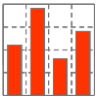


---

# Oracle Performance on Exadata X2-2 half-rack high-capacity with write back flash cache

Benchmark Report

December 2012

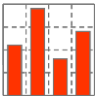


## **1 About Benchware**

2 Architecture Overview

3 Some Performance Tests

4 Conclusion

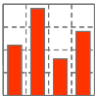


---

## Services and Products

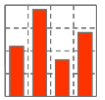
Strong foundation in core technologies such as Oracle database systems, server and storage systems

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

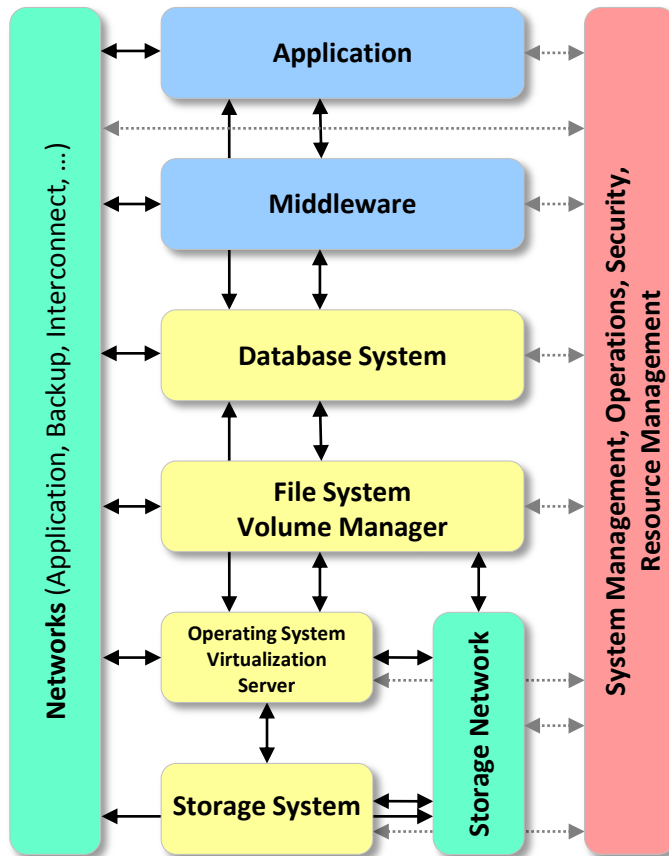


## Value proposition

- 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)



## Performance of complex Oracle platforms is not predictable



*System architects have a wide choice of components, technologies and configurations*

### Networks (IP-based)

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

### Oracle Database

Different versions, patches and options, about one 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

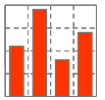
Different volume managers (VxVM, ASM) and file systems (UFS, VxFS, ext3, JFS, ZFS, raw devices), different I/O methods (async, direct), a lot of configuration parameters (#LUNS, queue depth, max i/o unit), software striping and/or mirroring, multipathing.

### Server & Operating System

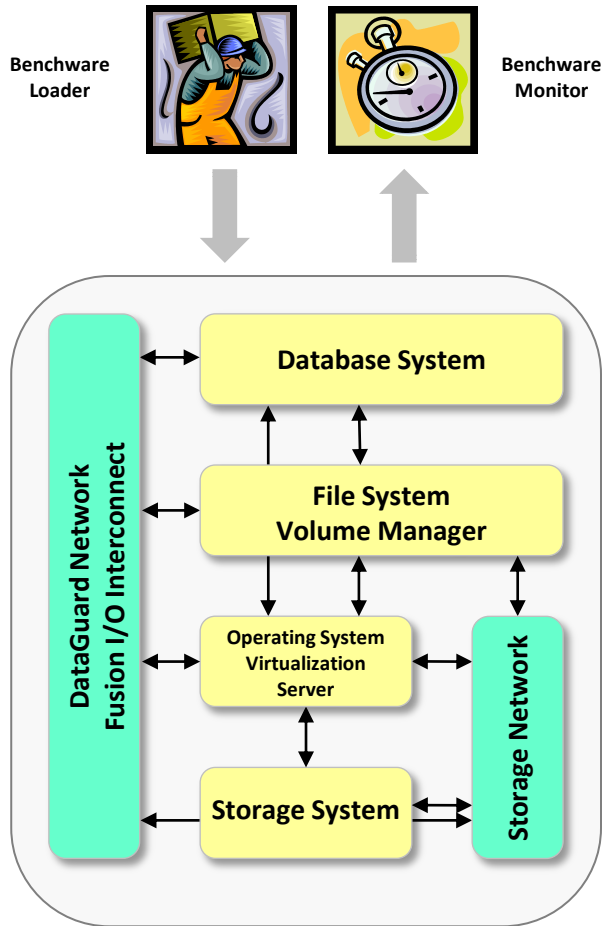
Different server systems, processors and CPU architectures, (x86, IA-64, UltraSparc, SPARC64, Power), #cores, multithreading, main memory, bus architecture. Different operating systems and patches, over one hundred configuration parameters, virtualization of resources.

### 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.



Complex architecture of Oracle platforms requires benchmarking



*Object of measurement: Oracle platform*

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



Key performance metrics of Oracle platforms should be as easy as key performance metrics used in the automobile industry

## Engine

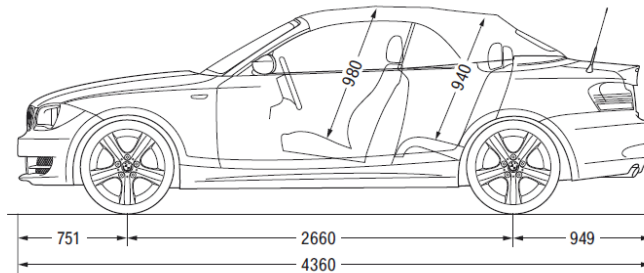
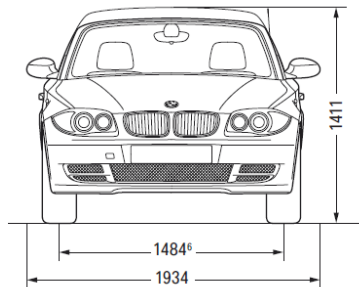
Cylinders/valves	4/4
Capacity in ccm	1,995
Stroke/bore in mm	90.0/84.0
Max. output in kW (hp) at 1/min	105 (143)/6,000
Max. torque in Nm at 1/min	190/4,250
Power-to-weight ratio (EU) in kg/hp	10.5

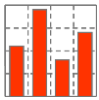
## Performance

Drag (cw)	0.32
Top speed (km/h)	210
Acceleration 0 - 100 km/h (in s)	9.3
Acceleration 0 - 1,000 m (in s)	30.6
Acceleration 80 - 120 km/h in 4th/5th gear (in s)	9.6/12.5



Source: [www.bmw.de](http://www.bmw.de)





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

CPU Performance CPU-bound PL/SQL operations	OLTP systems	DWH systems	Proof of Efficiency	Key Performance Metrics	Unit
<ul style="list-style-type: none"> <li>pl/sql basic operations arithmetic mix, string mix</li> </ul>	★★	★★	multithreading virtualization	speed throughput	[s] [ops]
<ul style="list-style-type: none"> <li>pl/sql algorithms fibonacci, prime numbers</li> </ul>	★★★★	★★			

Server Performance Server-bound SQL operations completely in SGA	OLTP systems	DWH systems	Proof of Efficiency	Key Performance Metrics	Unit
<ul style="list-style-type: none"> <li>in-memory SQL full scan, index access</li> </ul>	★★★★	★★	scalability cc-uma virtualization	speed throughput	[μs] [s] [bgps] [tps] [rps] [MBps]
<ul style="list-style-type: none"> <li>pl/sql algorithms quicksort</li> </ul>	★★	★			

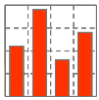
[s] seconds  
 [ms] milli seconds ( $10^{-3}$ )  
 [μs] micro seconds ( $10^{-6}$ )  
 [ns] nano seconds ( $10^{-9}$ )

[bgps] buffer gets per second  
 [rps] rows per second  
 [tps] transactions per second  
 [ops] operations per second

[MBps] mega bytes per second  
 [GBps] giga bytes per second  
 [iops] i/o operations per second  
 [qpm] queries per minute

★ less important  
 ★★ important  
 ★★★ very important





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

Storage Performance I/O-bound Oracle operations	OLTP systems	DWH systems	Proof of Efficiency	Key Performance Metrics	Unit
<ul style="list-style-type: none"> <li>sequential I/O 1 MByte, read and write</li> </ul>	★★	★★★★	data integrity tiering, pooling striping virtualization replication	throughput service time	[ms] [MBps] [GBps] [iops]
<ul style="list-style-type: none"> <li>random I/O 8 kByte, read and write</li> </ul>	★★★★	★			

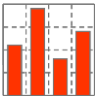
Database Performance Mixed resource usage: CPU, memory, storage	OLTP systems	DWH systems	Proof of Efficiency	Key Performance Metrics	Unit
<ul style="list-style-type: none"> <li>data load uncompressed, compressed</li> </ul>	★★	★★★★	scalability	speed throughput	[ms] [s] [rps] [tps] [qpm]
<ul style="list-style-type: none"> <li>data scan</li> </ul>	★	★★★★			
<ul style="list-style-type: none"> <li>data aggregation &amp; reports</li> </ul>	★★	★★★★		service time	[s]
<ul style="list-style-type: none"> <li>OLTP transactions insert, select, update</li> </ul>	★★★★	★			

[s] seconds  
 [ms] milli seconds ( $10^{-3}$ )  
 [μs] micro seconds ( $10^{-6}$ )  
 [ns] nano seconds ( $10^{-9}$ )

[bps] buffers per second  
 [rps] rows per second  
 [tps] transactions per second  
 [ops] operations per second

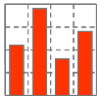
[MBps] mega bytes per second  
 [GBps] giga bytes per second  
 [iops] i/o operations per second  
 [qpm] queries per minute

★ less important  
 ★★ important  
 ★★★ very important

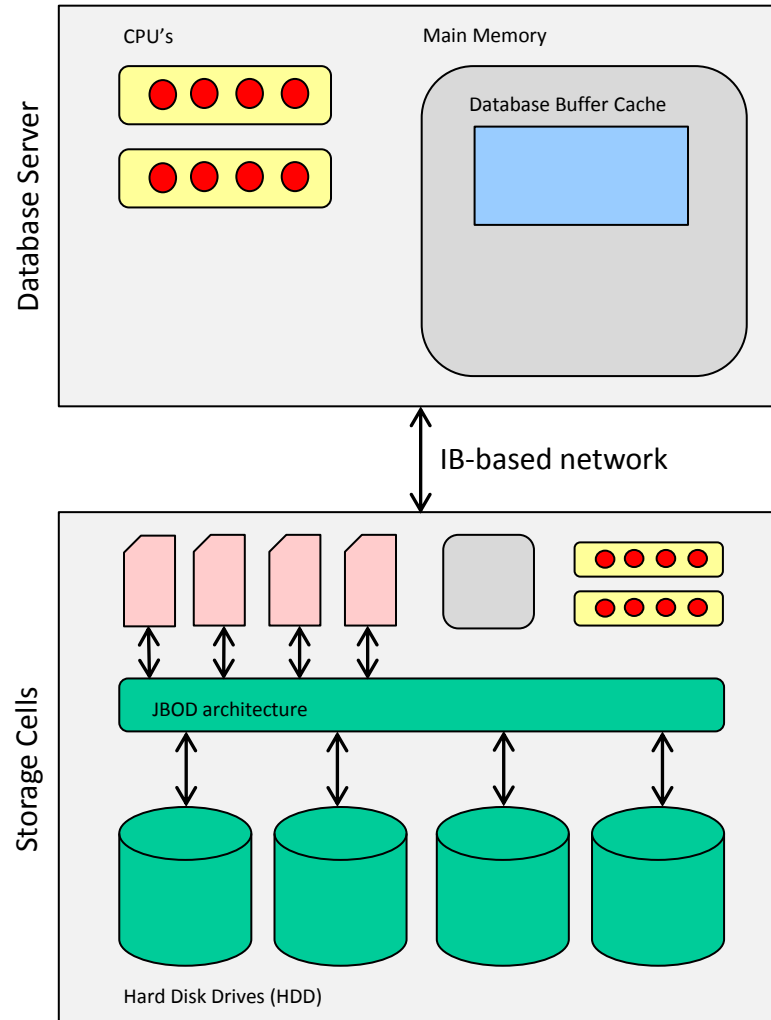


- 1 About Benchware
- 2 Architecture Overview**
- 3 Some Performance Tests
- 4 Conclusion

# Architecture Overview



## Oracle Exadata



Auto Tiering for read & write:

- RAM  $100 \times 10^{-9} \text{ s}$
- Flash Cache  $250 \times 10^{-6} \text{ s}$
- Hard Disk Drives  $1 \times 10^{-2} \text{ s}$

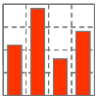
Exadata X2-2 HR HC (Sep 2010):

- HDD capacity raw 168 TByte
- HDD capacity usable  $\sim 80 \text{ TByte}$
- Flash capacity 2.6 TByte

Exadata X3-2 HR HC (Okt 2012):

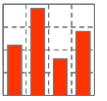
- HDD capacity raw 252 TByte
- HDD capacity usable  $\sim 80 \text{ TByte}$
- Flash capacity  $\sim 11 \text{ TByte}$

# Architecture Overview



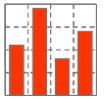
Oracle Exadata Software 11.2.3.2.0

- New write back cache instead of write through cache
- Auto tiering
- Still manual configuration possible

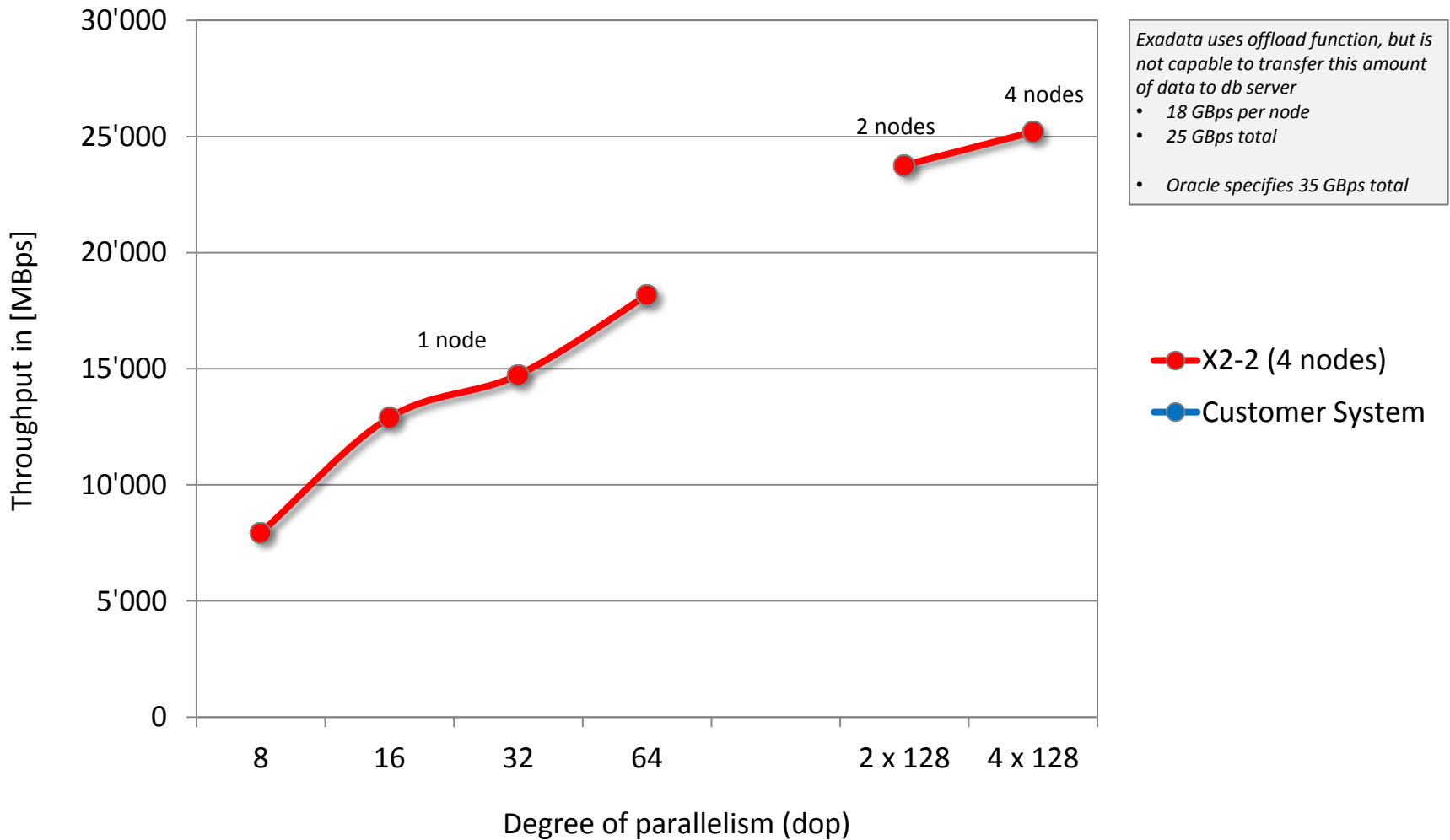


- 1 About Benchware
- 2 Architecture Overview
- 3 Some Performance Tests**
- 4 Conclusion

# Some Performance Tests



## Oracle sequential read (from flash cache)



# Some Performance Tests



## Oracle sequential read (from flash cache)

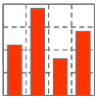
X2-2

Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
1	115	T426	1	1	8	2.626E+07	0.000E+00	6	7939	57	7996	7890	1	7890	0	39
	116	T426	1	1	16	4.267E+07	0.000E+00	10	12896	81	12977	12821	1	12822	0	24
	117	T426	1	1	32	4.876E+07	0.000E+00	11	14731	89	14820	14653	1	14654	0	21
	118	T426	1	1	64	6.024E+07	1.000E-01	14	18168	99	18267	18100	1	18101	0	17
	119	T426	1	1	128	5.689E+07	1.000E-01	13	17190	108	17298	17095	1	17096	0	18
	120	T426	2	2	128	7.877E+07	1.000E-01	19	23760	74	23834	23669	1	23670	0	26
	121	T426	4	4	128	8.359E+07	1.000E-01	20	25200	39	25239	25118	0	25119	0	49

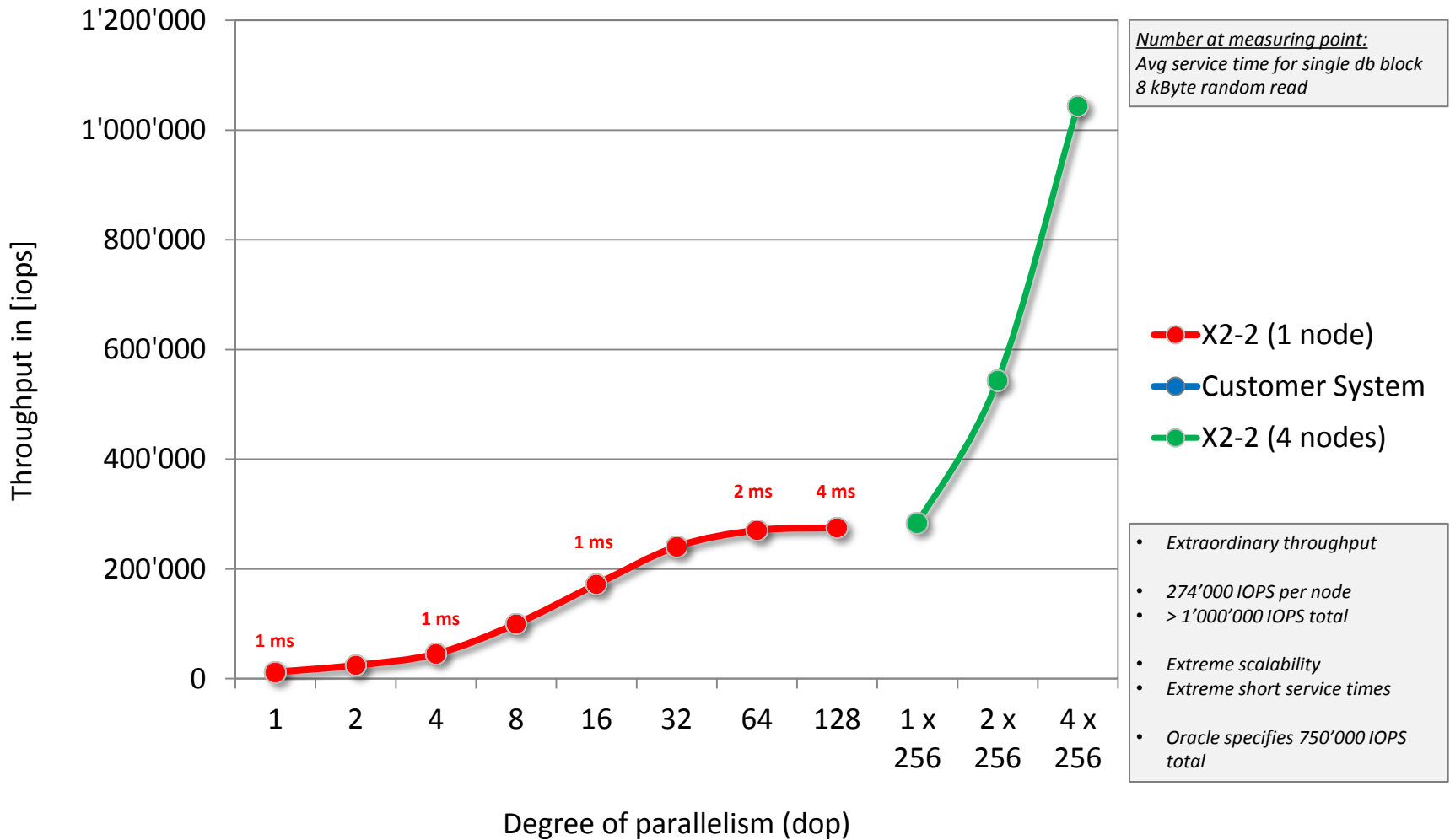
- Avg I/O size about 1 MByte

- #N - describes the number of RAC nodes involved in this test
- #J - describes the number of jobs running the workload, jobs are distributed round-robin over all configured node in #N
- CPU [%] – total CPU utilization in cluster

# Some Performance Tests



## Oracle random read (from flash cache)





# Some Performance Tests



## Oracle random read (from flash cache)

X2-2 1<sup>st</sup> test

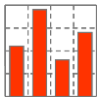
Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
1	152	T626	1	1	1	1.265E+04	1.200E+00	2	11855	52	11907	93	1	93	0	207
153	T626	1	2	1	1	2.597E+04	2.500E+00	3	24251	96	24347	190	1	191	1	202
154	T626	1	4	1	1	4.836E+04	4.700E+00	5	45046	165	45211	352	3	355	1	217
155	T626	1	8	1	1	1.076E+05	1.050E+01	6	99826	356	100182	780	6	786	3	195
156	T626	1	16	1	1	1.856E+05	1.810E+01	11	171834	598	172432	1343	9	1352	5	226
157	T626	1	32	1	1	2.600E+05	2.540E+01	18	240665	830	241495	1880	13	1893	7	323
158	T626	1	64	1	1	2.898E+05	2.830E+01	22	270388	931	271319	2119	21	2140	7	579
159	T626	1	128	1	1	2.965E+05	2.890E+01	25	274670	951	275621	2152	21	2174	7	1132
160	T626	2	256	1	1	3.758E+05	3.670E+01	35	353579	1205	354784	2770	27	2798	10	1785
161	T626	4	512	1	1	5.261E+05	5.140E+01	46	490000	1683	491683	3839	38	3877	13	2551

X2-2 2<sup>nd</sup> test

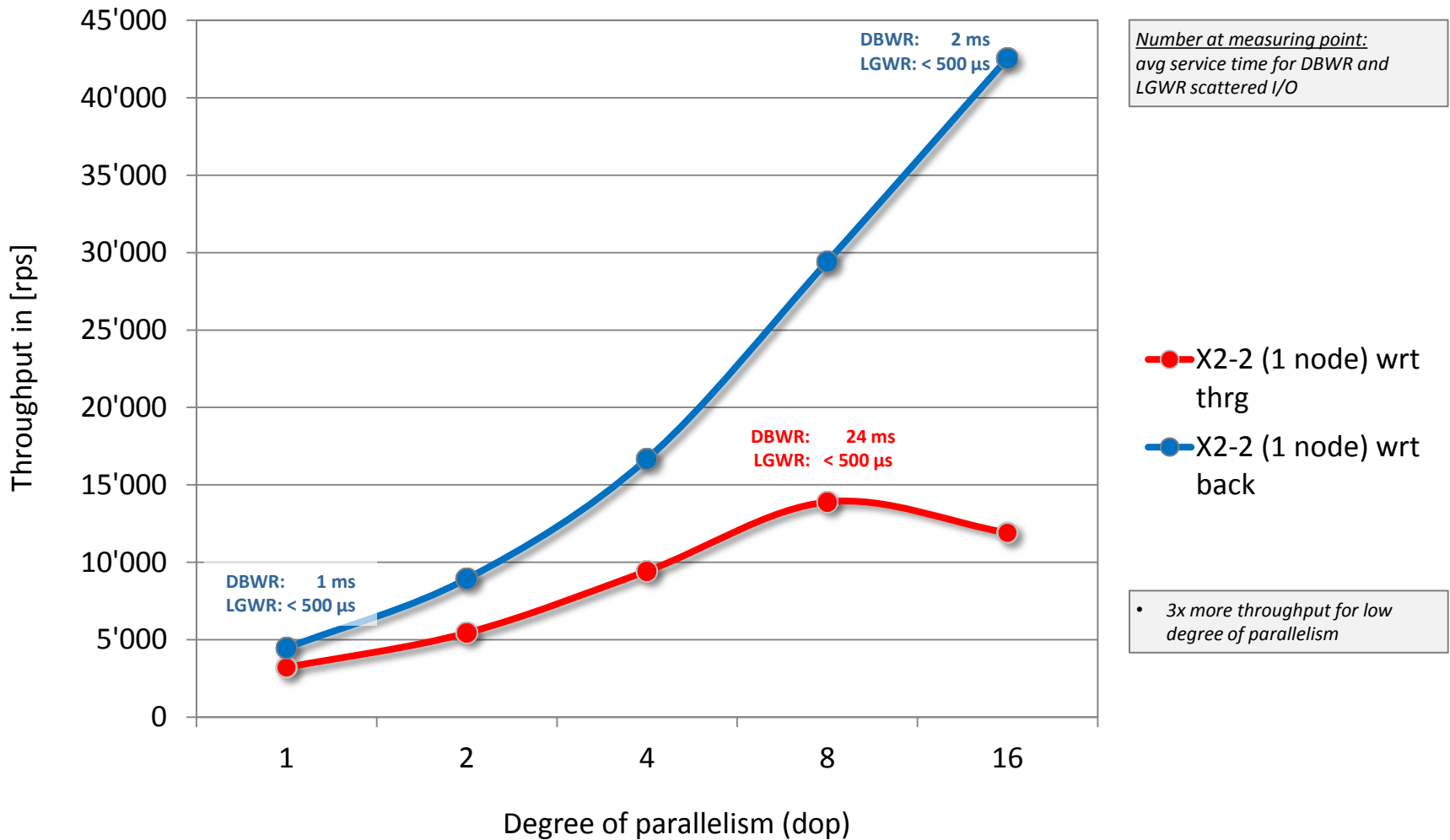
Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
2	4	T626	1	128	1	2.924E+05	2.850E+01	23	282411	939	283350	2213	22	2234	8	1148
5	T626	1	256	1	1	2.921E+05	2.850E+01	24	283473	939	284412	2222	23	2246	8	2297
6	T626	2	256	1	1	5.493E+05	5.360E+01	43	531680	1759	533439	4165	41	4206	15	1222
7	T626	2	512	1	1	5.602E+05	5.470E+01	46	542431	959	543390	4253	38	4291	15	2396
8	T626	4	512	1	1	1.077E+06	1.051E+02	83	1043861	85	1043946	8178	53	8231	29	1247
9	T626	4	1024	1	1	1.077E+06	1.052E+02	86	1044097	91	1044188	8186	60	8246	29	2434

• Test had been repeated several times to warm up cell flash cache

# Some Performance Tests



Oracle OLTP insert transaction, 2 rows per commit



# Some Performance Tests



Oracle OLTP insert transaction, 2 rows per commit

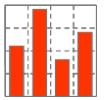
X2-2 write  
through cache

Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
1	60	T311	1	1	1	3.125E+03	1.563E+03	2	393	1914	2307	18	8	27	4	40
61	T311	1	1	1	1	3.205E+03	1.603E+03	2	39	1893	1932	0	7	8	4	39
62	T311	1	2	1	1	5.435E+03	2.717E+03	3	33	3764	3797	0	12	13	6	46
63	T311	1	4	1	1	9.434E+03	4.717E+03	4	31	4340	4371	0	24	24	12	53
64	T311	1	8	1	1	1.389E+04	6.944E+03	6	26	4935	4961	0	41	42	18	72
65	T311	1	16	1	1	1.196E+04	5.988E+03	6	18	5134	5152	0	45	46	16	167

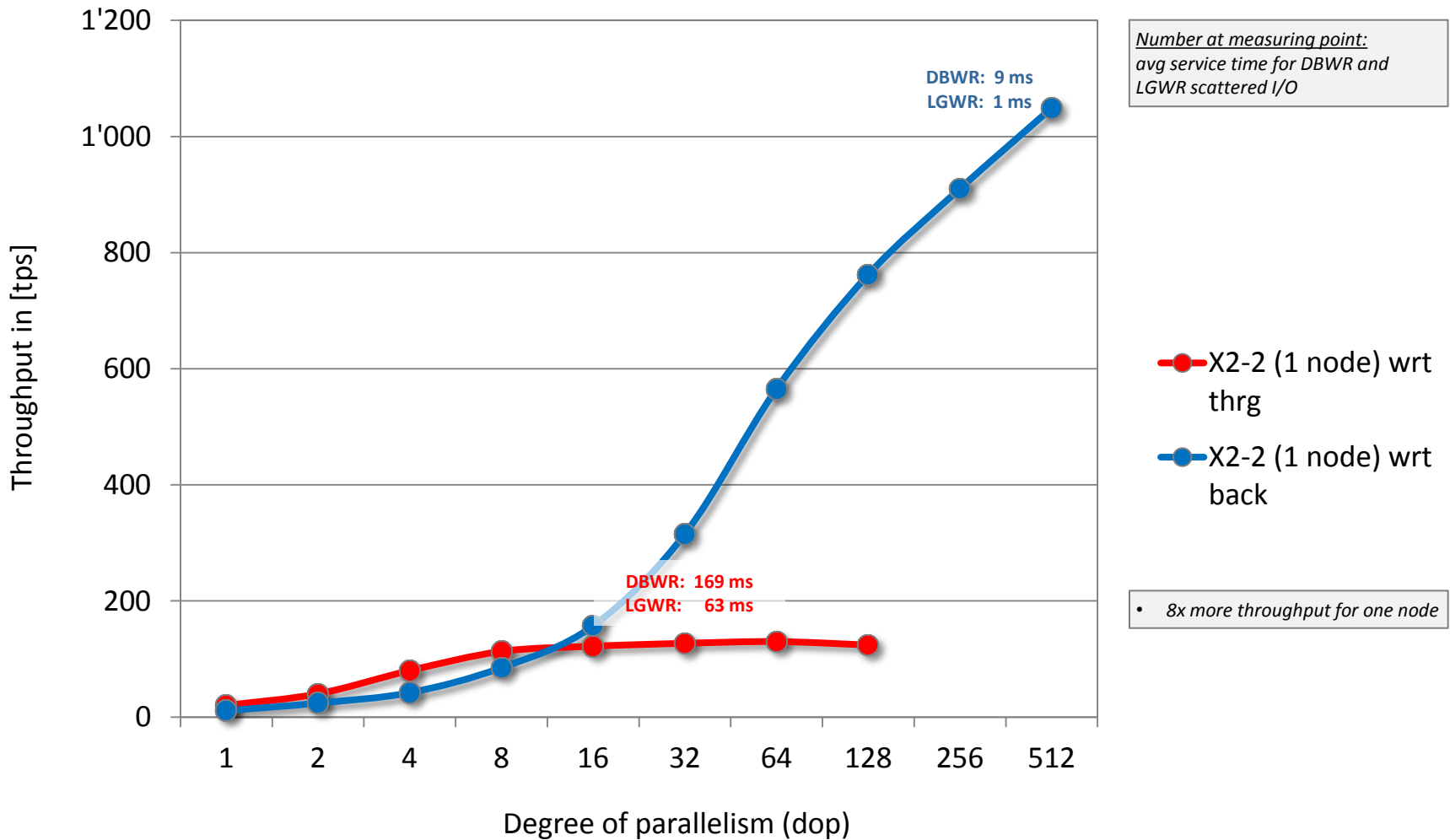
X2-2 write  
back cache

Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
7	7	T311	1	1	1	4.808E+03	2.404E+03	9	10	5288	5298	0	19	19	7	26
8	T311	1	1	1	1	4.464E+03	2.232E+03	8	7	4867	4874	0	17	18	6	28
9	T311	1	2	1	1	8.929E+03	4.464E+03	12	7	9759	9766	0	35	35	12	28
10	T311	1	4	1	1	1.667E+04	8.333E+03	19	8	11333	11341	0	66	66	24	30
11	T311	1	8	1	1	2.941E+04	1.471E+04	33	7	11311	11318	0	120	120	43	34
12	T311	1	16	1	1	4.255E+04	2.128E+04	55	6	12349	12355	0	182	182	66	47

# Some Performance Tests



Oracle OLTP update transaction, primary key access, 32 rows per tx



# Some Performance Tests



Oracle OLTP update transaction, primary key access, 32 rows per tx

X2-2 write  
through cache

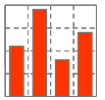
Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
1	142	T716	1	1	1	6.425E+02	2.010E+01	1	1001	697	1698	61	6	68	1	51
143	T716	1	2	1	1	1.285E+03	4.020E+01	2	1913	1375	3288	60	12	71	1	51
144	T716	1	4	1	1	2.570E+03	8.030E+01	3	3712	2678	6390	57	22	80	1	51
145	T716	1	8	1	1	3.641E+03	1.138E+02	3	5270	3760	9030	60	31	92	2	72
146	T716	1	16	1	1	3.913E+03	1.223E+02	4	5682	4007	9689	56	33	90	2	134
147	T716	1	32	1	1	4.080E+03	1.275E+02	3	5918	4140	10058	60	35	94	2	257
148	T716	1	64	1	1	4.178E+03	1.305E+02	3	6182	4280	10462	60	36	97	3	502
149	T716	1	128	1	1	3.961E+03	1.238E+02	2	5915	4020	9935	59	34	92	2	1059
150	T716	2	256	1	1	4.599E+03	1.437E+02	3	6965	4595	11560	63	38	101	2	1824
151	T716	4	512	1	1	4.903E+03	1.532E+02	4	7190	4760	11950	57	40	97	3	3422

X2-2 write  
back cache

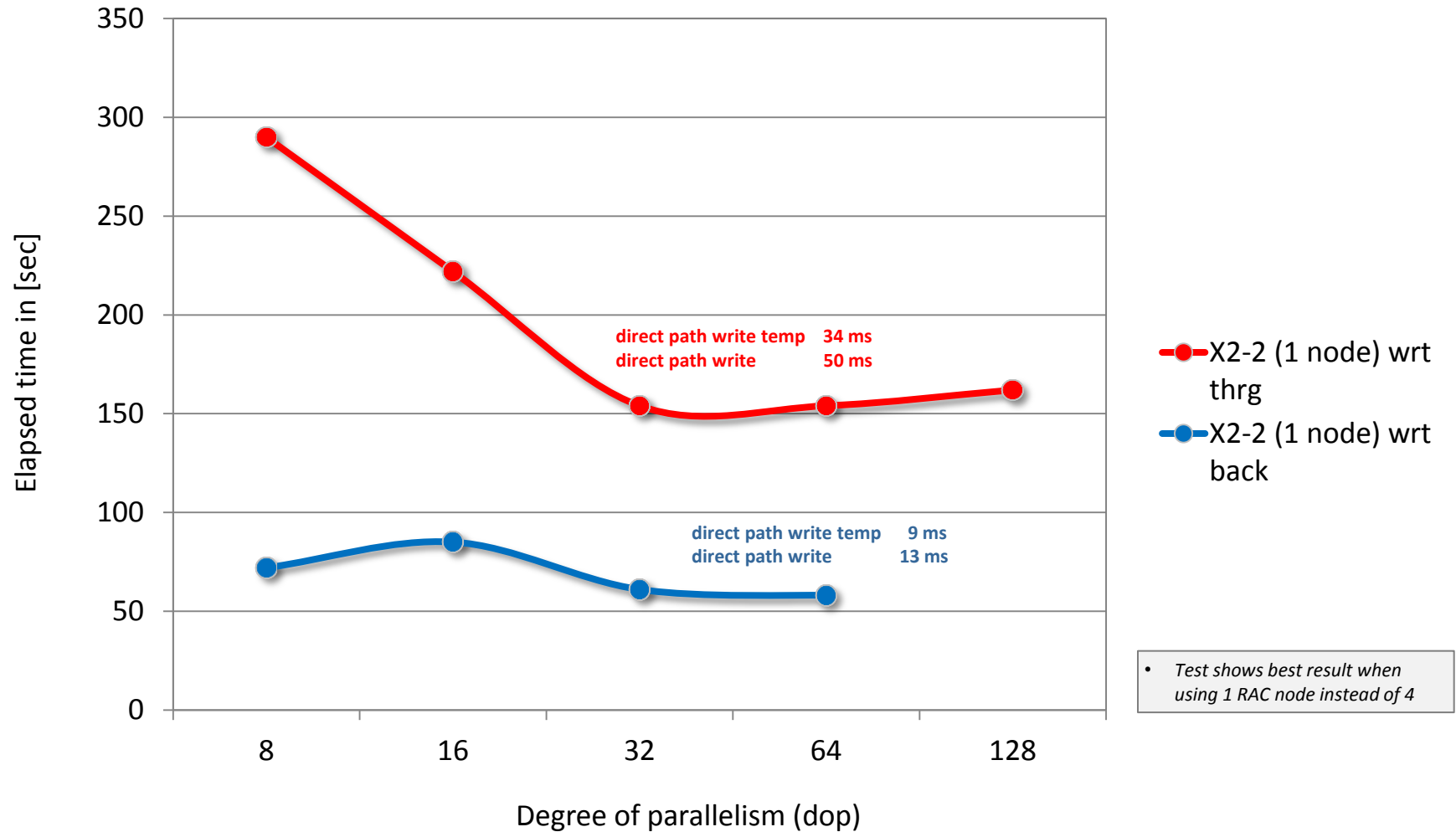
Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
22	1	T716	1	1	1	3.413E+02	1.070E+01	1	881	478	1359	10	4	15	0	24
2	T716	1	2	1	1	7.802E+02	2.440E+01	2	1644	1011	2655	28	9	37	0	21
3	T716	1	4	1	1	1.365E+03	4.270E+01	5	2693	1720	4413	74	14	88	1	24
4	T716	1	8	1	1	2.731E+03	8.530E+01	4	5070	3460	8530	178	28	206	2	24
5	T716	1	16	1	1	5.041E+03	1.575E+02	6	8249	6479	14728	224	51	275	3	26
6	T716	1	32	1	1	1.008E+04	3.151E+02	9	15905	13237	29142	126	104	230	6	26
7	T716	1	64	1	1	1.808E+04	5.650E+02	12	28416	25282	53698	222	201	423	10	29
8	T716	1	128	1	1	2.439E+04	7.620E+02	17	38426	31606	70032	300	258	558	14	43
9	T716	1	256	1	1	2.913E+04	9.102E+02	17	45843	31443	77286	358	273	631	16	72
10	T716	1	512	1	1	3.355E+04	1.049E+03	20	52698	35210	87908	412	308	720	19	125
1	T716	4	1024	1	1	7.109E+04	2.222E+03	49	93744	76287	170031	773	634	1407	38	118
2	T716	4	2048	1	1	8.265E+04	2.583E+03	58	95510	87705	183215	746	737	1483	44	203

• Up to 18x more random writes

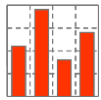
# Some Performance Tests



Database aggregation, create unique b-tree index,  $10^9$  rows



# Some Performance Tests



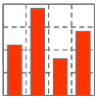
Database aggregation, create b-tree index 10<sup>9</sup> rows

X2-2 write  
through cache

Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
1	127	T512	1	1	8	0.000E+00	0.000E+00	6	1841	4318	6159	1255	368	1623	85	290
128	T512	1	1	16	0.000E+00	0.000E+00	6	2517	5796	8313	1667	483	2150	111	222	
129	T512	1	1	32	0.000E+00	0.000E+00	8	3373	8102	11475	2324	641	2965	160	154	
130	T512	1	1	64	0.000E+00	0.000E+00	8	3298	8123	11421	2326	645	2971	160	154	
131	T512	1	1	128	0.000E+00	0.000E+00	8	3130	7659	10789	2211	612	2823	152	162	

X2-2 write  
back cache

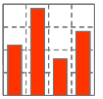
Run	Tst	Code	#N	#J	#T	Rows/sec [rps]	Ops/sec [ops]	CPU [%]	Physical read [iops]	Physical write [iops]	Total [iops]	Physical read [MBps]	Physical write [MBps]	Total [MBps]	REDO [MBps]	Time [sec]
7	64	T512	1	1	8	0.000E+00	0.000E+00	28	2407	6129	8536	2264	651	2915	171	72
65	T512	1	1	16	0.000E+00	0.000E+00	31	5582	6216	11798	2109	744	2853	145	85	
66	T512	1	1	32	0.000E+00	0.000E+00	51	4138	8546	12684	2941	1037	3978	202	61	
67	T512	1	1	64	0.000E+00	0.000E+00	54	4349	8979	13328	3088	1087	4175	212	58	



- 1 About Benchware
- 2 Architecture Overview
- 3 Some Performance Tests
- 4 Conclusion**



# Conclusion



Oracle Exadata Software 11.2.3.2.0

- Best of both worlds
  - HDD technology for capacity
  - Flash technology for performance
- Random write up to 18x more throughput
- Exadata is now much more suitable for OLTP systems with high volume of mutating transactions
- Even existing Exadata V2 and Exadata X2 customers benefit from this software update without any additional charges!

**BENCHWARE**

*swiss precision in performance measurement*

*[www.benchmarkware.ch](http://www.benchmarkware.ch)*

*[info@benchmarkware.ch](mailto:info@benchmarkware.ch)*