Software-Defined Storage from Symantec and Intel: Applications in a Flash

Solution Overview: Storage Management

Summary

Symantec and Intel bring together the flexibility of Software-Defined Storage with the performance of PCIe Flash to provide up to **184%** performance improvements at **5-10%** the cost of an all flash array[1].

Introduction

As myths around lifespan and reliability become a thing of the past and $/capacity continues to decline, the value of flash as a core component to any data center strategy is becoming more evident. At the same time, x86 servers are increasing the slots available for internal storage, CPUs are more dense and powerful, and the internals and network are carrying more data.

This performance and capacity is one of the key drivers of the emerging Software-Defined Storage market. Symantec™ Storage Foundation Cluster File System unlocks the capacity and performance of Intel P3700 Flash devices by creating a Software-Defined infrastructure to provide a fast, flexible environment for mission critical applications. Cluster File System ensures that the data is available when and where the application needs it.

Target Workload

Oracle with Online Transaction Processing (OLTP) workloads are a perfect fit for the randomized workloads that Flash storage excels in, and will be a focus of this paper. In reality, any application can gain benefits by holding its data on fast storage as close to the CPU as possible, removing expensive network transactions. While proprietary engineered solutions exist on the market to solve specific use cases, the architecture presented here can hold any database, application or custom solution.

Oracle is an especially interesting workload for performance testing and architecture validation because it presents mixed read/write patterns that pose challenges for traditional SAN storage. While the database writer processes performs random reads and writes, the data base redo log writer performs purely sequential writes. Oracle focused benchmarks are available that use all flash arrays to achieve over 1.5 million transactions. It is interesting to note however, that in some of those architectures a stripe of several Hard Disk Drives (HDDs) are required to hold redo log data[2]. Our workloads are pure flash.

Architecture

In our testing, single instance Oracle 11gR2 Enterprise instances are used. While Cluster File System also offers options for parallel applications, like Oracle RAC, in this exercise we are going to examine how using Cluster File System's Fast Failover capabilities, a single instance can be made available in the other server very quickly, avoiding costs and complexities of RAC. For those cases where instant recovery is required, RAC is an option, but is not discussed in this architecture.

Our servers are x86 Linux systems using the Intel E5 v3 processors, each with four Intel SSD DC P3700 2.5” PCIe drives with NVMe support.

1. Check Appendix for performance metrics
2. http://c970058.r58.cf2.rackcdn.com/fdr/tpcc/Cisco-Oracle C240 TPC-C FDR.pdf

Fig 1: Base Architecture

The database instances will be made highly available using the Symantec™ Cluster Server powered by Veritas component included with Cluster File System, while internal storage will be shared across the configuration using the Flexible Storage Sharing (FSS) feature of Cluster File System for data protection. Cluster Server and Cluster File System are tightly integrated to provide a single solution for both application and storage management.

Performance

A TPC-C benchmark emulates an OLTP type workload where users are running online transactions. Running two single instances in each of the servers, this configuration was able to achieve close to 1.5 Million transactions per minute.

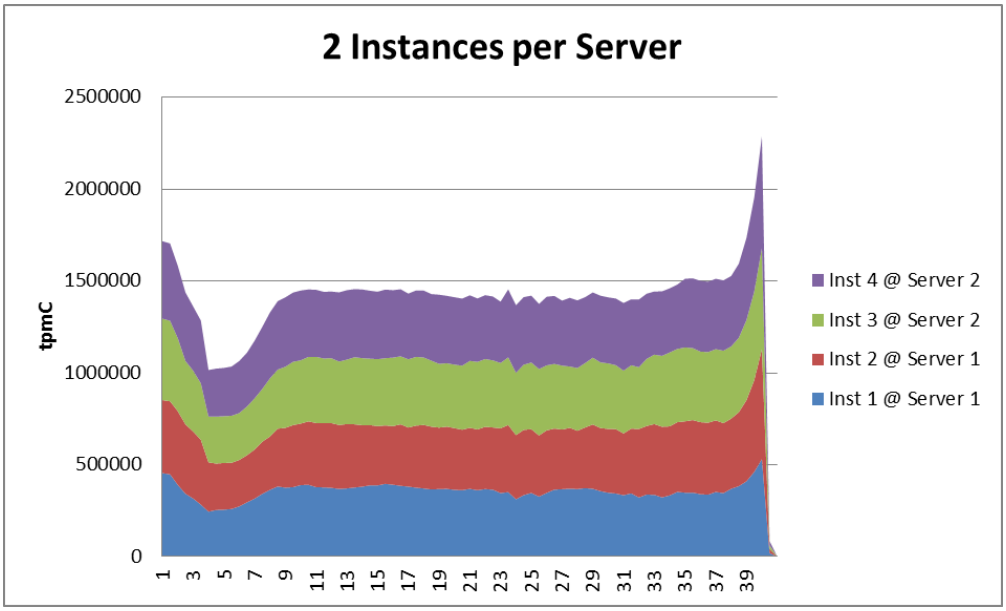

Fig 2: Four Instance Performance

These very high numbers in a 2U configuration show how distributed software can be used to maximize the performance of underline storage structure, with out impacting data availability. With just two servers and internal flash devices, Cluster File System is providing a high available solution with zero Recovery Point Objective (RPO). Any storage failure will be transparent to the database. A server or application failure will utilize fast failover technology to recover the service in the surviving node

very quickly. With over 1.5 million transactions per minute, this environment achieves the same results as published all flash arrays[3] that don't provide high availability and come at a higher cost.

Along with measured transactions, another way to measure performance is through I/O latency. The latencies seen under this configuration for extended periods are between two and four hundred microseconds sustained for steady, predictable performance for both reads and writes. **Figure 3** shows the read and write latencies seen by the data volumes for one instance. In our environment the data volume read latency of 0.37 ms is 60% faster than published EMC XTremIO metrics[4].

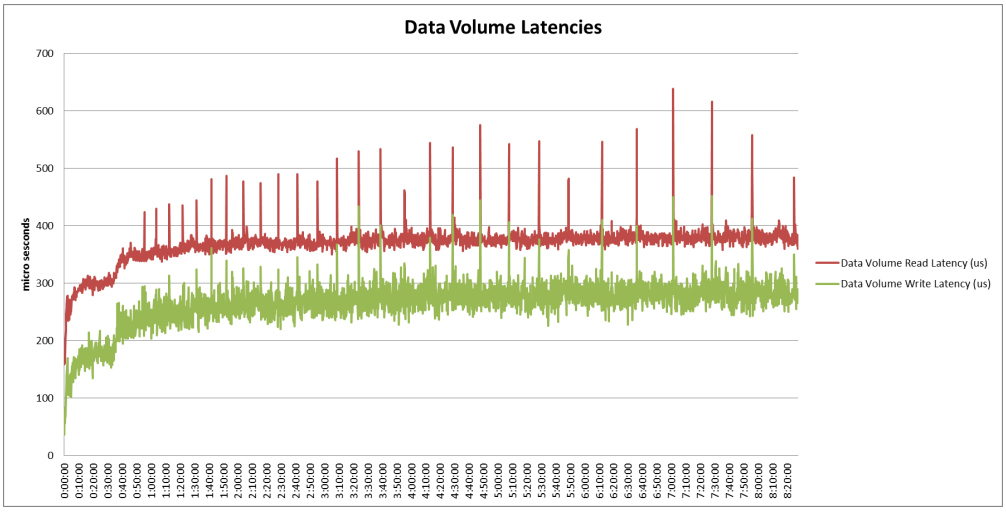

Fig 3: Data Volume Latencies

Application Availability

With the Oracle agent for Cluster Server, any instance failure will be instantly identified by Oracle, triggering the instance recovery very quickly without the needs of a traditional monitor interval. With the Active/Active configuration that Cluster File System brings, the volumes and file systems are accessible by all the nodes at the same time. The time needed to recover the database, therefore, is limited to the time needed to restart and recover the instance plus a minimal detection time.

To demonstrate this capability in this high performance environment, we have tested how much time is needed for one instance to be recovered while the database is running at full workload. To do that, we halted one Oracle process and measure the time needed until the database is recovered in the other node, taking also note on how much data had to be recovered from redo log.

3. https://access.redhat.com/sites/default/files/attachments/red_hat-violin_perf.brief_v5_0.pdf
4. http://xtremio.com/wp-content/uploads/2014/07/h13174-wp-optimized-flash-storage-for-oracle-databases.pdf

	RUN1	RUN2	RUN3	RUN4
Oracle starting in the other node	0:00:18	0:00:17	0:00:16	0:00:21
Database mounted	0:00:07	0:00:01	0:00:06	0:00:06
Crash recovery	0:02:10	0:02:09	0:00:50	0:02:06
Database online	0:00:04	0:00:02	0:00:02	0:00:02
Total Recovery Time	**0:02:39**	**0:02:29**	**0:01:14**	**0:02:35**
Read KB redo	15,970,970	15,984,448	5,654,785	15,854,764
Blocks need Recovery	1,031,287	1,015,626	985,103	1,026,460
Completed redo application (MB)	2,657	2,276	1,033	1,726

Conclusion

As enterprise data centers continue their quests to provide a high level of service for their internal customers and applications, there is the constant underlying goal to do so at reduced complexity and cost. To meet those goals, however, without sacrificing performance or availability has been elusive. Converged architectures based on flash arrays will provide high performance, but at a high cost. Scale-out Software-Defined Storage solutions can utilize Direct Attached Storage (DAS) for reduced cost, but lack built-in data and application management functionality leading to complex point solutions.

Cluster File System provides a full suite of data and application availability tools in a single solution to bring performance, availability, and flexibility to the data center. Combined with unmatched read and write capabilities from Intel Solid State Drive Data Center Family for PCIe, enterprises have the ability to achieve those results at a much lower cost than the alternatives.

Appendix

http://xtremio.com/wp-content/uploads/2014/07/h13174-wp-optimized-flash-storage-for-oracle-databases.pdf

Configuration	XtremeIO [1]	Symantec/Intel	% Improvements
Storage HW	Single X-Brick	**8 x Intel P3700**	**----**
Capacity	10TB	**8TB**	**----**
Fibre Channel Paths	4	**N/A**	**----**
Storage Costs (MSRP)	340,000	**56,656**	**17%**
Oracle Instances	1	**1**	**----**
Highly Available Config	No	**Yes**	**----**
TPmC	----	**608,069**	**----**
IOPs	98,868	**137,852**	**39%**
DB sequential reads latency (ms)	1.07	**0.377**	**184%**
DB parallel read (ms)	----	**0.862**	**----**
Redo Log Throughput (MB/s)	20.8	**75.8**	**264%**
DB writer response time (ms)	----	**0.267**	**----**

https://access.redhat.com/sites/default/files/attachments/red_hat-violin_perf.brief_v5_0.pdf

http://c970058.r58.cf2.rackcdn.com/fdr/tpcc/Cisco-Oracle%20C240%20TPC-C%20FDR.pdf

Configuration	Violin [1]	Cisco [2]	Symantec/Intel
Storage HW	Violin 6616	Violin V6000 VTrack J630sS JBOD	**8 x Intel P3700**
Capacity	8TB	32TB All-Flash + 12TB HDD	**8TB**
System RAM	128GB	768 GB	**256GB**
Fibre Channel Paths	6	16	**N/A**
Storage Costs (MSRP)	540,000	1,098,460	**56,656**
Oracle Instances	4	1	**4**
Highly Available Config	No	Yes	**Yes**
TPmC	1,500,000	1,609,186	**1,427,477**

More Information

Visit our website

http://enterprise.symantec.com

To speak with a Product Specialist in the U.S.

Call toll-free 1 (800) 745 6054

To speak with a Product Specialist outside the U.S.

For specific country offices and contact numbers, please visit our website.

About Symantec

Symantec Corporation (NASDAQ: SYMC) is an information protection expert that helps people, businesses, and governments seeking the freedom to unlock the opportunities technology brings—anytime, anywhere. Founded in April 1982, Symantec, a Fortune 500 company operating one of the largest global data intelligence networks, has provided leading security, backup, and availability solutions for where vital information is stored, accessed, and shared. The company's more than 20,000 employees reside in more than 50 countries. Ninety-nine percent of Fortune 500 companies are Symantec customers. In fiscal 2014, it recorded revenue of $6.7 billion. To learn more go to www.symantec.com or connect with Symantec at: go.symantec.com/socialmedia.

Symantec World Headquarters

350 Ellis St.
Mountain View, CA 94043 USA
+1 (650) 527 8000
1 (800) 721 3934
www.symantec.com

21344089 12/14