Dell EMC Ready Solutions for Oracle with XtremIO X2 - Frequently Asked Questions

Question: **What is XtremIO?**

Answer: Dell EMC XtremIO is a purpose-built all-flash array that offers consistently high performance with low latency; unmatched storage efficiency with inline, all-the-time data services; rich application integrated copy services; and unprecedented management simplicity.

Dell EMC also introduced X2-T. This entry price point model is a single X-Brick offering built on XtremIO's unique metadata centric architecture with full data services including inline data reduction (in-memory space efficient copies, deduplication, and compression), and enterprise grade availability. X2-T is ideal for competitive situations with capacities ranging from ~100TBe to 200TBe with predictable growth rates.

Question: **What are the key components of the new Ready Solution for Oracle with XtremIO X2?**

Answer: The key components of the new Ready Solution for Oracle with XtremIO X2 includes:

- PowerEdge 14G servers: The R940 / R840 and R740
- Dell EMC XtremIO X2-S (X2-R is also supported)
  - 2 x X-Bricks and a total of 36 drives
- Data Domain 6300
- Dell TOR networking S4048-ON, Connectrix B-Series 6X10 (16GB) SAN switches

The Ready Solution for Oracle is designed to allow flexible configurations based upon the workload requirements. Customers can design the Ready Solution for Oracle to meet their requirements with the confidence the configuration will support their database ecosystem.

Question: **What software components was the Ready Solution for Oracle validated with?**

Answer: The Ready Solution for Oracle have been validated with:

- Oracle Enterprise Cloud Manager 13c
- Oracle 12c R2 single instance and RAC
- Oracle 12c R2 grid infrastructure
- Oracle 11g R2 single instance
- Oracle Recovery Manager (RMAN) 12c
- Red Hat Enterprise Linux (RHEL) 7.4 for the Oracle 12c databases
- Red Hat Enterprise Linux (RHEL) 6.9 for the Oracle 11g databases

(intel)

- Dell EMC Live Optics
- Data Domain Boost for Enterprise Applications (DBEA)
- VMware 6.5 U2

Question: **What database workloads were part of the validation testing?**

Answer: The following database workloads were part of the validation testing:

- Silly Little Oracle Benchmark (SLOB) was use to generate the OLTP workload
- Custom SQL query was used for the OLAP workload

Question: **How is the Ready Solution for Oracle designed for high availability?**

Answer: The Ready Solution was designed for full redundancy to eliminate single points of failure. We encourage customers to review the design and

- Each of the database servers consists of highly available LAN and SAN connectivity with multiple network interface cards (NICs) and multiple host bus adapters (HBAs).
  - Two dual-port 10 GbE NICs and two dual-port 16 Gb/s HBAs in each database server provide sufficient bandwidth and high availability for Oracle database LAN and SAN traffic.
  - Redundant network ports are connected to separate 10 GbE switches for HA and bandwidth in case of failure.
  - HBA ports on the same HBA adapters are connected to separate FC switches for HA and bandwidth in case of failure.
- Redundant ToR S4048-ON 10 Gb Ethernet switches provide sufficient 10 Gb ports and bandwidth to share both Oracle database public and private interconnect traffic. These switches also provide sufficient 40 GbE uplink ports to connect the Ready Solution for Oracle to the datacenter core switches.
- Redundant DS-6510B 16 Gb/s Fibre Channel (FC) switches provide sufficient ports and bandwidth to share Oracle database SAN traffic between the database servers and the XtremIO X2 storage controllers.
- A dual X-Bricks XtremIO X2 cluster storage array with redundant and active-active storage controllers provide sufficient capacity, bandwidth, and high availability for all the Oracle databases tested in the Ready Solution for Oracle.
  - Each FC port in each XtremIO X2 storage controller is connected to separate FC switches for high bandwidth and high availability (HA).
- A management port from each major hardware component is connected to S3048-ON 1 GbE management switch.
  - On each of the database servers:
  - A dedicated iDRAC network port is connected to the management switch for out-of-band, bare-metal management of the server
- The 1 GbE LAN on motherboard (LOM) network ports is connected to the management switch for in-band management of the server from within the operating system

Question: **What are compute best practices?**

Answer: Compute best practices include:

- Set the BIOS System Profile to: `Performance`
- The PCIe network adapters and HBAs that are used for Oracle database public, private interconnect, and SAN traffic were populated based on the recommended PCIe slot priority for optimal power, bandwidth, and thermal performance of the adapters and the system. The populated and the recommended PCIe slots are wired to two separate CPUs in their respective servers, allowing for load balancing of I/Os across the two CPUs.

- For optimal performance, in the PowerEdge R940-based databases servers, the HBAs were populated in PCIe slots 2 and 5 and the NICs were populated in PCIe slots 3 and 6.
- For optimal performance, in the PowerEdge R740-based database server, the HBAs were populated in PCIe slots 1 and 7 and the NICs were populated in PCIe slot 3 and the rack Network Daughter Card (rNDC) slot

Question: **What are the VMware host virtualization best practices?**

Answer: **HBA queue depth settings** - The following table lists the default and the recommended HBA queue depth settings in ESXi 6.5 hosts connecting to XtremIO X2 storage arrays:

HBA queue depth settings in ESXi based hosts

Parameter	Default value	Recommended value
LUN Queue Depth	QLogic: 64 Emulex: 30	QLogic: 256 Emulex: 128
HBA Queue Depth	QLogic: N/A Emulex: 8192	QLogic: N/A Emulex: 8192 (maximum)

We set the LUN queue depth to the recommended value of 256 for the QLogic HBAs used in our virtualized databases server. This setting ensures that the XtremIO X2 storage arrays handle an optimal number of SCSI commands (including I/O requests).

Note: The QLogic HBA Queue Depth setting is no longer read by vSphere, therefore, it is not relevant when configuring a vSphere host with QLogic HBAs.

Multipath configuration - The virtual databases ESXi host was configured by using vSphere Native Multipathing (NMP). The following parameter values are recommended on the host for optimal performance with XtremIO X2 storage:

- Set the native path selection policy on the XtremIO volumes presented to the ESXi hosts to round-robin.
- Set the vSphere NMP round-robin path switching frequency to XtremIO volumes from the default value (1000 I/O packets) to 1.

Note: In ESXi 6.5, the default path selection policy is round-robin and the default path switching frequency is 1. Therefore, no change was needed in our virtualized databases server.

Host parameter settings - The following ESXi host parameters were configured:

- `Disk.SchedNumReqOutstanding`—Determines the maximum number of active storage commands (I/O) allowed at any time at the VMkernel. This parameter value for each XtremIO volume presented to the ESXi host was set to the recommended value of 256 by running the following CLI commands on the ESXi 6.5 host:

To get the list of all XtremIO volumes, type:

```
$> esxcli storage nmp path list | grep XtremIO -B1 | grep "\ naa" | sort | uniq
```

To set the value for each volume, type:

```
$> esxcli storage core device set -d <naa.xxx> -O 256
```

where `<naa.xxx>` is the XtremIO volume obtained from the previous command.

- `Disk.SchedQuantum`—determines the maximum number of consecutive "sequential" I/Os allowed from one VM before switching to another VM. This value was set from its default value of 8 to the recommended value of 64.

Question: **What are the virtual machine best practices?**

Answer: We used the following design principles and best practices to create the VMs in this Ready Solution for Oracle:

- SCSI controllers—We created multiple SCSI controllers to optimize and balance the I/O for the different database disks, as shown in the following table.

Table 1. SCSI controller properties set in VMs

Controller	Purpose	SCSI bus sharing	Change type
SCSI 0	Guest OS disk	None	VMware Paravirtual
SCSI 1	Oracle DATA disks	Physical	VMware Paravirtual
SCSI 2	Oracle REDO disks	Physical	VMware Paravirtual
SCSI 3	Oracle OCR, GIMR, FRA, TEMP	Physical	VMware Paravirtual

- **Hard disk drives**—All database-related virtual disks—for example, DATA, REDO, FRA, OCR/VD, and TEMP—were assigned the following properties:
  - Type: 'Thick provisioned eager zeroed'—Which ensures that the space required for the virtual disks are allocated at creation time and the data on the physical device on the storage is zeroed out
  - Sharing: 'No sharing'—Selected because the deployed databases are standalone databases that do not require sharing database virtual disks with another VM, unlike RAC VMs that share virtual disks between two or more VMs in the database cluster.

Question: **What are the Red Hat best practices for Oracle?**

Answer: **RHEL 7.4 as bare-metal operating system** - In the PowerEdge R940-based production database servers and the PowerEdge R740-based XVC databases server, we deployed RHEL 7.4 as the bare-metal operating system.

Note: We used the XtremIO Host Configuration Guide to apply best practices. See this guide for the complete list of best practices for XtremIO storage in a Linux environment.

HBA queue depth settings - The following table lists the default and recommended HBA queue depth settings for a Linux environment.

Table 2. HBA queue depth settings in Linux-based servers

Parameter	Default value	Recommended value
LUN Queue Depth	QLogic: 32 Emulex: 30	QLogic: Keep default value Emulex: Keep default value
HBA Queue Depth	QLogic: 32 Emulex: 8192	QLogic: 65535 (maximum) Emulex: 8192 (maximum)

Note: We kept the default queue depth values in our physical database servers because we used Emulex HBAs.

I/O elevator settings - The recommended I/O elevator setting in the RHEL operating system running in the database servers connecting to XtremIO X2 storage arrays is either `deadline` or `noop`. The `cfg` I/O elevator setting is not recommended. In our physical database servers, we used `deadline`, which is the default I/O elevator setting in RHEL 7.4.

Multipath configuration - We configured the physical database servers using Linux Native Multipathing available in the RHEL 7.4 operating system. We created the configuration file for the `/etc/multipath.conf` multipath daemon with the following recommended settings:

```
devices {
     device {
               vendor                XtremIO
                  product              XtremApp
               path_grouping_policy   multibus
                  path_checker           tur
                  path_selector          "queue-length 0"
                  rr_min_io_rq            1
                  user_friendly_names   yes
                  fast_io_fail_tmo       15
                  failback              immediate
          }
```

Partition alignment in Linux - We partitioned the database disks or XtremIO volumes presented to the Linux-based physical database servers by using fdisk with the default starting sector value of 2,048. This setting ensures that the starting sector number is a multiple of 16 (16 sectors, at 512 bytes each, is 8 KB). Therefore, each database disk is correctly aligned with the XtremIO storage LUN striping.

RHEL as VM guest operating system - In this Ready Solution for Oracle, we deployed RHEL 6.9 as the guest operating system in VM1 running Oracle 11g R2 database and RHEL 7.4 as the guest operating system in VM2 running Oracle 12c R2 database.

In each guest operating system , we adhered to the following recommended best practices.

PVSCSI driver - For optimal XtremIO X2 storage performance in a VMware environment, we recommend PVSCSI controllers and driver in the guest VMs. In this Ready Solution for Oracle, we ensured that the inbox RHEL `vmw_pvscsi` driver module was loaded and used in the guest operating systems.

Note: The PVSCSI driver is used only when the SCSI controller type is set to `VMware Paravirtual` in the VM settings.

PVSCSI LUN Queue Depth and ring-pages settings - The following table shows the default and the recommended `vmw_pvscsi` parameter settings.

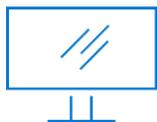**Table 3. PVSCSI parameter settings in guest operating systems**

Parameter	Default value	Recommended value
vmw_pvscsi.cmd_per_lun	RHEL 6: 64 RHEL 7: 254	RHEL 6: 254 RHEL 7: 254
vmw_pvscsi.ring_pages	RHEL 6: 8 RHEL 7: 8	RHEL 6: 32 RHEL 7: 32

The parameters and their respective recommended values in this table were appended to the kernel boot arguments:

- In the `/boot/grub.conf` file for the RHEL 6-based guest operating system

- In the `/etc/default/grub` file for the RHEL 7-based guest operating system

Other guest operating systems were configured as follows:

- Multipath was not configured because it is handled at the ESXi host level.

- In the 12c database VM, all database storage disks were set up by using Oracle ASM Filter Driver (ASMFD).

- In the 11g database VM, all database storage disks were set up by using `udev` rules.