Hammerdb Test for OracleRAC with Memblaze® PBlaze SSD

Executive Summary

Accompany with the Oracle Engineered Systems portfolio evolution, the NVMe SSD stepped into enterprise application total solution area. The higher bandwidth, bigger IOPS, and lower latency which are the native NVMe SSD characters, would accelerate the database performance.

In the whitepaper, we will describe the architecture of distributed data server and oracle database server which connected with InfiniBand network. In each data server, there is on piece of PBlaze4 SSD device. Namely, run an oracle RAC on distributed storage system which used all NVMe devices. What amazing. We run a Hammerdb application which could get the Oracle OLTP performance on the test bed, and do a check for the results.

InfiniBand Introduction

InfiniBand is a high-performance, multi-purpose network architecture based on a switch design often called a "switched fabric." InfiniBand was designed for use in I/O networks such as storage area networks (SAN) or in cluster networks. It has become a leading standard in high-performance computing.

InfiniBand implements a messaging service for applications called Channel I/O that bypasses network operating systems in order to achieve high performance in specialized environments. It provides the ability for two InfiniBand-enabled applications to create a direct communication channel having send and receive queues called Queue Pairs. The queues map to memory spaces accessible to each application for data sharing (called Remote Direct Memory Access or RDMA).

About the Test

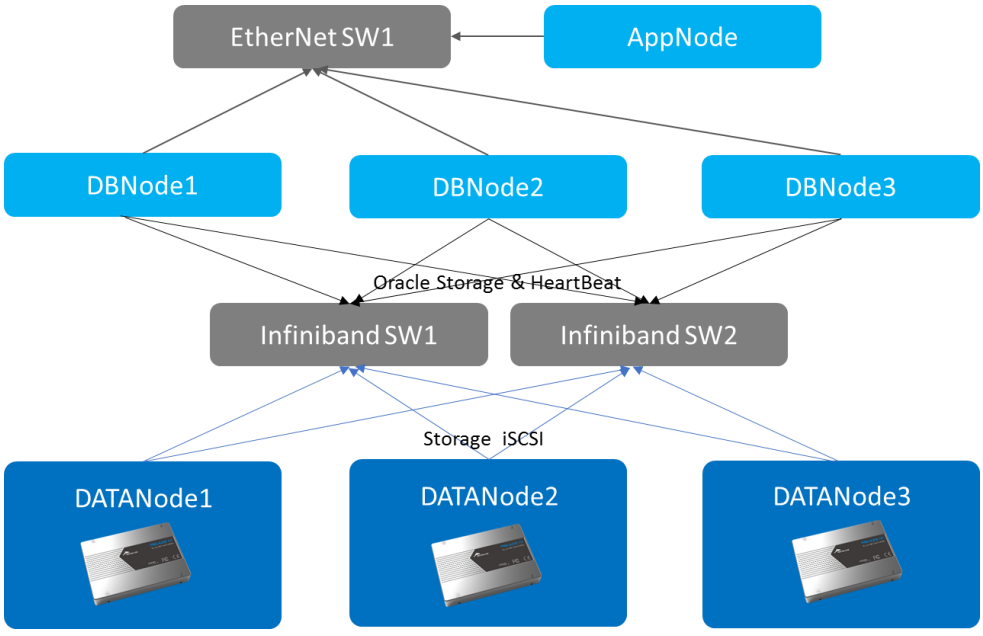

Figure 1. Test bed logical topology

In application server, run Hammerdb which will produce TPCC transactions, and push the jobs to the Database server by the Ether Network. The Database Server will transfer the data in the Storage server by the InfiniBand network. The data of database was stored in the PBlaze SSD on the Storage server, and the partition of the PBlaze SSD will be export by the STGT software. In the Database server, the block device used as Oracle ASM disk, and in the Oracle ASM instance, the data mirroring was deployed.

The overall test environment is as follows:

Application Server * 1

- Dell PowerEdge R720
- 1* Intel XeonE5-2630(6 cores) v3 CPU
- 2* 8GB DRAM
- CentOS 6.5

Database Server * 3

- Dell PowerEdge R720
- 2* Intel XeonE5-2680(12 cores) v3 CPU
- 8* 16GB DRAM
- CentOS 6.5

Storage Server * 3

- Dell PowerEdge R730xd
- 2* Intel XeonE5-2630(6 cores) v3 CPU
- 8* 8GB DRAM
- 1 x Memblaze 3.2T PBlaze4
- CentOS 6.5

Benchmark Tool

- HammerDB 2.19, TPC-C test tool

Software

- ORACLE RAC 12C, database software
- STGT, block device export

Test Results

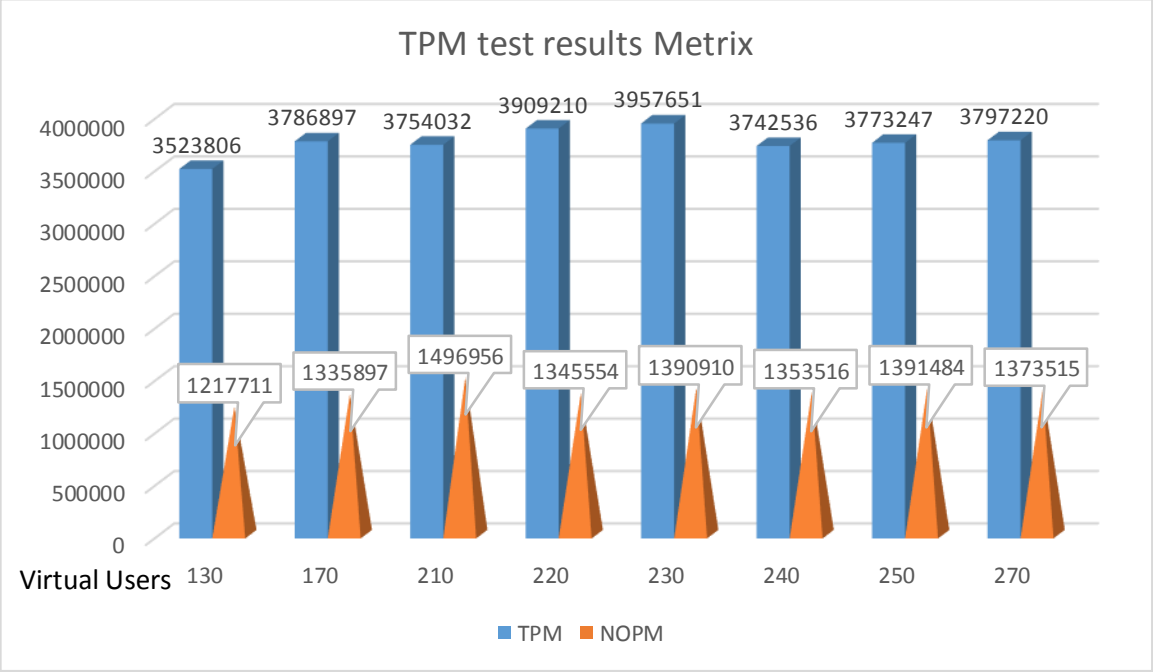

Figure 2, Benchmark Test Results on Different Virtual Users

Figure 2 clearly demonstrates the TPM performance difference of virtual users. The TPM value is the summary of user commits transactions and user rollbacks transactions Per Minute. The NOPM (New Orders Per Minute) is taken from within the schema itself in the district table.

Load Profile

	Per Second	Per Transaction	Per Exec	Per Call
DB Time(s):	70.7	0.0	0.00	0.01
DB CPU(s):	29.4	0.0	0.00	0.00
Redo size (bytes):	93,830,550.3	5,667.5		
Logical read (blocks):	1,955,351.1	118.1		
Block changes:	527,323.3	31.9		
Physical read (blocks):	4,807.9	0.3		
Physical write (blocks):	8,039.1	0.5		
Read IO requests:	4,799.6	0.3		
Write IO requests:	4,713.9	0.3		
Read IO (MB):	37.6	0.0		
Write IO (MB):	62.8	0.0		
Global Cache blocks received:	2,585.6	0.2		
Global Cache blocks served:	2,657.3	0.2		
User calls:	12,687.0	0.8		
Parses (SQL):	7,062.7	0.4		
Hard parses (SQL):	0.3	0.0		
SQL Work Area (MB):	3.0	0.0		
Logons:	0.4	0.0		
Executes (SQL):	341,660.9	20.6		
Rollbacks:	27.8	0.0		
Transactions:	16,555.9			

Figure 3, one instance AWR report with 230 virtual users

As showed in Figure3, the summary TPS(Transactions Per Second) of the single instance is 16583.7. The IOPS second is 9513.5, and the data transferred size is 100.4MB. The most IO request is cached in the memory, so the IO pressure which PBlaze4 SSD encounter is not very high. The CPU utility of database nodes are all full, so the bottle neck located in the CPU number.

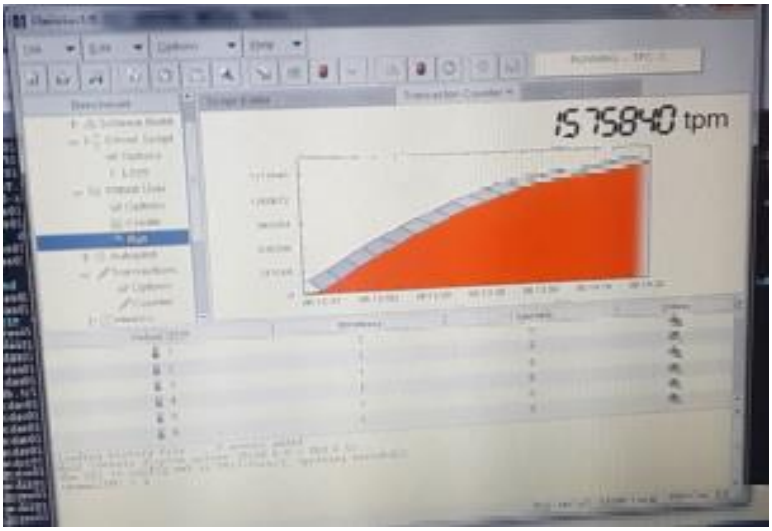

Figure 4, Hammerdb autopilot test results for different virtual users' number

As illustrated in Figure 4, according by the virtual user sequence, the hammerdb run the test. The TPM value draws a smooth curve, and it is almost linear. When there are 230 virtual users concurrently do the transaction task, the peak of TPM was arrived. The TPM in the Figure 3 was the TPCC's transaction Per Minute.

Conclusions

This practice clearly demonstrates how to build a distributed system like the Oracle Exadata machine. According by the Hammerdb test results, although the PBlaze SSD performance did not fully utilized, the entire system's performance and scale out ability are perfect.