Exadata, the new Oracle database machine

7 October 2008

Anton Topurov Eric Grancher Elzbieta Gajewska-Dendek







- Oracle Database Machine
- Exadata Storage Server Hardware
- Exadata Architecture and Features
- Beta Testing
- Test Results
- Conclusions





Oracle Database Machine

- 8 DL360 Oracle Database servers
 - 2 quad-core Intel Xeon, 32GB RAM
 - Oracle Enterprise Linux
 - Oracle RAC
- 14 Exadata Storage Cells (SAS or SATA)
 - Up to 14 TB uncompressed user data on SAS
 - Up to 46 TB uncompressed user data on SATA
- 4 InfiniBand switches
- I Gigabit Ethernet switch
- Keyboard, Video, Mouse (KVM) hardware
- Hardware Warranty
 - 3 YR Parts/3 YR Labor/3 YR On-site
 - 24X7, 4 Hour response time



Exadata Storage Server Hardware

Exadata Storage Server



Racked Exadata Storage Servers



- Building block of massively parallel Exadata Storage Grid
 - Up to 1GB/sec data bandwidth per cell
- HP DL180 G5
 - 2 Intel quad-core processors
 - 8GB RAM
 - Dual-port 4X DDR InfiniBand card
 - 12 SAS or SATA disks
- Software pre-installed
 - Oracle Exadata Storage Server Software
 - Oracle Enterprise Linux
 - HP Management Software
- Hardware Warranty
 - 3 YR Parts/3 YR Labor/3 YR On-site
 - 24X7, 4 Hour response



Exadata Important Features

Database aware storage – does:

- Predicate filtering
- Column projection filtering
- Join processing (star-joins for DWH)
- Tablespace creation
 - eliminates the I/O associated with the creation and writing of tablespace blocks
- I/O resource management inter and intra database

Exadata Architecture I







Exadata Architecture II





Testing Exadata Storage

[JST]: Exadata beta testing



- Beta testing since April
- 3 day learning event at Oracle UK site
- Developed PVSS swingbench benchmark
- Remote tests done in August
- More remote tests to be done this month
- Possibility to get the hardware onsite



CERN SUSER10 SUSER02 openlab Operators SUSER01 **EVENTLASTVAL1** ELEMENT_ID Inserts+Queries TS VALUE_NUMBER STATUS MANAGER USER TEXT VALUE_STRING VALUE_TIMESTAMP CORRVALUE_NUMBER OLVALUE NUMBER CORRVALUE_STRING OLVALUE STRING CORRVALUE_TIMESTAMP OLVALUE TIMESTAMP Valarch.InsertHistory() EVENTHISTORY_001000041 PK ELEMENT_ID PK TS PK SYS ID VALUE NUMBER STATUS MANAGER USER BASE TEXT VALUE_STRING nnn VALUE_TIMESTAMP CORRVALUE NUMBER OLVALUE_NUMBER CORRVALUE STRING **PVSS** Hardware OLVALUE_STRING CORRVALUE TIMESTAMP OLVALUE TIMESTAMP Clients tools

PVSS Workload Simulator



Created within Swingbench Framework



Swingbench in action



子 Applications Places System 国 🔮 🚳 oracle Thu Aug 28, 7:54 PM 🐗 SwingBench 2.3.0.381 (SWPVSS1) SwingBench 2.3.0.381 (SWPVSS2) Time Remaining : 0:00:00 Time Remaining : 0:00:00 15 15 Users Users 420 Transactions per Minute 486 Transactions per Minute Transactions per Second : 16 an bhi san dr. 1 tit t t-bh Ba Transactions per Second 10 CPU CPU Disk Activity Disk Activity Property Value Property Value Benchmark Name PVSS Benchmark Benchmark Name "PVSS Benchmark" Connect String SWPVSS2 Connect String SWPVSS1 Coordinator Coordinator Driver Type Oracle10g Type II jdbc driver (oci) Driver Type Oracle10g Type II jdbc driver (oci) Maximum Think Time 0 Maximum Think Time 0 Minimum Think Time 0 Minimum Think Time 0 Query Time Out 600 Query Time Out 600 User Count 15 User Count 15 User Name SUSER03 User Name SUSER01 ٧ SwingBench 2.3.0.381 (SWPVSS2) _ = × SwingBench 2.3.0.381 (SWPVSS1) Time Remaining : 0:00:00 Time Remaining : 0:00:00 Users 15 15 Users 432 Transactions per Minute Transactions per Minute 488 Transactions per Second i talana a dala fili dia natatri di 🕄 indo Transactions per Second 13 CPU CPU Disk Activity Disk Activity Property Value Property Value Benchmark Name "PVSS Benchmark" Benchmark Name "PVSS Benchmark" Connect String SWPVSS2 SWPVSS1 Connect String Coordinator Coordinator Driver Type Oracle10g Type II jdbc driver (oci) Oracle10g Type II jdbc driver (oci) Driver Type Maximum Think Time 0 Maximum Think Time 0 Minimum Think Time 0 Minimum Think Time 0 Query Time Out 600 Query Time Out 600 User Count 15 User Count 15 User Name SUSER04 User Name SUSER02 🛃 SwingBench 2.3.0... [oracle@sof:~/sw... & SwingBench 2.3.... 🛃 SwingBench 2.3.0... 🛃 SwingBench 2.3.... [root@sof:~]



CERN setup test results



CERN openlab – 2008

Exadata test results



CE

oper

CET Time

Conclusions



CERN setup

- 2 node RAC can sustain 30-40 000 changes/s
- Degradation of performance due to tablespace creation

Exadata setup

- We have only intermediate results, more testing is needed
- 4 node RAC, could cope with 150 000 changes/s
- Performance degradation due to small SGA (1G)
- With sga_target=10G, much better performance
- Redo logs were the bottleneck
- More tests scheduled to identify the full performance of Exadata



