- Write-back cache
- Data protection
  - RAID
  - Batteries



Standalone



Cluster



Hot swap



RAID 0



RAID 1



RAID 5



RAID 0+1

## Storage Systems Explained





# Operating system

- XFC
  - No-cache IO
- RMS
  - Global buffers
  - File header and directory information
  - Installed images / known files
  - RAH & WBH
    - RMS\_SEQFILE\_WBH
- XQP





# Generally

- The further the data is from the database engine...
- ...The slower the access may be



# Local Buffers

- Per-process
- Non-shared
- Using a default buffer count from the disco-era?





# Global Buffers

- SHARED MEMORY IS...
  - PROCESS
  - PROCESS RESIDENT
  - SYSTEM
  - LARGE MEMORY IS ENABLED
- PAGE TRANSFER
  - VIA DISK / MEMORY
- Increases locking operations vs. local buffers



# Fast Commit

- Avoid writing pages to database at commit
- Write & release/demote locks at blocking AST
- Write checkpoint



# Row Cache

- Single-computer
- SHARED MEMORY IS
  - PROCESS
  - PROCESS RESIDENT
- WRITE-BACK
  - Less likely without snapshots in cache
  - RCS tends to do much of the writing with snapshots in cache



# Tradeoffs

- Eventually we'd like data to get to more permanent storage
- Recovery time
- Flushing / Checkpointing




# Summary

- Memory tends to be faster than mechanical disks
  - Modern “disks” are becoming “memory”
- Don't be afraid of memory!
  - 99 Luftballons or 99 Mega Bytes – A long time ago



**For More Information**

[search.oracle.com](https://search.oracle.com)



**or**

[\*\*oracle.com/rdb\*\*](https://oracle.com/rdb)



**ORACLE IS THE INFORMATION COMPANY**