

STORAGE SWITZERLAND

ADDRESSING STORAGE CHALLENGES IN COST EFFECTIVE ORACLE ENVIRONMENTS



George Crump, Senior Analyst

Every Oracle powered enterprise wants to run their applications more cost efficiently. The usual goals are to decrease license and maintenance costs while increasing overall performance. At the same time there is a need to push some Oracle applications to new performance heights, those for which the faster they process data and respond to requests the more revenue the organization generates. These objectives are typically accomplished by either running more Oracle instances on larger stand-alone servers or on a Oracle Real Application Cluster (RAC) populated with small, fast servers creating a scale out application grid.

Oracle costs can get out of control when the environment demands too many physical servers running Oracle applications or when an application has to be “partitioned”, or split up across multiple physical servers to meet performance demands. In the first case there can be dozens of application servers that are slightly above average consumers of CPU and storage resources, so much so that they’re not considered for traditional server virtualization projects. The second case is at the other end of the spectrum. These are specific Oracle applications that consume so much of the storage I/O resources that database administrators are looking for ways to spread

them across more servers and storage devices. In both cases there are solutions to address the impact on the CPU resource, but overcoming the storage performance challenges to Oracle consolidation is critical for these projects to be successful.

Other than a mainframe the ideal solution for the first consolidation would be traditional server virtualization, but virtualizing an Oracle application has its own unique challenges. The early stages of server virtualization initially focused on consolidating the dozens, if not hundreds, of servers in the data center that were simply performing a utility function. The next step for IT managers and where the bigger potential payoff lies is virtualizing application servers. Here again the initial focus has been on more lightweight, important, but not mission critical applications. Virtualizing just these business important applications has put a well-documented strain on the storage that supports the virtual infrastructure. The largest payoff in server virtualization is the consolidation of mission critical applications, typically those are Oracle environments that are performance sensitive but not performance critical. The storage challenges are even worse when considering the virtualization of these applications.

At the other end of the spectrum are environments where the database application processing demands have become so high that the application becomes processor and/or storage I/O constrained. These are critical environments in which the performance of the application is directly responsible for the revenue generation potential of the organization. To overcome the performance constraints of the current environment the application is broken up, known as partitioning, or re-written and parallelized to take advantage of Oracle RAC. These two options spread the processing and storage loads across multiple servers and storage systems. In both instances this leads to further Oracle server (applications) sprawl in the enterprise.

The problem with these approaches is that they require expensive per-core licensing and a significant commitment of re-development time to take advantage of the partitioned parallel architecture. If the right type of solid-state storage is not used neither of these workarounds address the disk latency issue. In fact they make it worse by consolidating all the I/O traffic into a single legacy HDD based storage system.

This has led to vendors like Oracle creating “database machines”. Akin to mainframes these systems integrate servers, storage hardware (including disk and Flash SSD) as well as a unique (Infiniband) network interconnects. While an interesting concept these systems have their own set of issues.

The biggest challenge is that this scale out appliance is provided exclusively by a single vendor: Oracle. For the first time since the days of the IBM mainframe we have a vendor trying to offer the entire software and hardware stack. Even IBM doesn't do that anymore. The result is a 'siloed' solution that doesn't fit into the typical data center infrastructure, instead it tries to create a single solution that can address all the various types of applications that Oracle is used for. As was the case in the days of the IBM mainframe, when a single vendor was in control of everything, the price of IT goes up significantly and

innovation is put at risk. Fortunately, innovation is still alive and well and there are alternatives to this scale out, single vendor, and mainframe-like architecture.

The alternative is to move to a scale up instead of a scale out architecture. The logic in scale up is that since the performance capabilities are there, Oracle environments can be made more cost effective while performing better, without having to re-develop the entire application. A scale up architecture supports the current infrastructure and does not require a massive reprogramming effort. A scale up system would also have the advantage of supporting ALL current versions of Oracle and not force an upgrade to 11g, which would be expensive and something the organization may not have the resources for.

Finally, a scale up storage solution would be able to support any OS strategy, not just Linux, and would also support the entire application stack as well as other databases, like Microsoft SQL. It seems reasonable then that scale up solutions would be more affordable and more flexible than a scale out solution. But can a scale up solution out perform a closed, scale out solution?

There are two key performance components to scaling up an Oracle environment; processing power and storage I/O performance. Given the powerful server technology available from companies like HP, a scale up solution for Oracle consolidation and/or performance acceleration is no longer limited to addressing a shortage of processing resources. Server manufacturers like HP are now providing massive multi-core (64 in some cases) architectures that should provide the processing resources required by the most demanding environments. Interestingly these environments provide one of the theoretical advantages of scale out architectures, the ability to granularly add more performance when needed. These server systems are typically fully upgradeable, allowing the purchase of more performance or processing cores as needed. Scale up systems provide scale out flexibility but are less expensive than closed, scale out architectures when all the nodes, storage and especially re-programming is factored in.

The second key to affordable scale up architectures is to combine these powerful servers with storage that can handle the extremely high I/O density and inherently random nature of the environment, without requiring a massive re-development effort. In these environments many vendors have suggested moving to PCIe based architectures to allow for on-host locality of performance. Meaning that the PCIe solid state is directly channeled to the processor, there is no storage network to add latency. The advantage of locality can't be argued with but the challenge is most vendors are using PCIe based cards inside of the host servers. PCIe SSD cards, while fine for light to medium workloads, are space constrained (<1 Terabyte) and because they don't have enough flash cells, they often suffer performance issues when placed under heavy read/write loads. These environments may be better served by Memory Arrays instead of limited use PCIe cards.

Under duress of a high end Oracle workload a flash cell failure may bring down the entire card and potentially the entire node. While vendors may suggest that this is acceptable when leveraging Real Application Cluster (RAC)(there are many servers), the customers have to ask themselves if it really is. Once the customer and the user of the application have come to count on a certain performance profile, any loss in that performance might be considered the equivalent of the application being completely down.

In the scale up environment Memory Arrays can be attached to the single server or shared across many servers as performance and capacity requirements dictate. These systems, like, [Violin Memory's](#), have an architecture that allows for resiliency by populating the system with enough flash memory cells that any single cell failure can be compensated for by using a technique similar to global hot spares. The flash modules can also be replaced while the system is active with no impact to performance. Another key factor is performance predictability. Most flash systems experience a drop off in performance after the device has been completely written to and garbage collection techniques have to be invoked in the same bandwidth/user space as customer data to alleviate the problem. This

cleanup is driven by the flash controller, which allocates some of its processing resources to the pre-clearing of cells so that writes are not delayed. The challenge in a high I/O environment, like Oracle, is that write traffic can become so great that the garbage collection processes gets overwhelmed and flash performance begins to suffer and reads (20us) queue up waiting for flash blocks to be erased (5+ ms), resulting in large latency spikes.

There are two solutions to this problem. First, have plenty of flash memory (oversubscribe) from which garbage collection can work and make sure that there is enough flash controller muscle to churn through the garbage collection task quickly. Memory Arrays like Violin's, while space efficient, are not space constrained like PCIe flash boards. They have plenty of room for sufficient flash memory modules, flash controllers and algorithms (vRAID) to provide consistent performance under load for the life of the storage system. Many PCIe based flash systems do not, they simply don't have the space to perform the task. PCIe flash boards have their place as a memory extension technology, however a high end Oracle consolidation project is not likely one of them. This is echoed in Oracle's own Database Appliance, Exadata, where PCIe Flash Cards are relegated to providing caching mechanism, allowing all writes to be executed at the much slower disk layer.

It is important to understand the business goals for consolidation and performance acceleration. With Oracle Database a scale up solution should be the first consideration, as in most cases it is less expensive and requires less application re-development and infrastructure re-architecting, something that most IT departments are running too thin to even consider.

Starting with a single server with a single storage system tends to be more reliable and easier to maintain. This then becomes the foundation for building a Highly Available Architecture, that may use Data Guard, GoldenGate, Veritas Cluster Services or even Real Application Cluster, these are all solutions that you leverage once the environment is stabilized.

About Storage Switzerland

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: <http://www.storage-switzerland.com>

Copyright © 2011 Storage Switzerland, Inc. - All rights reserved