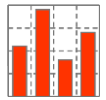
BENCHWARE

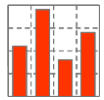


Oracle Performance
on
M5000 with F20 Flash Cache

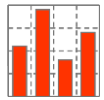
Benchmark Report

September 2011

Contents



Benchware Ltd

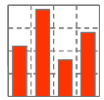


Services and Products

Strong foundation in core technologies like
Oracle database system, server and storage systems

- System Architecture, Component Evaluation, Reviews
- Performance Analysis & Optimization
- Benchmarking
- Database Engineering

Benchware Ltd

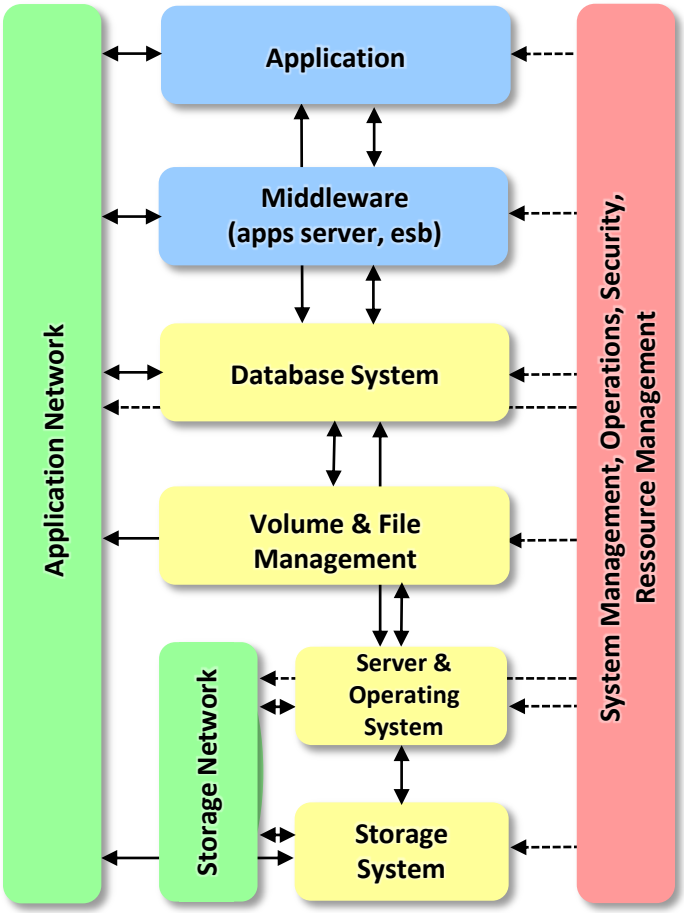


Value proposition

- Vendor-independent company
  - Benchware is completely committed to customers' interests
- Holistic approach in designing, tuning and benchmarking Oracle systems
- Long experience track record
  - Responsible for system architecture of largest DWH and OLTP systems, mainly telecom and finance industry
  - Oracle since 1984 (Oracle Version 3)
  - Performance tuning and benchmarking since 1993 (Oracle Version 7)

Benchware Ltd

Complex architecture of Oracle platforms needs benchmarking



Performance of complex technology stack is NOT predictable – unless running a benchmark

Application Network (IP-based)

Bandwidth, latency during remote database mirroring (sync, async) due to switches and sql*net and tcp/ip stack (frame size, …).

Oracle Database

Different versions, patches and options, about hundred configuration parameters.

Storage Network (FC-, IB- or IP-based)

Bandwidth, latency during remote storage mirroring (sync, async) due to switches, hubs and distance.

Volume & File Management

Different volume managers (VxVM, ASM) and file Systems (UFS, VxFS, ext3, JFS, ZFS, raw devices), different I/O methods (async, direct), a lot of config parameters (#LUNS, queue depth, max i/o unit), software striping and/or mirroring, multipathing.

Server & Operating System

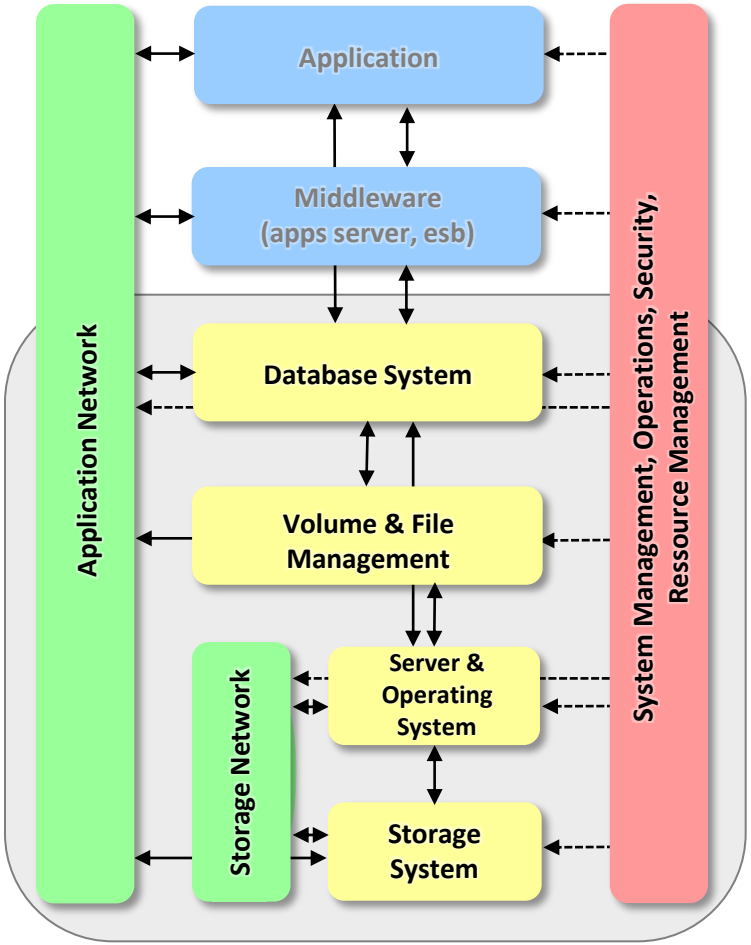
Different server Systems, processors and CPU architectures, (x86, IA-64, UltraSparc, SPARC64, Power), #cores, multithreading, main memory, bus architecture. Different operating Systems and patches, over hundred configuration parameters, virtualization of resources.

Storage System

Different storage Systems, storage tiers and storage technology: spindle count and speed, RAID management, cache management, server interface technology, storage system options like remote copy, hardware striping and/or mirroring, virtualization of resources.

Benchware Ltd

Benchware Performance Suite



Object of measurement

- Benchware Performance Suite
  - Benchware Monitor
  - Benchware Loader
- Performance measurement at the interface between application and technology stack
- Key Performance Metrics can be used for SLA between IT operation and business
- Benchware uses Oracle Database stack to generate all kind of loads for cpu, server, storage and database

Benchware Ltd

Library of Oracle benchmark tests - implemented in PL/SQL, Java and SQL

CPU Performance CPU-bound Oracle operations All operations in Level 1, 2, 3 CPU cache	OLTP systems	DWH systems	Efficiency	Metrics	Unit
• pl/sql basic operations	★★	★★	multithreading virtualization	speed throughput	[s] [ops]
• pl/sql algorithms fibonacci, prime numbers	★★★	★★			

Server Performance Server-bound Oracle operations All operations in RAM - no I/O operations	OLTP systems	DWH systems	Efficiency	Metrics	Unit
• in-memory SQL	★★★	★★	scalability cc-numa virtualization	speed throughput	[µs] [s] [bps] [tps] [rps]
• pl/sql algorithms quicksort	★★	★			

[s]	seconds	[bps]	buffers per second	[MBps]	mega bytes per second	★	less important
[ms]	milli seconds (10^{-3})	[rps]	rows per second	[GBps]	giga bytes per second	★★	important
[µs]	micro seconds (10^{-6})	[tps]	transactions per second	[iops]	i/o operations per second	★★★	very important
[ns]	nano seconds (10^{-9})	[ops]	operations per second	[qpm]	queries per minute		

Benchware Ltd

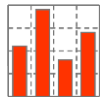
Library of Oracle benchmark tests - implemented in PL/SQL, Java and SQL

Storage Performance I/O-bound Oracle operations	OLTP systems	DWH systems	Efficiency	Metrics	Unit
• sequential I/O 1 MByte, read and write	★★	★★★	RAID tiering striping virtualization	service time throughput	[ms] [MBps] [GBps] [iops]
• random I/O 8 kByte, read and write	★★★	★			

Database Performance Mixed resource usage: CPU, memory, storage	OLTP systems	DWH systems	Efficiency	Metrics	Unit
• data load uncompressed, compressed	★★	★★★	scalability	speed throughput service time	[ms] [s] [rps] [tps] [qpm]
• data scan	★	★★★			
• data aggregation & reports	★★	★★★			
• OLTP transactions insert, select, update	★★★	★			

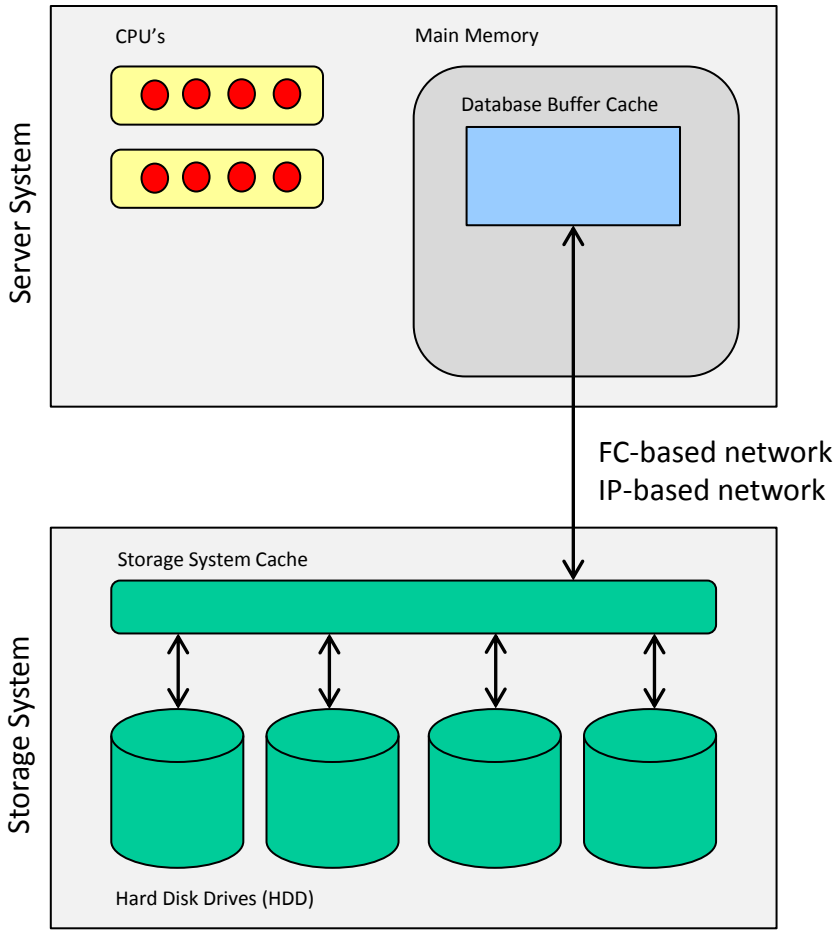
[s]	seconds	[bps]	buffers per second	[MBps]	mega bytes per second	★	less important
[ms]	milli seconds (10^{-3})	[rps]	rows per second	[GBps]	giga bytes per second	★★	important
[µs]	micro seconds (10^{-6})	[tps]	transactions per second	[iops]	i/o operations per second	★★★	very important
[ns]	nano seconds (10^{-9})	[ops]	operations per second	[qpm]	queries per minute		

Contents



Flash Cache Technology

Architecture without Flash Cache



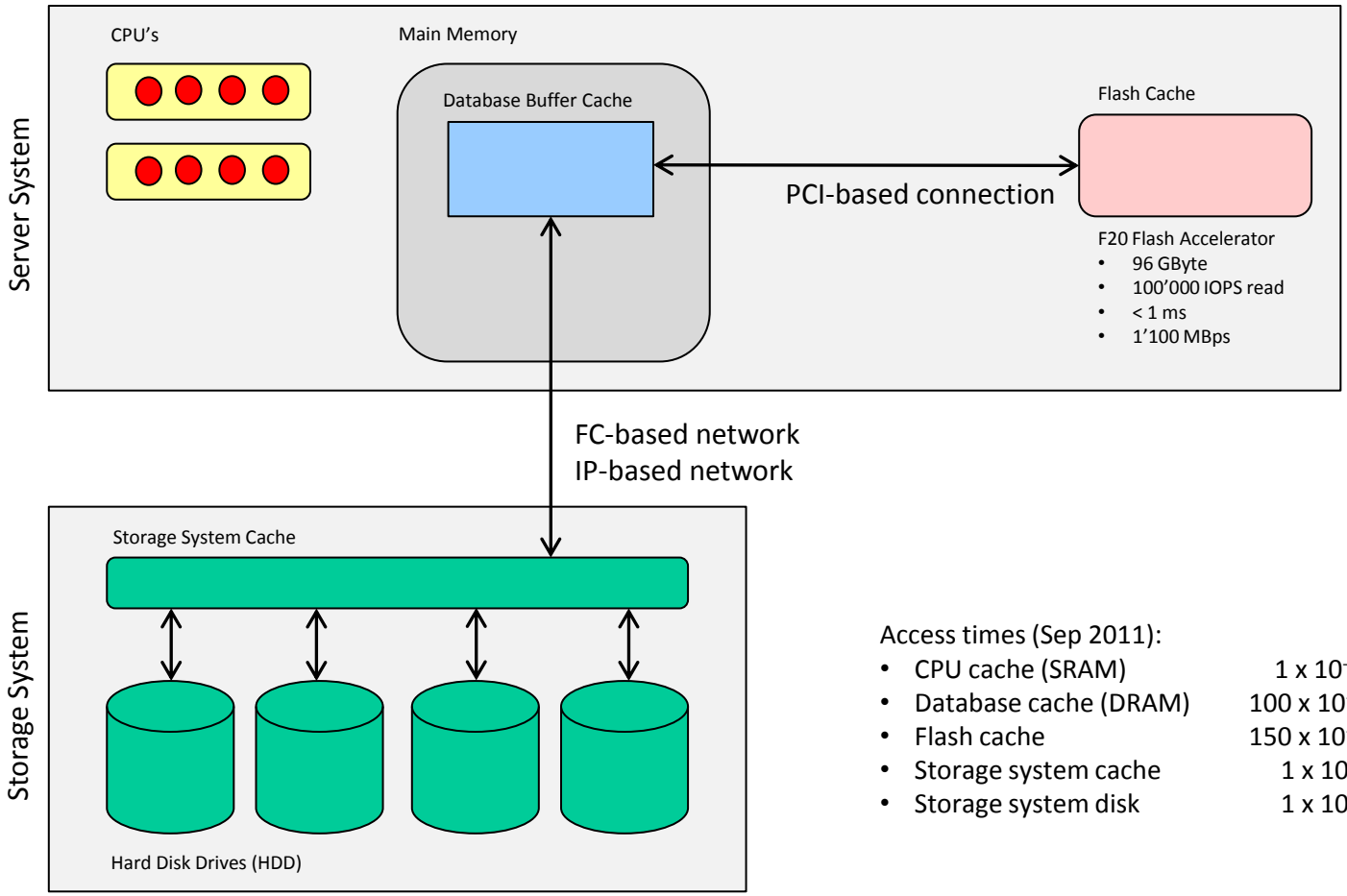
Access times (Sep 2011):

- CPU cache (SRAM) 1×10^{-9} s
- Database cache (DRAM) 100×10^{-9} s

- Storage system cache 1×10^{-3} s
- Storage system disk 1×10^{-2} s

Flash Cache Technology

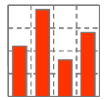
Architecture with F20 Flash Cache



Access times (Sep 2011):

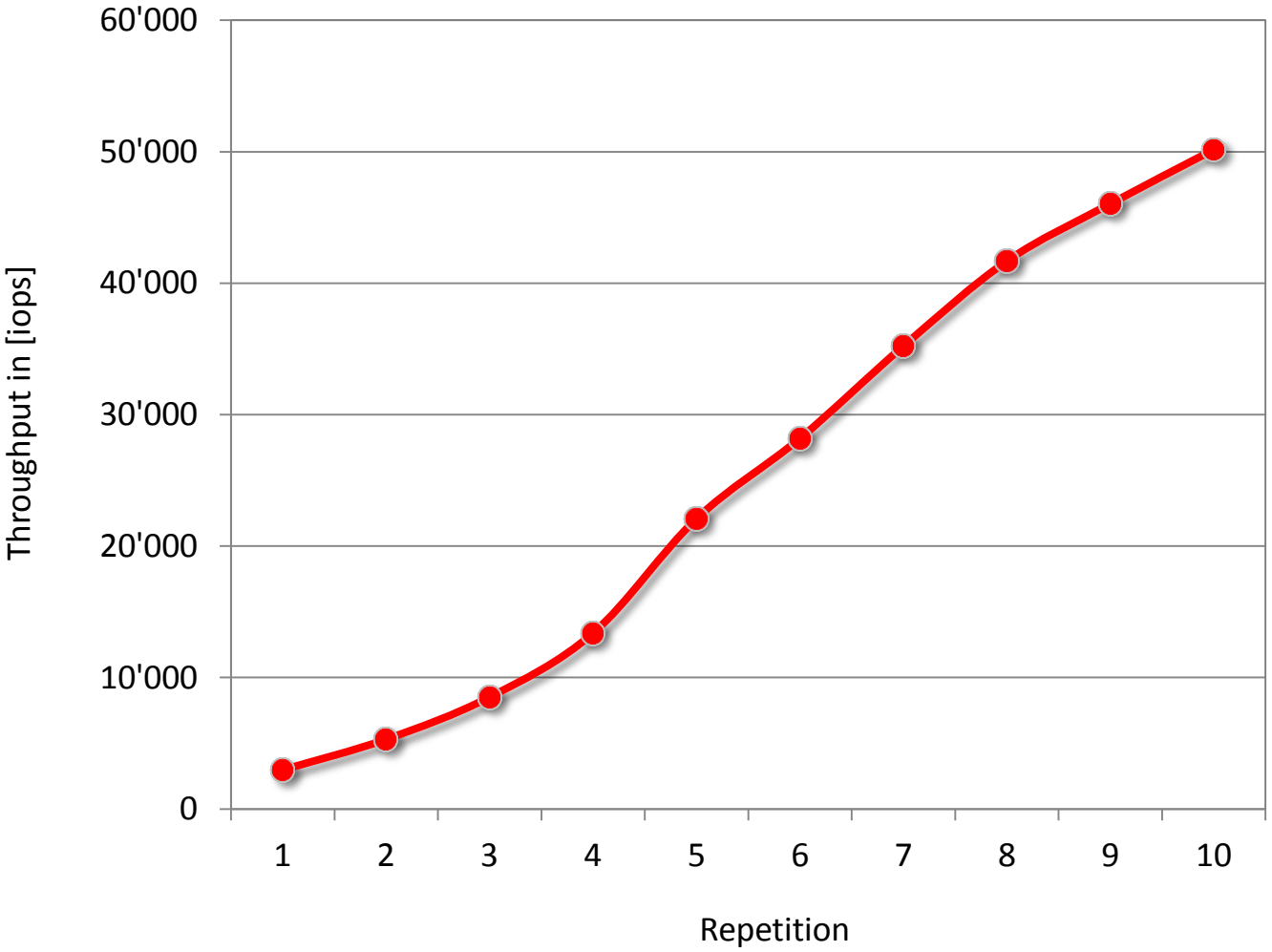
- CPU cache (SRAM) 1×10^{-9} s
- Database cache (DRAM) 100×10^{-9} s
- Flash cache 150×10^{-6} s
- Storage system cache 1×10^{-3} s
- Storage system disk 1×10^{-2} s

Contents



Storage Performance Test

Random read after Oracle instance startup



When repeating the test series several times, we achieved a max value of 50'000 random IOPS.

This is equivalent to
~200 hdd with 15k rpm (250 IOPS) or
~280 hdd with 10k rpm (180 IOPS)

Storage Performance Test

Random read after Oracle instance startup

F20 Flash Cache

Run	Tst	Code	#N	#J	#T	Rows/sec	Ops/sec	CPU [%]	Read [IOPS]	Write [IOPS]	Total [IOPS]	Read [MB/s]	Write [MB/s]	Total [MB/s]	REDO [MB/s]	Time [sec]
7	1	T626	1	64	1	2.91E+03	2.27E+00	9	2965	1962	4928	23	15	39	0	902
	2	T626	1	64	1	5.20E+03	4.07E+00	10	5302	1840	7142	41	14	56	0	503
	3	T626	1	64	1	8.36E+03	6.52E+00	12	8508	2046	10554	66	16	82	0	314
	4	T626	1	64	1	1.31E+04	1.02E+01	19	13371	2197	15568	104	17	122	0	200
	5	T626	1	64	1	2.17E+04	1.69E+01	19	22082	2493	24574	173	20	192	0	121
	6	T626	1	64	1	2.77E+04	2.16E+01	22	28186	2160	30346	220	17	237	0	95
	7	T626	1	64	1	3.46E+04	2.69E+01	29	35211	2009	37220	275	16	291	0	76
	8	T626	1	64	1	4.09E+04	3.20E+01	33	41676	1832	43508	326	14	340	0	64
	9	T626	1	64	1	4.52E+04	3.53E+01	33	46035	1487	47522	360	12	371	0	58
	10	T626	1	64	1	4.92E+04	3.86E+01	33	50154	1113	51268	392	9	401	0	53

Run 7 Tst 1

Top 5 Timed Foreground Events

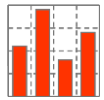
Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db file sequential read	2,025,346	52,705	26	46.7	User I/O
db flash cache single block ph	645,759	769	1	.7	User I/O

Run 7 Tst 10

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db flash cache single block ph	2,653,213	2,921	1	41.0	User I/O
db file sequential read	5,026	95	19	1.3	User I/O

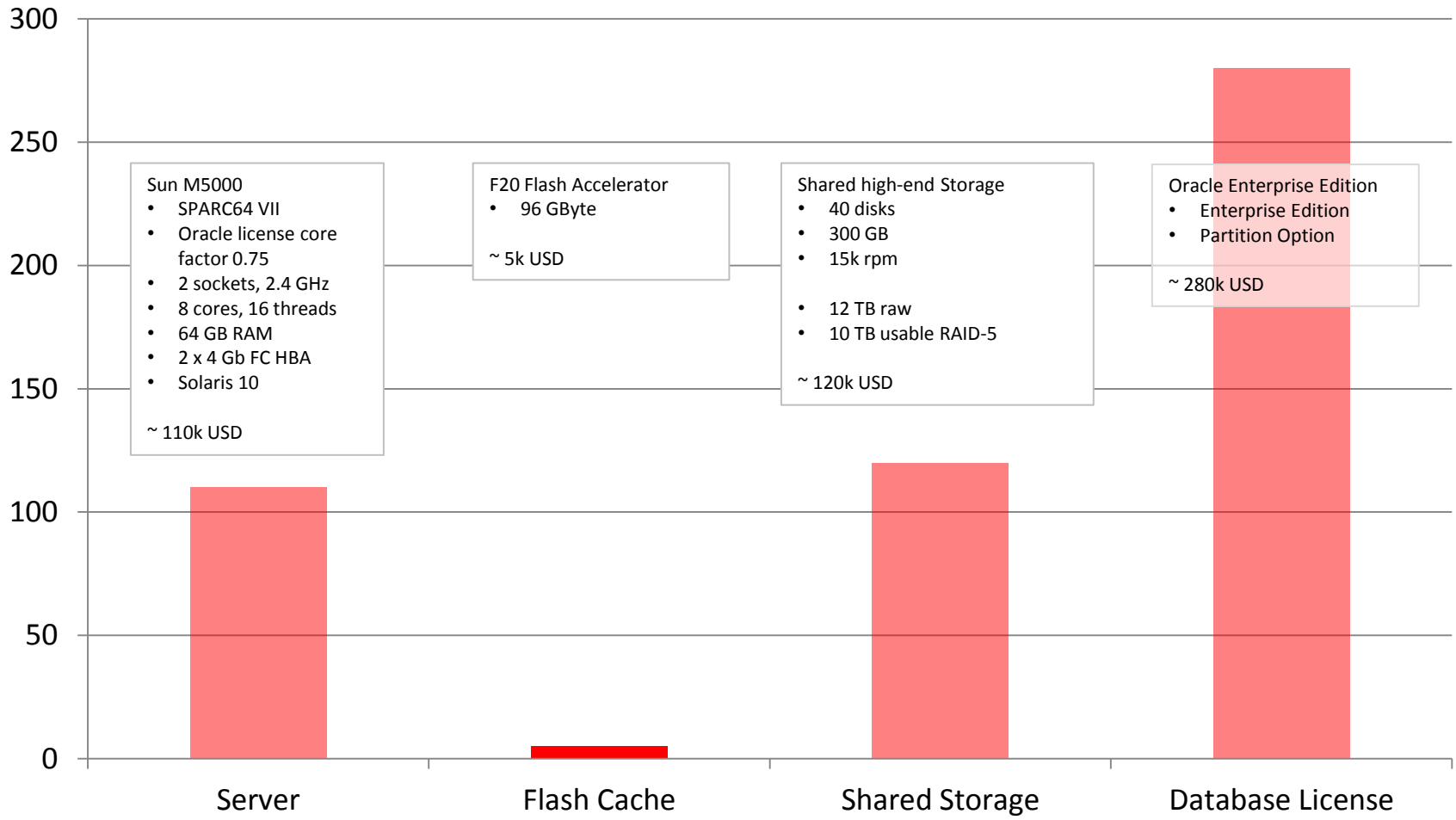
Contents



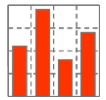
Conclusion

Sun M5000 with F20 Flash Cache

All prices are list prices



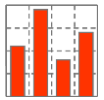
Conclusion



Sun M5000 with F20 Flash Cache

- Flash Cache advantages
  - Very high random and sequential I/O throughput
  - Service times for single block access < 1 msec
  - Unburden shared storage system

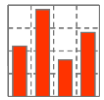
Conclusion



Sun M5000 with F20 Flash Cache

- Flash Cache management
  - Automatically managed by Oracle
    - Oracle instance parameter `db_flash_cache_size`
    - All database objects benefit from Flash Cache
  - Manually managed by application
    - Objects can be pinned to Flash Cache
    - Useful for hotspot database objects (needs capacity planning)
    - `SQL> ALTER TABLE <name> STORAGE (FLASH_CACHE KEEP);`

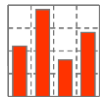
Conclusion



Sun M5000 with F20 Flash Cache

- Flash Cache limitations
  - For read operations only
  - Available for Oracle 11.2 on Solaris and OEL only
- Bottom line
  - A small hardware investment dramatically increases I/O throughput and decreases I/O latency by factors

Conclusion



Sun M5000 with F20 Flash Cache

- Benchware uses fair, reproducible and representative benchmark tests delivering understandable *key performance metrics* (KPM)
- Benchware uses a list of defined *price performance ratios* (PPR) to evaluate platform cost
- Benchware publishes *price performance ratios* (PPR) to its customers only

swiss precision in performance measurement

www.benchware.ch

info@benchware.ch