Oracle Virtualization Best Practices
EMC Oracle Solutions Marketing

• Storage Virtualization Concepts
  – Fibre Channel, NAS, iSCSI
• Provisioning Storage in vSphere
  – Physical RDM, Virtual RDM, and VMFS
• Performance, Consolidation, Ease of Provisioning, Manageability, and Availability

Abstract
This white paper describes best practices for virtualizing Oracle databases using EMC and VMware™ solutions. A board range of topics are covered to provide the reader with guidelines to use in virtualizing Oracle and enhance their current virtualized infrastructure.

May 2014
## Contents

**Executive summary** ................................................................. 4
  Audience .............................................................................................................. 4

**Benefits of Virtualizing Oracle using VMware vSphere** ....................... 4
  Cost Savings .............................................................................................................. 4
  Virtualization Agility ................................................................................................. 6
  High Availability ....................................................................................................... 7
  Support and Certification ........................................................................................... 7
  Customer References ............................................................................................... 8

**Storage Virtualization** ................................................................................. 9
  Dedicated Storage Infrastructure for Databases ......................................................... 10
  Multipathing ............................................................................................................... 10
  Storage Protocols ....................................................................................................... 12
  Provisioning Storage in vSphere ............................................................................... 16

**EMC Best Practices for Virtualizing Oracle** ............................................. 21
  Performance ............................................................................................................... 21
    FAST VP .................................................................................................................. 21
    FAST Cache .............................................................................................................. 23
    Paravirtual SCSI (PVSCSI) Adapters ...................................................................... 24
  Consolidation .............................................................................................................. 24
  Ease of Provisioning .................................................................................................... 25
    vCloud Automation Center ...................................................................................... 25
    vCenter Orchestrator ............................................................................................... 26
  Manageability .............................................................................................................. 27
    OEM 12c Plug-in for EMC VMAX and VNX .............................................................. 27
    EMC’s Oracle Workload Profile Assessments – AWR & Statspack ......................... 29
    DBClassify: A Professional Services Option ............................................................ 30
  Availability .................................................................................................................. 31

**Everything Oracle at EMC Community** .................................................... 33
  Conclusion .................................................................................................................. 33

**References** ........................................................................................................ 34
  Other References ....................................................................................................... 35
Executive summary

EMC proven solutions and technologies offer businesses, IT data centers, DBA teams, and application owners’ great value and seamless integration with virtualization. Success in virtualizing Oracle™ databases and applications is based upon best practices which are established by EMC and our partners. These best practices pave the way for our customers to add more agility, flexibility, resiliency, and performance to their infrastructure. In this paper we review best practices for implementing Oracle on an EMC infrastructure using VMware virtualization.

Audience
This white paper is intended for Oracle database administrators (DBAs), VMware administrators, storage administrators, IT architects, and technical managers responsible for designing, creating, and managing Oracle databases, infrastructure, and data centers.

Benefits of Virtualizing Oracle using VMware vSphere
IT-as-a-Service (ITaaS) is a goal of many organizations as they seek to become more agile in providing services and more competitive in improving their time-to-market. The trend is to develop a virtualized infrastructure to accomplish cost reductions in database maintenance and add greater manageability, flexibility and high availability. To realize these goals IT data centers are designing virtualized infrastructures that offer consolidation of physical resources and automation of routine maintenance tasks.

Cost Savings
Virtualization has been a disruptive technology wave for several years now. However, the conversation has shifted from why virtualize to architecting IT-as-a-Service (ITaaS) and Database-as-a-Service (DBaaS). Consolidation is one of the key tenets of virtualization enabling the data center to use less hardware and experience both operational and capital savings. While the value of consolidation and the cost savings of using less hardware were convincing enough to get businesses to start the virtualization journey, some DBA teams managing mission critical databases pushed back against virtualizing Oracle, questioning, among other things, performance.

Oracle databases represent a significant investment and too many it seemed the cost savings from virtualization didn’t justify introducing another layer to an already stable physical infrastructure. Over time, however, the early adopters transformed into the early majority and are now actively pursuing DBaaS. This transformation is a process that starts with a few applications and grows with the acceptance of virtualization as a stable consolidation platform for mission critical Oracle databases. Another trend called Performance Consolidation has accelerated virtualization as these trends work together to provide the optimal platform for production systems. EMC offers performance solutions like FAST VP and FAST Cache that deliver greater storage performance using fewer hard drives. For example, in a recent paper called, EMC VNX™ Scaling Performance for Oracle 12c RAC on VMware vSphere the multicore FAST Cache using only five Flash drives was able to increase by 93% the number of Transactions per Minute (TPM) and
lower latency 91% for db file sequential read wait events. The combined benefits of virtualization and performance trends have DBA teams virtualizing databases and architecting the optimized database infrastructure.

An article by David Floyer of Wikibon entitled, Virtualization of Oracle Evolves to Best Practices Production Systems that discusses the cost impact of an optimized database infrastructure. In the article the cost of a traditional Oracle database core was compared to the cost of an optimized database core. The analysis of the traditional database core was less expensive ($27,716 per core) as it did not include costs related to virtualization and optimized storage architecture using flash drives. The optimized database core was 17.8% more expensive ($32,671 per core) as the virtualization and flash drives were included in the cost. According to Wikibon, “… the overall cost is (speaking of the cost of the Latency Optimized Database System Deployment) significantly reduced because the number of cores can be reduced from 192 to 120, a decrease of 37.5%.”[15]

The findings of the article illustrate how companies can save costs by blending the virtualization and performance trends into a latency optimized database system which simultaneously uses less hardware and achieves greater performance.

Figure 1: Shows the cost savings of the latency optimized database system deployment compared to the traditional database system. Source: Wikibon, April 2013, Virtualization of Oracle Evolves to Best Practice for Production Systems.
Virtualization Agility
Technical agility is the ability to change the application stack to increase resiliency, efficiency, coordination, and automate common tasks to free time for other growth initiatives. The virtualization feature called live migration enables movement of a virtual machine from one server to another while the application is running and without data loss. VMware vMotion is one of the best examples of how virtualizing Oracle adds agility to mission critical databases. In a recent study by Principled Technologies entitled, *Demonstrating vMotion Capabilities with Oracle RAC on VMware vSphere* made the case for agility by showcasing how fast very large virtualized databases could be vMotioned in parallel without data loss.

Under heavy workloads these virtualized Oracle RAC nodes were migrated, using vMoton, and across all the tests there wasn’t a single ejection or fencing operation. This shows VMware virtualization provides 100% abstraction between the operating system and hardware and is a very mature platform for mission critical applications. The following is the configuration tested in this study:

![Architecture diagram](image)

**Figure 2: Architecture from Principled Technologies study “Demonstrating vMotion Capabilities with Oracle RAC on VMware vSphere”**

The study demonstrated:

- vMotion of a single large virtualized Oracle RAC node took 130 seconds
- vMotion of two large virtualized Oracle RAC nodes took 155 seconds
- vMotion of three (that is all nodes) virtualized Oracle RAC nodes took 180 seconds

According to Principled Technologies, “As we showed in our tests using Cisco UCS server hardware and EMC VMAX™ storage, VMware vSphere with vMotion made it easy to shift large databases and other workloads from one server to another in a cluster, without application downtime. By choosing to run your large databases on VMware vSphere virtual machines, you can reap the benefits of VMware vMotion for the ultimate agility and flexibility in managing your mission-critical database workloads.” Virtualization adds resiliency to the infrastructure and enables proactive and automated hardware maintenance. vMotion is the cornerstone to other technologies such as High Availability (HA) and Distributed Resource Scheduler (DRS) clusters. Together vMotion, HA, and DRS address in large part the components of agility: resiliency, efficiency, coordination, and automation.

**High Availability**
To explore the resiliency side of agility, VMware High Availability (HA) will be discussed. Using VMware the administrator can create HA clusters. Clusters are the logical grouping of services to servers. For example, a virtualization administrator can create a logical cluster consisting of three servers running virtualized Oracle databases. The newly created HA clusters monitors for failures such as the loss of a network card and server and will automatically restart the virtual machines on other hosts in the cluster without manual intervention. Using HA clusters reduces application downtime and automates protection for the virtualized application without modifications to the guest operating system.

Using VMware HA clusters with VPLEX™ Metro provides a high availability solution within or across data centers. VPLEX Metro is a synchronous connection between two sites to federate the storage arrays. In the VPLEX Metro architecture all data is synchronously mirrored across the two storage arrays. With VPLEX Metro, a VMFS datastore can be distributed or “stretched” across multiple storage arrays, and VMware HA can be active across both locations. This solution of using VPLEX Metro with VMware HA clusters is called, “federated HA.” In the event of an unplanned outage, the virtualized Oracle databases automatically restart in the surviving data center without any data loss and no recovery required. The white paper called, [*Using VPLEX Metro with VMware High Availability and Fault Tolerance for Ultimate Availability*](#) provides a discussion on this for DBAs, VMware administrators, and others interested in geographical resiliency and automation.

**Support and Certification**
The question of Oracle's support and certification of virtualization has been addressed elsewhere, but we discuss it here to provide a brief overview and to reference supporting documentation. When discussing the support of Oracle perhaps best to start by asking the question, “What will Oracle support?” The answer is simply that Oracle supports its own products. Only if Oracle’s support team believes the support request is not the related to their product will they direct remediation to another part of the stack. Oracle can redirect the support to the operating system, network cards (and their associated drivers), security software, change control software, HBA (and their associated drivers) and even virtualization including Oracle Virtual Machine. This is not unique to Oracle support, as most companies will remediate issues with their products unless the
cause of the problem is not in their product. Generally, Oracle support has been good to customers who have virtualized Oracle and most of the fear, uncertainty, and doubt are part of sales discussions designed to keep customers from virtualizing. In the technology adoption curve, we are entering the late majority and generally support is not the issue that it has been.

The relevant question concerning certification is very similar to support: “What does Oracle certify?” As many experienced DBAs know, we have to look to the Oracle certification matrix to determine what operating system the database or other product is certified for. With the exception of Oracle’s own hardware, virtualization and partnerships, the certification matrix only uses the operating system in its selection criteria. This means Oracle validates its products with operating systems and does not go further down the stack. This is not new and DBAs have been working with network and storage teams to validate that all the components like NICs and HBAs also work with the certified operating system. The same is true with virtualization as the Oracle DBA will have to work with the vSphere administrators to validate if the target operating system is certified by VMware. Unless you are using Oracle hardware, this might be true of your current database infrastructure, including your NICs, HBAs, and other software, all of which cannot be found on the Oracle certification matrix. EMC, along with many of our partners, have been socializing and working with customers to build consensus on virtualizing Oracle and will help customers as well. Some supporting documentation includes:

- **VMware’s Oracle Support Policy**
- **Understanding Oracle Certification, Support and Licensing for VMware Environments**
- **Virtualizing Oracle with VMware**
  - **Understanding Oracle Certification, Support and Licensing for VMware Environments**
  - **Webcast: House of Brick and GE Discuss Implementing a 100% Oracle Virtualization Strategy**
  - **Read Virtualizing Business Critical Applications with VMware**

In the paper, **Understanding Oracle Certification, Support and Licensing for VMware Environments**, Appendix 3 provides an overview of VMware support for Oracle products running on VMware. To quote, “VMware Support will accept accountability for any Oracle-related issue reported by a customer. By being accountable, VMware Support will drive the issue to resolution regardless of which vendor (VMware, Oracle, or others) is responsible for the resolution.” This means VMware support will work on the customer’s behalf to resolve issues for Oracle software running on VMware. This is called out because the statement shows VMware is standing behind its technology and its customers. This also eliminates the finger pointing or what we sometimes refer to as blame storms in an effort address any impact thought to be because of virtualization.

**Customer References**

Virtualization of Oracle environments using VMware vSphere has become extremely common, if not standard, especially for non-production (for instance, test / dev) environments. A very large percentage of Oracle databases and RAC servers are virtualized using VMware vSphere at this
point. A complete list of the public customer references for EMC’s Oracle customers can be found here. The following is a sampling:

**American Tire Distributors**
American Tire Distributors is one of the largest US suppliers of tires to the replacement tire market, and is experiencing explosive growth. To manage this growth, ATD embarked on an aggressive journey to private cloud. They chose EMC and VMware to help consolidate their Oracle database servers onto a private cloud environment. Using EMC storage they achieved 65% faster back-end processing times and improved database utilization by 97% using VMware on VNX. American Tire Distributors also use Data Domain backup for 2X faster backups and 9X faster restores. Customer Profile, Webcast, and Video Testimonial

**Esprinet**
Esprinet is a leading wholesale distributor of IT and consumer electronics in Italy and Spain. Due to business growth, Esprinet evolved a complex environment, spanning HP, IBM and EMC storage to support different aspects of the business. This complex infrastructure was time-consuming and costly to manage. Performance was poor, and more capacity was needed. Esprinet decided to consolidate most of its IBM and all of its HP storage onto EMC VNX, in the process implementing EMC FAST Suite, including the Oracle infrastructure which is virtualized using VMware vSphere. With FAST Cache, the company meets 80% of its I/O demand from Flash drives while using 33% fewer drives. Customer Profile

**Storage Virtualization**
Traditionally, production databases and mission critical applications are among the last to be virtualized as their importance to the business demands a transition without disruption or risk. VMware, founded in 1998, is a very mature virtualization platform, and has proven to be an enterprise technology platform that leads the cloud transformation of Tier-1 production DB applications for over 55% in 2013 and projected to reach 75% in 2016.[1] Rarely are application owners asking, why should they virtualize; rather, the focus is on “how can automated application provisioning be achieved” commonly referred to as IT-as-a-Service (ITaaS) or Databases-as-a-Service (DBaaS) for databases.

One of the key considerations in virtualizing Oracle database is an understanding of the virtualization layers. The bottom most layer of the virtualization model is storage and for Oracle databases storage is one of the most important aspects of designing the cloud infrastructure for your applications. Oracle DBAs are highly concerned about: storage performance, reliability, manageability and protection. VMware Administrators work closely with the storage administrators and DBAs to align the selection of physical Raw Device Map (pRDM), virtual Raw Device Map (vRDM), Network Attached Storage (NAS), Internet Small Computer System Interface (iSCSI), and Virtual Machine File System (VMFS) with management and performance requirements of the application. The design of the storage architecture as it relates to virtualization and most importantly the database is a key determining factor for successfully virtualizing mission critical
applications. Exploring storage protocols will give us a starting point in designing our virtual storage architecture.

**Dedicated Storage Infrastructure for Databases**
Production databases and applications should have a dedicated storage infrastructure across HBAs, switch ports, cables, and disks. While virtualization is very good at consolidation of server resources and driving increased utilization, the same consolidation practices should not apply to the storage design supporting your productions databases. Dedicated storage achieves predictable performance, speeds performance remediation efforts, and eases storage management.

**Multipathing**
To use multipathing, two or more HBAs are required, in addition to a storage array which supports multipathing. For example, beginning with Enginuity 5876 and Solutions Enable v7.4 for the Symmtrix™ VMAX arrays support Native Multipathing (NMP). It is recommended that the storage and virtualization administrators collaborate on which of these three types of multipathing are supported and best for their Oracle databases:

- **Active/Active:** I/O requests can be sent to a LUN via any storage array process or port. The new generation of VNX storage arrays have Symmetric Active/Active path management for optimal multipathing. Traditionally, the ownership of a LUN by a storage processor meant Pseudo Active/Active was used, meaning there was an optimal and non-optimal path down to storage which could cause thrashing (when data is sent down the non-optimal path causing a trespass and must be transferred to the owning SP before a commit can occur). Heavy trespassing is called thrashing and is comparatively not as good as true Active/Active multipathing. Symmetric Active/Active in the new VNX improves multipathing by managing only the metadata between storage processors. It works like this: The SP communicates over a management interface to request a lock on a Logical Block Access (LBA). This new LBA locking mechanism means only metadata flows between processors and the owning SP can maintain write-order fidelity which is extremely important for Oracle databases.

- **Pseudo Active/Active:** also commonly called Asymmetric Logical Unit Access (ALUA). As described above data can flow down any path however, if a non-optimal path is used this leads to trespassing. This means data flowing down the optimal path is much faster than data that is not.

- **Active/Passive:** data flows down one active path and in the event of a path failure data the passive path takes over.

Now that we briefly reviewed the three types of multipathing there is the matter of overlaying the Path Selection Policy (PSP). VMware enables the virtual administrator to choose from these three PSP polices:

- **Fixed:** Sends data down the optimized path and is commonly used for Pseudo Active/Active or ALUA arrays
- Most Recently Used (MRU): commonly used with Active/Passive because it sends data down one path until a failure and the passive path takes over.
- Round Robin (RR): Sends a fixed amount of data, I/O, or bytes down one path then does the same for the next path. Normally, favored for true Active/Active multipathing.

Learning what multipathing your array supports and matching the path selection policy will assist in validating the configuration for your virtualized database. Beyond the native path management in VMware there is EMC PowerPath Virtual Edition (VE) that extends performance with features such as dynamic load balancing, ease of management and failover and recovery capabilities. PowerPath VE resides in the ESXi hypervisor, below the virtual machines and above the storage, and because of its position in the stack PowerPath VE is a one–time server installation that benefits heterogeneous VMs. Oracle databases and production applications normally demand the best resiliency and recovery options to prevent loss of service making these two PowerPath VE features key considerations:

- Automatic failover and recovery: If a path error occurs PowerPath automatically redirects data to the best surviving path and then rebalances I/O across all remaining paths. No manual intervention is required in the case of a failure and when the failed path is available PowerPath automatically adds the path to the pool and again rebalances I/O.

- Intelligent load balancing: statistics and algorithms are used by PowerPath to analyze paths and determine the optimal path. This means PowerPath detects bottlenecks and dynamically redirects I/O to maintain the level of performance expected by Oracle databases.

The following graphic shows the use of PowerPath VE:
In the ESG Lab Validation report entitled, Automated Path Optimization for VMware Virtual Environment both the automation failover and recovery and the performance benefits of PowerPath VE were validated. For example, PowerPath VE achieved 25% more IOPS in an OLTP workload when compared to Native Multipathing. The resiliency and performance benefits of PowerPath VE are key considerations for architecting your Oracle database storage architecture.

**Storage Protocols**
There are several storage protocols that storage and database administrators can choose for their applications and each has their own strengths and challenges. iSCSI is a protocol that uses TCP to transport SCSI commands over the network to connect to block devices in storage. Network Attached Storage (NAS) is a protocol that uses TCP to present file services on the storage array that operates as a file server. Finally and perhaps most commonly used with Oracle databases is Fibre Channel a protocol for direct connection between the host and the enterprise storage.

**Fibre Channel**
A host server connects to the Storage Area Network (SAN) typically using two Host Bus Adapters (HBA) to a high performance storage system. The Fibre Channel (FC) protocol is used to transfer SCSI commands, enabling the server to attach to the underlying storage. Components of the SAN architecture include:

- **HBAs**: PCIe cards in the server that offload the FC protocol from the host to connect via switches to the storage array
- **Switches**: assist with routing traffic to and from storage array and are used to restrict server access to storage which is called, “zoning.” Zones are created for a group of
servers that need to access the same shared storage. For example, Oracle Real Application Clusters (RAC) is a group of servers that need to access the same database files which can be accomplished by creating a zone. Thus, zoning enables servers to share the same storage and prevents servers from accessing other storage.

- LUN Masking: is used by the storage administrator to configure servers to access LUNs or conversely not to see LUNs.
- Cables, storage processors, and mechanical and solid-state disk devices

Using two HBAs provides the opportunity to enable multipathing, a technique in which multiple paths are used between the host and storage, to increase bandwidth and scale application performance. Oracle databases and other enterprise applications are performance dependent upon the fast transfer of information, thus making the use of multipathing a key consideration for the storage architecture.

**Fibre Channel Guidelines**

In this section we will review in order of priority, top being the most important, the relevant guidelines to consider when architecting Oracle database storage.

- Use the fastest supported HBA cards and ensure speeds are consistently the same across the cards. Reference VMware Knowledge Base (KB) article 1006602.
- Check with the HBA card vendor for the recommended queue depth setting for the HBA cards. Reference VMware KB article 2072070 called, *EMC VMAX and DMX Symmetrix Storage Array recommendation for optimal performance on VMware ESXi/ESX*.
- For Oracle RAC research Change Maximum Outstanding Disk Requests for Virtual Machines. Reference [VMware KB article 1268](https://kb.vmware.com/kb/1268) and ESXi and vCenter Server 5 Documentation.

**Network Attached Storage**

NAS is file-level data storage accessed over a network to a gateway on a computer or enterprise systems like EMC's VNX series. One of the strengths of using NAS is it enables heterogeneous group of clients to access the same files on storage. Enterprise systems like the VNX allow Microsoft Windows, Linux, and UNIX clients to share files in multiprotocol Network File System (NFS) and Common Internet File System (CIFS) environments. Recently, NAS has become increasing popular for some of these reasons:

- Simplified storage design: Choosing NFS reduces complexity as the business does not have to explore options like pRDM, vRDM, and VMFS. Additionally, using enterprise storage means all array based features like snaps, clones, replication, and acceleration technologies like auto-tiering and dynamic flash caching are available.
• **Fast Ethernet**: The rapid development of fast Ethernet means customers will enjoy the performance offered by 40GbE or even 100GbE. With broad adoption of fast Ethernet prices will drop making the cost per gigabyte very competitive.

• **Direct NFS (dNFS) from Oracle**: implementing dNFS is simple and because the NFS connection is managed by the database binaries database performance is very good.

In the proven solution entitled, [EMC VNX Scaling Performance for Oracle 12c RAC on VMware vSphere 5.5](#) the storage and RAC cluster interconnect networks used 10 GbE and the Oracle database dNFS. Performance testing in the solution showed a four RAC node architecture was able to sustain 239,000+ Transactions Per Minute (TPM) without FAST Cache and 744,000+ TPM with FAST Cache (Fully Automated Storage Tiering (FAST) uses Solid State Disks (SDD), commonly referred to as flash drives, to dynamically cache very active data). The checklist below has some of the best practices for using NAS with virtualized Oracle databases.

### NAS Guidelines
The below has some of the best practices for using NAS with virtualized Oracle databases.

• **Isolate NFS traffic**: With network storage solutions like NFS a major consideration is predictable performance and this can be achieved by dedicating the production network to your production servers. This means isolating the server traffic using dedicated switches or Virtual LANs (VLAN).

• **Fast Ethernet**: Use the fastest supported Ethernet for both VMware vSphere and EMC storage.

• **Jumbo frames**: Implementing jumbo frames means changing the default MTU size of 1500 to a MTU size of 9000. A larger MTU increases throughput, generates fewer network packages of greater size, and could lower server CPU utilization. These layers have to be configured to support jumbo frames: Virtual Machine, Virtual Distributed Switch (VDS), physical switch and the VNX data mover.

• **Implement Oracle dNFS**

Oracle has optimized its dNFS client to increase throughput and overall performance by bypassing the operating system and using asynchronous I/O. The dNFS client uses Direct I/O to improve performance by eliminating OS write-ordering locks, bypassing OS caches and reducing system CPU utilization. Asynchronous I/O enables the processing of reads and writes to continue without waits normally associated with write-ordering locks.

Other benefits include:

• **High Availability**: dNFS supports up to four parallel network paths to the NAS system and automatically load balances I/O across all available network paths. In the case of a network failure any lost packets will automatically be resent over surviving network paths.

• **Ease of Administration**: dNFS enables the Oracle DBA administrator to setup the client for both single instance and Oracle RAC with little effort.

Other NAS recommendations:
If using more than eight NFS volumes review the VMware TCP/IP network heap minimum and maximum sizes. Definition of advanced NFS options can be found in VMware KB article 1007909.

iSCSI
Internet Small Computer System Interface (iSCSI) protocol uses TCP to transport SCSI packets to connect to block devices on the network. Benefits of iSCSI includes the fact that it uses the existing network and uses lower cost hardware (network cards) making iSCSI comparatively less expensive than other protocol solutions. Adding additional bandwidth with iSCSI is easily achieved by using port aggregation and bonding links to scale with growing workloads. There are two implementations of iSCSI the business can choose from:

- **Software iSCSI**: The processing of iSCSI traffic over TCP/IP is performed by the servers CPUs. Server processing of iSCSI traffic can demand significant processor utilization especially when network paths are oversubscribed. The behavior of TCP/IP on oversubscribed network paths is to drop packets and have the packets resent thus driving additional server processor utilization.

- **Hardware assisted iSCSI**: TCP/IP Offload Engine or TOE is a chip in the network adapter that processes the iSCSI traffic. The offloading of the TCP/IP from the server processors to the TOE chip in the network adapter is the recommended path for supporting iSCSI.

**iSCSI Guidelines**
- **Isolate iSCSI traffic**: As mentioned above oversubscription of network paths will cause dropped TCP/IP packets leading to congestion and impacting performance. Designing a dedicated network for the iSCSI traffic assists in reducing if not eliminating the impact of non-production network traffic that could cause congestion.

- **Fast Ethernet**: Use the fastest supported Ethernet for both VMware vSphere and EMC storage.

- **Jumbo frames**: Implementing jumbo frames means changing the default MTU size of 1500 to a MTU size of 9000. A larger MTU increases throughput, generates fewer network packages of greater size, and could lower server CPU utilization. The follow layers have to be configured to support jumbo frames: Virtual Machine, Virtual Distributed Switch (VDS), physical switch and the VNX data mover.

- **Use round robin path policy** – Both the VMAX and next generation VNX storage arrays support true active/active NMP for use with the round robin path select policy. The prior generation of VNX storage arrays support pseudo Active/Active (ALUA) meaning a fixed path selection policy applies: I/O will go down the active optimized path.

- **VNX OE for Block application tuning using VMware ESX Server with iSCSI Datastore review VMware KB article 1002598 entitled, ESX/ESXi hosts might experience read or write performance issues with certain storage arrays.**

[14]
Provisioning Storage in vSphere
VMware provides great flexibility in architecting the storage. Common methods for storage provisioning include VMFS and RDM. Each storage method has its strengths and selection of VMFS and RDM can be complex. Many customers start with reviewing how storage is used today. For example, customers might review procedures for provisioning databases using storage based snaps and clones. Having a good picture of how DBA teams use storage features today will assist in deciding to go with RDM or VMFS. Historically, RDM was slightly faster than VMFS, but today there is little to no difference in performance between the two. Below we review VMFS and RDM and provide guidelines we hope will help in selecting the best storage type for your business.

vSphere 5.5 Storage Maximums
vSphere storage maximums are important to designing the storage infrastructure and a common topic of discussion with customers. Included are some storage maximums that may apply to virtualizing database storage. Please review the paper, *VMware Configuration Maximums vSphere 5.5* for a complete list of maximums supported by vSphere 5.5.

- iSCSI LUNs per server: 256
- Fibre Channel LUNs per server: 256
- NFS mounts per server: 256
- VMFS Volumes per host: 256
- Virtual Disks per host: 2048

VMFS
VMFS is VMware’s implementation of a high-performance clustered file system optimized for virtual machines. By clustered, VMware means that multiple VMs can read and write to the same VMFS datastore making storage consolidation and management very easy. VMFS works on any SCSI-based protocol including Fibre Channel, Fibre Channel over Ethernet, and iSCSI. Choosing VMFS gives the Oracle DBAs access to features like Distributed Resource Schedule (DRS), High Availability (HA), vMotion, and Storage vMotion. It is even possible to use storage based snapshots to clone VMFS datastores as shown in the EMC proven solution, *Upgrade to Oracle Database 12c with Oracle Multitenant Option (Pluggable Database)*. In this proven solution step-by-step instructions take the reader through how the clone the VMFS datastore using VMAX TimeFinder™ and use auto-provisioning groups to configure ESXi access to the target host. Recommended reading for companies using VMFS with their Oracle databases.

Below are considerations for determining if VMFS is right for virtualizing your Oracle databases:

- Storage consolidation: VMFS volumes can host one or many virtual machines however Oracle databases tend to be among the most demanding of storage when compared to other applications and the application users also expect high service
levels so storage consolidation is less of a benefit with regard to databases. Oracle DBAs should collaborate closely with the storage and VMware administrators to validate that the VMFS storage layout has been architected to deliver expected performance and application SLAs for the business. Generally, production should have a dedicated VMFS datastore for predictable performance and less critical databases like those in development are candidates for greater consolidation.

- Ease of administration: Generally, VMware administrators find managing VMFS datastores easier than RDMs. For example, adding a VM to a VMFS datastore is an easy administrative task that can be completed very quickly.

- Support for disabling simultaneous write protection: VMware KB article 1034165 entitled, Disabling simultaneous write protection provided by VMFS using the multi-writer flag has a very good technical overview of when to use this feature and is recommended reading particularly with implementing Oracle Real Application Clusters (RAC). By default multiple VMs in the same datastore cannot write the same vmdk file as this could cause data corruption. For clustering solutions like Oracle RAC that maintain write consistency the recommendation is to disable simultaneous write protection. There is a maximum of eight physical servers supported for disabling simultaneous write protection. There are some caveats that come with disabling simultaneous write protection as the VMware features below are unsupported:
  - Most snapshots and other action which utilize snapshots[5]
    - Cloning a virtual machine one or more disks configured with the multi-write flag[5]
    - Storage vMotion[5]
    - Change Block Tracking (CBT)[5]
    - Suspending a virtual Machine[5]
    - Hot-extending a virtual disk[5]

- Oracle RAC Node Live Migration: Using VMFS means the DBA can use vMotion to non-disruptively migrate the VM from one server to another. A recent paper by Principled Technologies entitled, Demonstrating vMotion Capabilities with Oracle RAC on VMware vSphere is highly recommended reading for DBAs interested in third party validation proving that vMotion of heavily utilized RAC nodes will not result in data loss and can be done very quickly. In the study, vMotion of all three heavily utilized RAC nodes took only 180 seconds to complete with minor impact to database performance.

- Flash Read Cache (vFRC): A new feature in vSphere 5.5 is Flash Read Cache which enables managing flash storage located in the server. For example, EMC's XtremSF cards are a server-based storage solution to accelerate workloads like Oracle databases and work with vFRC. At the time of this writing, the VMware Compatibility Guide shows five EMC XtremSF cards which are compatible with vFRC on ESXi 5.5 U1 including, the 1.4 TB PCI SSD card. In the recent study by Principled Technologies
entitled, *VMware vSphere 5.5 with vSphere Flash Read Cache: Performance with Oracle Database 12c* a very demanding Online Analytical Processing or OLAP workload was used with Oracle 12c, Cisco UCS server blades and PCIe SSD card. Results in the paper showed a 14% performance improvement in the OLAP workload using vFRC and after a vMotion the vFRC cache warmed up rather quickly.

**VMFS Guidelines**

Below are considerations for implementing VMFS with Oracle:

- Architect the VMFS datastore for performance, not consolidation.
- Virtual Storage Integrator (VSI) is a VMware vCenter plug-in that simplifies management of virtualized storage by enabling the ability to map virtual machines to storage and to self-provision storage. EMC provides VSI plug-in for free.
- Distribute your LUNs/disks evenly and use a free tool like the Oracle Workload Profile Assessment to assist in designing the storage layout.

**Physical RDM**

Raw Device Mapping (RDM) is a method VMware uses to connect, using a mapping file, the Virtual Machine to the physical storage device. If you are familiar with Unix or Linux, then think of the mapping file as a symbolic link to a file that provides address resolution to the Logical Unit Number (LUN). The mapping file nicely allows user-friendly name for a mapped device. For example, it is not necessary to refer to the device by its device name but instead the mapping file name can be used, for example:

```
/vmfs/oracle/oradata/finprod01/finProdDisk.vmdk
```

In physical compatibility mode minimal SCSI virtualization is used. The VMkernel passes all SCSI commands to the device except for the REPORT LUNs. The REPORT LUNs command is virtualized because the VMkernel needs to isolate the LUN to the Virtual Machine. Save for the REPORT LUN command, all SCSI commands are passed directly to the device making the use of pRDMs a strong consideration for IT departments that rely upon storage activities like clones, snapshots, replication, and backups. It is common for DBAs to use storage features to quickly provision copies of production databases and for resiliency. Changing the storage procedures could represent a significant investment in time, and thus the use of pRDM enables both virtualization and the retention of procedures based upon storage features.

Below are considerations for determining if pRDMs is right for virtualizing your Oracle databases:

- pRDM enables IT continued use of storage array features like clones, snapshots, replication and backups without the need for modification.
• pRDM enables the virtual-to-physical and physical-to-virtual migrations. This means an Oracle database very quickly with minimal effort can move to VMware virtualization in a physical-to-virtual migration using pRDM. Conversely, if required the reverse is true, the Oracle database can very quickly with minimal effort move from VMware virtualization to physical by disabling the use of the mapping file(s).

• pRDM can be migrated to vRDM: See VMware Knowledge Base article 1006599 entitled, “Switching a raw data mapping between physical and virtual compatibility modes in ESX/ESXi.”

• pRDM can be migrated to VMDK via cold migration if the goal is to eventually have the database in a VMFS datastore.

• vMotion requires consistent LUN IDs for all RDMs across all participating ESXi hosts or vMotion will not be available.\[6\]

• VMware High Availability is supported with pRDMs.

The benefits below apply to physical compatibility mode and virtual compatibility mode:

• The 1-to-1 mapping between virtual machine and LUN means IOPS are not impacted by other virtual machines.

• User-Friendly Persistent names: As mentioned the mapping file enables the use of a friendly name for the device name.

• Dynamic Name Resolution: The physical configuration of the storage devices can change but these changes are transparent as the VMFS files automatically resolves mapping to the SCSI devices.

• Distributed File Locking: Oracle RAC and other clustering solutions require two or more virtual machines to access the same LUNs. Distributed file locking on a raw device mapping makes it safe for the two or more virtual machines to access the same LUN.

Physical compatibility mode limitations:

• Using VMware snapshots is not supported however, storage array based snapshots and clones will work.

• Flash Read Cache (vFRC) is not supported but XtremCache will work.

**Virtual RDM**

In virtual compatibility mode, the VMkernel virtualizes nearly all the SCSI commands sending only READ and WRITE commands to the mapped device.\[6\] Because most of the SCSI commands have been virtualized, masking the hardware characteristics of the LUN, the mapped device is exactly the same as a virtual disk file in a VMFS volume. Virtual compatibility mode offers nearly all the
same benefits of VMFS. Below we highlight just a few benefits of vRDM to show the differences between vRDM and pRDM:

- vMotion of Oracle databases and RAC nodes (requires disable simultaneous wire protection flag to be set).
- VMware snapshots are supported.
- vRDM migration (virtual-to-physical) has conflicting documentation. For example, in this VMware vSphere Storage discussion entitled, “pRDM vs vRDM virtual / physical interchangeability” customers have tested and found that moving from vRDMs to physical hosts does work. The recommendation is to log a support ticket with VMware to determine if it is supported.
- Flash Read Cache (vFRC) is supported.

In comparing vRDM to pRDM the major decision point is whether to retain storage array procedures for managing databases or move towards virtualization based storage management. As can be seen in the table below both vRDM and pRDM offer snapshot capabilities the difference is with vRDM it’s done using vSphere and with pRDM through the storage array. Performance considerations follow a similar path as a virtualization team would use vFRC with vRDM and the storage team would use XtremCache with pRDM. Generally, selection of vRDM, pRDM, and VMFS is dependent on which is the best for the business and DBA team as there isn’t significant loss of functionality in choosing any one of the three.

<table>
<thead>
<tr>
<th></th>
<th>vRDM</th>
<th>pRDM</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCSI Commands Passed Through</td>
<td>NO</td>
<td>YES (Notes)</td>
<td>Except for REPORT LUNS</td>
</tr>
<tr>
<td>Disable Simultaneous Write Protection Required for Oracle RAC</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>vCenter Support</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Distributed Locking</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Storage Array Snaps, Clones Replication</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Virtual-to-Physical Interchangeability</td>
<td>YES (Notes)</td>
<td>YES</td>
<td>Customer validated, no supporting VMware KB article</td>
</tr>
<tr>
<td>VMware Snapshots Supported</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>VMware Flash Read Cache Support</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>XtremCache Support</td>
<td>NO (Notes)</td>
<td>YES</td>
<td>vRDM implementation should use vFRC</td>
</tr>
<tr>
<td>vMotion Supported</td>
<td>YES</td>
<td>YES</td>
<td>Requires consistent LUN IDs for all RDMs across all participating ESXi hosts</td>
</tr>
<tr>
<td>VMware High Availability</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Comparing virtual compatibility mode (vRDM) to physical compatibility mode (pRDM)
EMC Best Practices for Virtualizing Oracle

This section briefly reviews some of the key solutions from EMC that benefit virtualized Oracle performance, consolidation, ease of provisioning, manageability, and availability.

Performance

Virtualizing Oracle databases improves resiliency, manageability, and consolidation, but performance is primarily driven by technologies that accelerate access to your data. In a recent IOUG study, Efficiency Isn’t Enough: Data Center Lead the Drive To Innovation, the authors commented, “What types of database management activities are taking up most of respondents’ time each week? The activities on which data managers spend the most time are performance tuning and diagnosis, following by maintaining uptime and availability.” While virtualizing databases, one of the primary concerns of DBAs is adding to the time spent on tuning performance and diagnosis. This need not be the case if the DBA collaborates closely with the VMware and storage administrators to architect for performance. EMC has several proven technologies to accelerate databases that are totally transparent to the DBA, and at the very least, means no additional management overhead and at the very best, may significantly minimize the need for further performance tuning. The best place to start is with a proven technology that has been around for years, and depending upon the EMC storage you are using, Fully Automated Storage Tiering Virtual Pools (FAST VP) may be available to you today.

FAST VP

Prior to automated tiering, storage arrays were typically configured as a single storage tier. For example, a storage array might have all Fibre Channel 10K RPM (Revolutions Per Minute) hard drives. Performance was largely determined by the number of hard drives and the RAID (redundant array of independent disks) type supporting the database(s). Architecting storage was a one-time event, and the DBA had to drive as much performance as possible from this deterministic storage model. Challenges included having to manually manage datafiles, resolve data skewing issues and tuning poor SQL to improve database performance. Fortunately, FAST VP turns the complexity of manually managing the database on storage to a fully automated solution that identifies the very active data and moves it to the fastest specified tier of storage.

In today’s storage arrays, Solid State Drives (SSD) also called flash drives and Enterprise Flash Drives (EFD), are the fastest tier of drives. Just below EFD drives are all drives with a rotational speed of 10,000 and 15,000 RPM and are categorized as FC and can be combined into a single tier. Similarly, all drives with a rotational speed of 7,200 RPM, or lower, are categorized as SATA. The term storage tier means a group of similar resources (EFD, FC, or SATA) with the same characteristics like RAID protection type (RAID 1, 5, 6, or unprotected). A VMAX FAST VP storage tier is used for moving data, referred to as relocation granularity, in sizes of 768 KB or more commonly 7,680 KB (or 7.5 MB) between the VP tiers. In the new generation of VNX OE storage arrays, the relocation granularity has been reduced from 1 GB to 256 MB so FAST VP is more efficient and reduces storage space overhead. There is a great deal of flexibility in configuring FAST VP with options like FAST policies enabling the storage administrator to set usage rules that provide guidelines for the placement of data. Many of the proven solution papers on the Everything Oracle at EMC community and the Technical Notes for FAST VP on the Symmetrix VMAX...
and VNX are recommended reading as they go into great detail on installation, function, and configuration of FAST VP.

Below are some of the guidelines for using FAST VP with Oracle databases but unlike other recommendations we strongly suggest you work closely with your EMC Engineer to validate the configuration. Each database has a unique I/O pattern and workload profile so these guidelines are meant to be the starting point of collaborating with an EMC Engineer to customize the FAST VP configuration.

- **Use the Oracle Workload Profile Assessment (OWPA):** The OWPA is a tool used by EMC Engineers to analyze database AWR reports and is free to customers. This evidence-based approach takes statistics from AWR reports and assists with determining the number and types of drives to use in FAST VP pools.

- **Use DBClassify:** DBClassify is a service for DBA and Storage Administrators in which an EMC Engineer installs DBClassify and then collaborates with the two groups to show how FAST VP will improve performance though evidence based approach. The delivery includes detailed recommendations for optimizing database and storage infrastructure, accelerating the adoption of intelligent storage tiering into production databases, and training for the team for the continued use of DBClassify.

- **Tier Advisor:** Tier Advisor is a storage feature that will assist in determining how many of each type of drive you may need.[11]

- **Across most of our Oracle proven solutions, we use RAID 5 and RAID 10:**
  - RAID 5 has good write performance and good resiliency with 12.5% capacity overhead. In our papers we use RAID 5 for data files, CDB (applies to Oracle 12c), temp files, and FRA files.
  - RAID 10 has the best write performance and better resiliency than RAID 5 with a 50% capacity overhead. In our papers we use RAID 10 for Redo logs, Archive logs, and CRS files.
  - RAID 6 is not recommended due to the write overhead except for use with the SATA and NL-SAS tier where the majority of database data may reside and the need for double disk protection is deemed necessary. Consider placing FRA, archive logs, and RMAN backups on RAID 6 storage pools.

- **VMAX general best practices for FAST VP:**
  - RAID recommendations change based on the VMAX model. For example, the VMAX 10K (959) and 20K the optimum number of EFDs is 8 per engine.[11]
  - For the VMAX 10K (987) or 40K VMAX the optimum number of EFDs is 16 per engine.[11]

- **VNX general best practices for FAST VP:**
  - RAID 5 in a 4 +1 configuration is optimal for performance and is used in many Oracle proven solutions.[12]
RAID 10 in a 4 + 4 configuration is optimal for performance and in our papers we have also used RAID 10 in a 2 + 2 configuration.\[12\]

FAST Cache
FAST Cache is unique to the EMC VNX series of storage arrays. For the purpose of this review, we will focus on the multicore FAST Cache in the new generation of VNX storage arrays, but for those interested in the difference between the original implementation and the performance benefits of the new multi-core FAST Cache, we recommend reading the blogs, “Part 1: Next Generation VNX – Oracle Database Storage Performance” and “EMC Multicore FAST Cache A Detailed Review”.\[13\] The way Multicore FAST Cache works is by positioning flash drives between the storage processor’s DRAM primary cache and the hard drives. In bound I/O requests are serviced by the primary DRAM cache, called a read hit, for the fastest possible performance. In case the DRAM cache cannot service the I/O request, called a read miss, the multicore FAST Cache will then service the request using the Flash drives. In the case of a read miss, the I/O request will be satisfied by the hard drives, and if frequently used, then promoted to both the DRAM cache and FAST Cache performance tiers. FAST VP and FAST Cache are complementary and work together to benefit the entire storage array. Generally, FAST Cache should be positioned for applications with random I/O that bursts frequently as most of the I/O read and write requests will be serviced by FAST Cache, thereby improving performance and reducing workload on the back-end hard drives.

Some of the attributes of FAST Cache that make it ideal for databases are its relocation granularity of 64 KB, real-time monitoring decides which data needs to be promoted, and that there are no relocation cycles attached to promoting the hot data to multicore FAST Cache. IO calls greater than 128 KB exceed the granularity boundary and are not promoted to FAST Cache. FAST VP on the VNX moves hot data, using relocation granularity of 256 MB, to flash drives providing another layer of performance before having to service requests from the slower high capacity hard drives. In summary, use FAST Cache for dynamically changing data patterns with frequent I/O bursts and FAST VP for optimizing disk utilization and efficiency. EMC recommends that if you have a limited number of Flash drives and an option to use them either for FAST VP or Multicore FAST Cache, that you use FAST Cache Optimized drives to create Multicore FAST Cache.\[13\]

Below are some of the guidelines for using FAST Cache with Oracle databases, but unlike other recommendations, we strongly suggest you work closely with your EMC Engineer to validate the configuration.

- VNX storage arrays use Service Processors (SP) and I/O should be balanced across the SP for optimal performance.
- Spread Flash drives across all available buses.
- Use the Oracle Workload Profile Assessment for sizing of FAST Cache for your Oracle databases.
- If the active dataset size is unknown, size FAST Cache to be 5 percent of your capacity and adjust after observation.\[14\]
- Avoid enabling FAST Cache for these workloads:
The workload is primary sequential.
- Oracle Redo Log files.
- Oracle Archive Log files.
- The workload is large-block I/O.

**Paravirtual SCSI (PVSCSI) Adapters**
As the name implies the Paravirtual SCSI adapter uses paravirtualization enabling the OS Kernel to communicate directly with the virtualization layer, in this case the ESXi hypervisor. This is why the first important step in creating a virtual machine is to select the operating system as the PVSCSI adapter only works for OSs like Windows Server, Red Hat Enterprise Linux, and some others. For a list of the operating systems that can support the PVSCSI adapters refer to VMware KB 1010398 entitled, “Configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters.” The benefits of paravirtualization are greater performance and lower CPU utilization. In the technical white paper, “Achieving a Million I/O Operations per Second from a Single VMware vSphere 5.0 Host,” the PVSCSI adapter averaged more IOPS and had lower latencies than the LSI adapter. In figure 5 of the paper entitled, “CPU Cost of an I/O Operation with LSI Logic SAS and PVSCSI Adapters” the finding showed an 8% better throughput at 10% lower CPU costs. Below are some of the guidelines for using PVSCSI adapters:

- Use the PVSCSI adapter for large to medium workloads that require better throughput, lower latency, and less CPU cost.
- If the virtual machine is performing a low number of IOPS the there is no need to change from the default LSI or BusLogic Parallel adapter.
  - If using the BusLogic Parallel virtual SCSI adapter and using a Windows guest operating system consider use of the custom BusLogic driver included in the VMware Tools package.[20]
- The queue depth of the SCSI drive can impact disk performance. For example, in the white paper entitled, “Microsoft SQL server Best Practices and design Guidelines for EMC Storage” a benefit was observed in setting the queue depth of the HBA to 64 (the default is 32 in EXSi 5.0, and 64 in ESXi 5.1 and 5.5).[21]

**Consolidation**
EMC storage acceleration technologies, like FAST VP and FAST Cache, combined with VMware virtualization, decrease the number of hard drives and processors for an overall value to the business through consolidation. Starting the discussion at the top of the stack, the general principles of virtualization consolidation are to use the fastest and fewest processors and to drive greater processor utilization by consolidating virtual machines to fewer servers. In the Wikibon article entitled, Virtualization of Oracle Evolves to Best Practices for Production Systems, the author compares the cost of a traditional Oracle Database deployment versus an optimized database infrastructure. The primary difference between the traditional Oracle database deployment and the optimized database infrastructure is virtualization and latency optimized...
primary storage. What is truly interesting is the author links the importance of CPU IO wait time to reducing the number of IOs and reducing IO latency by improving physical reads and writes to storage. The ultimate storage design is one in which storage is not a bottleneck or in the top five wait events of your AWR reports. For example, if DB Time is the top wait event, this means a significant part of processing is spent with sessions running at near 100% CPU. DB Time is defined by Oracle as the amount of elapsed time (in microseconds) spent performing database user-level calls. The key finding in this article was “the additional cost of server and storage enhancements increased the cost/core from $27,716/core to $32,661/core, and increase of 17.8%. However, as shown in Figure 3, the overall cost is significantly reduced because the number of cores can be reduced from 192 to 120, a decrease of 37.5%.”[15] The advent of optimized converged infrastructures like the VCE Vblock™ and others have built solutions around delivering extreme performance for greater consolidation and cost savings to the business.

According to Wikibon, consolidation considerations include:
- Servers have become powerful enough and the virtualization management software stable enough that best practice for deploying x86 Oracle Databases is to create an optimized and virtualized server and storage infrastructure for each level of Oracle Service required.” [15]
- “IT operations should ensure that the best quality servers are used for databases services”
- “…flash is aggressively used to minimize I/O response times and variance.”
- Using VMware to create dedicated clusters for Oracle to prevent the virtual machines moving to unlicensed servers as audits will likely show IT has managed licensing and no additional costs will be incurred.

Ease of Provisioning
The virtualization journey in part is the opportunity for DBA teams to automate manual procedures for faster provisioning time. Manually provisioning of databases is a complex process that could take several weeks or days for the DBA team. We will review solutions that automate the manual tasks enabling the DBA team to accelerate database provisioning. The components below are used my EMC IT for their DBaaS architecture.

vCloud Automation Center
VMware vCloud Automation Center (vCAC) provides a secure portal and self-service catalog to authorized users like administrators, business users, and developers via Microsoft Active Directory. To use an analogy, vCAC enables the business users to shop, much like they would on the internet, for internal IT services. For example, an Oracle developer starting a new project needs a new database. Prior to using vCAC, this process was time consuming as the developer would have to make a request for a new database and perhaps discuss with the DBA team scope, timeline, patch level and support requirements. It wasn’t unusual for this provisioning process to take a week or more because processing the request took planning and manual effort. vCAC changes the provisioning of databases and applications by presenting to the developer and
others a self-service portal of consumable IT services with embedded automation, vastly reducing the provisioning time. Catalogs are logically grouped by service categories meaning related offerings are grouped together. The use of service categories speeds navigation and select as the business user can quickly find the service they need. There is also the ability to expose costs related to ordering IT services and an approval process so application owners can manage their teams IT services. For example, the developer selects a catalog item called, “Medium Oracle Database” consisting of two vCPUs, 8 GB of RAM, and 500 GB of storage array capacity, and the manager receives an automated email detailing the request, initial build cost, and ongoing monthly costs.

There is much more to vCAC than this overview provides, and we recommend reading the VMware manual entitled, Foundations and Concepts, vCloud Automation Center 6.0 for a technical overview of this unified cloud management software.

**vCenter Orchestrator**

VMware vCenter Orchestrator (vCO) automates using a workflow engine tasks like provisioning databases. Most DBAs have procedures in documentation on how to install, manage, and backup databases. For example, in provisioning a copy of Oracle’s E-Business suite, some of the steps include verifying technical requirements, running AutoConfig on the applications tier, synchronizing the appsutil on the database tier nodes, running AutoConfig on the database tier and finally maintaining snapshot information as some of the standard tasks.[16] Many companies then have custom steps before provisioning a copy of the E-Business suite, such as masking sensitive data, removing database links, and deactivating user accounts before opening the applications to a developer. Using vCO, the DBA provisioning procedures can be captured into a workflow to automate and streamline creating copies of the production E-Business suite. Development of complex workflows with vCO is a drag-and-drop build process that includes a scripting engine for granular orchestration, management, and exception handling. The principles of good code management, such as versioning, checkpointing, and central management, are included with vCO making the use of this tool easy for administrators familiar with change control. vCenter Orchestrator is very extensible as the solution provides plug-in Software Development Kits (SDK) and REST API enhancements.

At the time of this writing, the latest version of vCOs is 5.5.1. We suggest reading the datasheet entitled, VMware vCenter Orchestrator, Executing Complex IT Operations Faster and at Lower Cost,[17] as well as vCenter Orchestrator Documentation, for a more detailed overview.[18]

**Puppet**

It is not uncommon for an Oracle DBA team to have to manage ten or more copies of production databases, and depending upon how many different production systems the business has, this could mean larger teams might manage well over a hundred databases. There could be a high degree of complexity in managing many databases, and the automation of tasks will become necessary to efficiently execute on business tasks. Puppet is a solution that centralizes the automation of technical tasks for automation across many different teams. Puppet has been developed to help the sysadmin community move to building and sharing mature tools that avoid
the duplication of everyone solving the same problem.\textsuperscript{[19]} Sysadmins use Puppet because they can use its framework to automate the majority of the technical tasks they perform, and it’s extensible. Puppet has its own custom language to script and share the creation of tasks.

To learn more about Puppet, we suggest reading the Puppet Labs documentation.

**Manageability**
EMC has three Oracle solutions to bridge the gap in managing databases on VMAX and VNX storage arrays, and two of them are free. In this section, we will review the free solution of OEM 12c and the Oracle Workload Profile Assessment (OWPA) and cover the professional delivery of DBclassify.

**OEM 12c Plug-in for EMC VMAX and VNX**
The goal of our OEM 12c plug-in is to enable the DBA to have a view into storage configurations, performance metrics, and to publish reports. Using the plug-in means the DBA can perform storage analysis together with database AWR reports for a more comprehensive approach to remediating performance issues and planning for database growth. With the OEM 12c plug-in installed, the DBA will have access to VMAX and VNX storage configuration details and over 50 performance metrics. The DBA, after some investigation, can then collaborate more closely with the storage administrator. The key difference is that the DBA will know or have a very good idea of the problem because the OEM 12c plug-in provides access to storage information. In a study by IOUG entitled *Efficiency Isn’t Enough: Data Center Lead the Drive to Innovation*, the importance of the DBA collaborating with the storage administrator was discussed:

“Respondents wholeheartedly agree that it is important to improve DBA-to-storage administrator communication and productivity in their organizations. This is something that 80% of data managers agree needs to be improved within their organizations. More than two-fifths feel very strongly about this, indicating that such communication is “very important.”\textsuperscript{[8]}

Indeed, enabling stronger collaboration between the DBA and storage administrator can mean a savings of hours, days, or even weeks in tackling performance issues and eliminating blame storms. Additionally, when working with Oracle support or EMC support, the DBA can more accurately direct the support analyst to the potential problem. For example, the ability to determine that storage performance is not the cause of an issue can equate to a big times savings too.
Figure 4: The OEM 12 plug-in from EMC for the VMAX showing the home screen

Here is a list of some of the features:

- Graphically review the storage configuration for the database.
- View real-time and historical storage performance enables the DBA as well.
  - Immediately launch into storage performance analysis.
  - Find reoccuring performance patterns.
- Set storage alerts called, “thresholds” to proactively tackle performance and meet SLAs.
- Independently conduct analysis AWR and storage analysis.
- More effectively collaborate with the storage administrators.
- Use reports that come with the plug-in or create custom reports.
- Perform comparisons between the current state of the storage configuration to saved configurations.
- Use OEM (minimize the learning curve).

Here are some links to assist with installing and using the OEM 12c plug-in from EMC:
- Download the Plug-installation Guide from [EMC Support](#)
- [Download the Cloud Control Documentation Administrator's Guide](#)
- [Oracle Enterprise Manager 12c Extensibility Exchange](#)
• Read several of the blogs and discussions on the Everything Oracle at EMC community

**EMC's Oracle Workload Profile Assessments – AWR & Statspack**

The Oracle Workload Profile Assessment is a service your EMC Sales Teams and Partner Sales Teams offer to all EMC customers and potential customers. The process involves providing database AWR or Statspack reports to your EMC or Partner contact, and they in turn run the reports through an online analysis tool. After processing the database performance reports, the OWPA tool creates an extensive Microsoft PowerPoint presentation. With a set of AWR or Statspack reports covering several days, the WPA builds a time-series view of the database focusing on the IO characteristics of the database. Metrics such as read and write IOPs, and MB/s, read and write IO sizes, read/write ratio, IO latencies and many other metrics are included in the presentation. Collaboration is the key to this tool as the EMC Sales or Partner Sales person will work with both the DBA and storage administrator to develop consensus on the findings. It’s been our experience that the DBAs may be interested in database analysis like hard parsing, redo log switching, and the storage administrator might be interested in IO latency analysis, EFD and XtremCache indicators. However, the OWPA tool is promoted as a way for performance collaboration across teams.

This is a free service, and if you are interested in getting started, here are some guidelines:

• A set of 48 to 96 Statespack or AWR reports with a 30-minute or 1-hour sample-interval capturing peak activity timeframes. A shorter sample interval and more reports are fine.

• Make the best attempt to identify peak activity periods for the reports: include prime heavy usage business hours and batch processing times.

• We provide sample scripts to ease the generation of the AWR reports in text format (HTML reports are accepted).
  - Link on this link for a zip file with the scripts: **[EMC’s Oracle Workload Profile Assessments blog by William Gaynor](EMC's%20Oracle%20Workload%20Profile%20Assessments%20blog%20by%20William%20Gaynor).**

There is more in an OWPA that a DBA may be interested in seeing: hard parsing, redo log switching, redo log writes and bytes per second, host CPU utilization by system, user, and IO wait, most active tablespaces over time and their read response times in milliseconds, a list of non-IO related timed events showing any latency and finally transactions, executions and commits per second. Again, all metrics are graphed over time.

Here is another link to assist with using the Oracle Profile Assessment:

• **[EMC Workload Profile Assessment for Oracle AWR Report / Statspack Gathering Procedures Instructions](EMC%20Workload%20Profile%20Assessment%20for%20Oracle%20AWR%20Report%20%2F%20Statspack%20Gathering%20Procedures%20Instructions).**

This collaboration extends to both the DBA and storage administrator as the focal point of the OWPA tool is analyzing database reports and overlaying storage considerations. For example,
the OWPA, there are sections covering IO Latency Analysis, EFD and XtemCache Indicators, top tablespaces and top timed events for IO and non-IO. With a set of AWR or Statspack reports covering several days, the OWPA builds a time-series view of the database.

**DBClassify: A Professional Services Option**

*DBClassify* is a tool delivered as part of a service offering used to monitor database I/O for both Oracle and SQL Server environments. The differentiator is that DB Classify integrates with FAST VP to enable the DBA to collaborate with the storage administrator in setting FAST VP policies tuned for databases. There are two types of service offerings: EMC database performance tiering assessment and database performance tiering residency. The primary difference between the DB Classify assessment and residency service offerings is scope:

- Performance Tiering Assessment is a comprehensive databases performance analysis of 3 databases.
- Performance Tiering Residency is only limited by time (160 hours delivered over 4 weeks or spread over 3 months).

The reason this is delivered only as a service is that the EMC database performance resident works side-by-side with the DBAs and storage administrators in transferring knowledge and to accelerate the intelligent storage tiering into production database environments. After the assessment or residency, the DB Classify tool remains at the customer which is important as new workloads and changes in existing workloads means continued analysis will be critical to take full advantage of FAST VP. Of all the tools covered thus far, DB Classify has the greatest degree of database to storage integration as it can:

- Analyze the database to make recommendations for creating storage groups.
  - Storage administrator uses the analysis to intelligently define the FAST VP storage tiers.
  - DBA leverage storage tiers for the database layout.
- Optimize database performance at the lowest costs.
- Enables database policy influence for FAST VP environments.

What is most interesting about DB Classify is the ability to influence the FAST VP policy engine. What this means is DBAs can use DB Classify to make sure that certain tables or processes would be allocated on the fastest storage tier. For example, if there is a process or report that is not frequently run but its quick completion is critical to the business, the DBA can use DB Classify to place the database tables used by the report in the Flash tier and when the report runs, it’s very fast as all the supporting tables are in the fastest tier of storage. The importance of this feature could have a significant impact to the business because it is frequently easier to use hardware to accelerate a process than it is to have the development team tune the code. For more information on this, please read the overview “EMC Database Performance Tiering Assessment” or contact your local EMC Services representative.
Availability
Using EMV VPLEX Metro, the DBA team can use Oracle RAC stretched over data centers for continuous operations. There are different VPLEX configurations:

- **VPLEX Local** provides seamless, non-disruptive data mobility across heterogeneous arrays within the datacenter.
- **VPLEX Metro with AccessAnywhere** enables active-active, block level access to data between two sites within a synchronous distance of five milliseconds round-trip-time (RTT).
- **VPLEX MetroPoint with RecoverPoint** can be used to build a three-site continuous protection from local and regional events. Using MetroPoint, your Oracle RAC can survive two site failures.
- **VPLEX Geo** provides asynchronous availability across heterogeneous arrays between two sites that exceed 5 milliseconds of RTT.

For the purpose of this discussion, we will discuss VPLEX Metro with Extended Oracle RAC. DBA teams can extend Oracle RAC between data centers for continuous availability. Using VPLEX, the business can federate heterogeneous storage arrays to mirror all writes to both storage arrays. VPLEX Metro ensures write-order fidelity so all writes across the Oracle RAC nodes will be acknowledged in the same way as using one storage array. Both from a DBA and database perceptive VPLEX Metro is transparent meaning no additional management, patching, or complexity. Extended Oracle RAC with VPLEX Metro can survive failures like the loss of a HBA port, server(s), building or site, network disconnect, and others making this solution a very resilient option for mission critical applications that cannot have unplanned downtime. We suggest reading, “Oracle Real Application Clusters (RAC) on Extended Distance Clusters with EMC VPLEX Metro” for a technical overview of the solution.
The guidelines below are from the previously mentioned white paper for using Extended Oracle RAC with VPLEX Metro:

- Generally, a RTT of 3 ms or less is preferred with moderate database workloads although the continuous availability solution is certified for a maximum of 5ms RTT.
- Use RAID 5 for the data files as this protection level offers good write performance, resiliency, and only requires 12.5% capacity for protection.
- Use RAID 1 for the redo log files as it offers the best write performance and better resiliency than RAID 5. Note in the paper only external RAID protection and not duplexing was used. Duplexing adds additional IO to the stretched Oracle RAC configuration and external protection should be strongly considered in place of duplexing. For example, for even greater resiliency RAID 6 protects against multiple drive failures and is a good option for both your redo logs and archive logs.
- Consider using Multiple ASM disk groups but ensure they are all in the same consistency group to guarantee write ordering. For example, in the paper the following ASM disk groups were used:
  - There are no specific guidelines for the placement of Temp files. The DBA can include with +DATA or separate for monitoring purposes.
    - It is recommended to create a unique disk group for CRS alone, for example: +GRID. This approach of separating the clusterware LUNs from the database is useful when there are plans to use storage technologies such as clones or snapshots to create additional database copies for repurposing.
  - In the paper +DATA, +LOG, and +FRA ASM disk groups were separated to allow storage based snaps and clones to be used for offloading backups from production. For example, cloning production to a backup host is advantageous because the IO from the backup operation will not impact the Stretched Oracle RAC cluster.
- VPLEX does not natively perform network compression, but compression on a network switch can be utilized.
- Create a dedicated VPLEX logging volume for each Oracle RAC cluster
  - A dedicated logging volume means resynchronizing I/O across the storage arrays after a failure will go faster.
  - Use RAID 1 for VPLEX logging volumes.
  - Configure at least 1 GB (preferably more) of logging volume space for every 16TB of distributed device space

Virtualizing an Oracle single instance and RAC extends resiliency of the VPLEX Metro architecture as the DBA can use federated HA. For example, in the case of a server or a site failure, the Oracle single instance databases and RAC nodes, using VMware HA, would automatically restart in the surviving data center. Stretched Oracle RAC would remain running but the lost nodes would be
restarted at the surviving site to minimize loss of performance. Upon recovery of the server or site, the virtualized databases and RAC nodes could be non-disruptively vMotioned back to the original equipment and site. For a detailed review of adding resiliency and automation to an Oracle VPLEX Metro architecture, we recommend reading, “Using VPLEX Metro with VMware High Availability and Fault Tolerance for Ultimate Availability.”

Everything Oracle at EMC Community
The Everything Oracle at EMC (EO@EMC) community is a place where Oracle professionals gather to discuss, collaborate, blog and exchange ideas. It might be interesting to know that many of the best practices reviewed in the paper were selected based upon customer discussions and blogs on the EO@EMC community. Becoming a member of the community is free and easy as following, Starting with the ECN: Register for an Account. Becoming a member has its benefits including participating in discussions, asking questions, and, if you really become social, you might be invited to join the EMC Elect team. The EMC Elect team is a group of community-driven people interested in engaging on EMC solutions and showing leadership by engaging others on the communities.

In the Oracle community are all the proven solutions, white papers, reference architectures that EMC has developed for Oracle over the last three years. The author of this paper and other authors, engineers, and DBAs can be found on the community. By joining the community you will have access to a broad range of database, virtualization, and storage professionals who share a passion for technology. Registering once also provides access to many of the other communities such as: Microsoft, SAP, VMware, and Isilon. We hope you join the Oracle community and share your feedback on this paper and engage others on using EMC solutions for Oracle.

Conclusion
This paper describes some of the best practices and deployment guidelines for Oracle databases using VMware virtualization and EMC proven solutions. EMC, VMware, and our partners want your virtualization of Oracle journey to be a success and built upon a foundation of proven practices that provide greater agility, flexibility, resiliency, and performance. Once your mission critical applications have been virtualized the next stage could be ITaaS.

Oracle DBAs and other application owners have a stake in ITaaS and adding automation to cover provisioning and routine tasks. Working towards building DBaaS involves combining all the best practices together: virtualization, databases, storage, and more. It’s the consolidation of best practices that can lead to standardization and eventually to self-service provisioning. The foundation is the automation of deploying your best practices in the delivery of databases and applications.
References
[1] “Production Oracle Databases reaches 55% and is Growing” Fast by David Floyer and published in September, 17, 2013
[3] ESG Lab Validation report entitled, “Automated Path Optimization for VMware Virtual Environment” published in April 2012. Reference to Figure 6: Performance Comparison, IOPS-intensive Workload
[4] “EMC VNX Scaling Performance for Oracle 12c RAC on VMware vSphere 5.5” published December 2013
[6] “vSphere Storage, ESXi, vCenter Server 5.5” published in 2013 by VMware
[10] EMC Technical Notes “FAST Cache and FAST VP for VNX OE for Block” release notes Rev 01 published August 16, 3013
[18] VMware vCenter Orchestrator Documentation URL: 
https://www.vmware.com/support/pubs/orchestrator_pubs.html

[19] “Puppet Labs, Puppet documentation” generated on July, 2013. URL: 
http://downloads.puppetlabs.com/docs/puppetmanual.pdf


[22] “Oracle Real Application Clusters (RAC) on Extended Distance Clusters with EMC VPLEX Metro” published in September 2011

**Other References**

- Using EMC Symmetrix Storage in VMware vSphere Environments, Version 9.0, URL: 
- Using EMC VNX Storage with VMware vSphere Environments, Version 3.0, URL: 

**Acknowledgments**

The author, Sam Lucido, would like thank Jason Kotsaftis for proposing the idea of this paper and his mentorship along the way. With Jason’s vision and support the idea transformed into a paper we hope adds value to our customers. Sometimes two hands aren’t enough and turning to someone for help can make a big difference. That was George O’Toole as he was always ready to help and without his assistance the writing of this paper may have stalled a few times. Jeff Browning is among the top thought leaders on virtualizing Oracle databases and his insights helped drive many of the topics in the paper.

Finally, the number of contributors and reviewers of this paper was amazing and their feedback raised the level of the paper. I believe EMC is a great place to work because of teamwork. My thanks to a very special team of contributors and reviewers that made this paper a fun and engaging process.