Reference Architectures for designing and deploying Microsoft SQL Server Databases in Active System800 Platform

Discuss database workload classification for designing and deploying SQL server databases on the Dell Active System 800 platform.

Dell SQL Database Solutions Engineering

May 2013| Rev 1.0

Contents

Tables

Figures

Executive summary

Virtualization provides highly efficient data center computing environments. To optimally design a virtual environment, an initial understanding of the underlying platform and its capabilities is important. This white paper proposes a virtual converged infrastructure using Dell Active System 800 for virtual database deployments. Dell Active System 800 is a member of the Dell™ Active Infrastructure family. Dell Active System 800 is a pre-engineered converged infrastructure system that is designed to address the typical data center challenges such as design, procurement, rapid deployment and testing. Dell Active System 800 benefits from the latest technology advances in the Dell™ server, storage, networking portfolios and virtualization era. Dell Active System 800 combines servers, storage, networking and management into an integrated and optimized system aimed to reduce the overall Total Cost of Ownership (TCO).

This white paper provides design guidelines and recommendations for the design and deployment of Microsoft SQL Server databases on the Dell Active System 800 converged virtualized platform. In addition, this white paper discusses a standardized classification of workloads based on database workload requirements assumptions for a wide range of applications.

Introduction

Many customers are using a technique called workload consideration to reduce the overall Capital and Operational IT costs. Database consolidation is achieved in many different ways depending on the systems and circumstances involved. Database consolidation on to a virtualized platform is gaining more popularity because of its unique advantages such as:

- Rapid deployment
- Centralized management
- Scalability
- Portability
- Highly available features

These advancements are attributed to server, storage and network technologies that support more bandwidth and computing horsepower.

It would be useful to evaluate a standardized classification of database workloads for design, deployment and consolidation on virtualized platforms. This activity provides an idea of the expected database performance using the specific configuration and workload. This standardization may also ease the deployments and it presents a common set of maintenance and configuration requirements. Use these configurations as a baseline in designing databases for the customer-specific requirements. The configuration must be fine-tuned to achieve the optimal performance for different workloads.

The goal of this paper is to design the virtual machine configurations based on predetermined database workload requirements for the Dell Active System 800 platform. In this white paper, we discuss a set of assumptions, which help us classify the standard OLTP database workloads. These assumptions reflect the important aspects to consider while determining the configuration of virtual machines for the Microsoft SQL Server database deployments. This white paper also discusses the observed database performance numbers for the proposed Dell Active System 800 reference configuration. This helps us understand the virtualization capabilities of the proposed infrastructure.

Objectives of this white paper

The main objectives of this technical paper are:

- Introduce the Dell Active System 800 converged platform for designing and deploying Microsoft SQL Server databases to virtualized infrastructure.
- Discuss a set of assumptions for classifying OLTP database workloads and provide an example of the database design for virtual environments.
- Discuss the design considerations for classifying SQL Server database workloads, which meet the requirements of small, medium and large OLTP-based applications.
- Discuss the virtual machine configurations needed to meet the workload classification requirements.
- Evaluate the performance of the virtual machine configurations for the proposed reference configuration.

Latest virtualization technologies and their impact

In this section, we discuss some of the new software and hardware technologies that recently emerged that have significant impact on virtualization.

Windows Server 2012 virtualization features

Windows Server 2012 comes with many improvements and new features in virtualization:

- Windows Native NIC teaming
- VHDX format
- Extensible virtual switch
- SR-IOV
- Integrated DCB support
- Improved cluster scalability and cluster aware updating

All these features and enhancements provide a robust virtual platform for deploying virtual machines running database workloads. For more details on Windows 2012 Hyper-V features, see: http://download.microsoft.com/download/2/C/A/2CA38362-37ED-4112-86A8-FDF14D5D4C9B/WS%202012%20Feature%20Comparison_Hyper-V.pdf

The emergence of converged infrastructures

Converged infrastructures help customers increase IT speed, agility, data center efficiency and decreases data center operation costs. The goal behind the converged infrastructure is the convergence of different types of network traffic such as LAN, SAN, and IPC over single a network fabric. All network traffic shares the same physical connections.

Data Center Bridging: Data Center Bridging (DCB) is an IEEE standard that helps in building the converged fabrics. DCB provides the following set of functionalities.

Priority Flow Control (PFC): The goal of PFC is to make sure of zero frame loss under congestion in DCB networks. It provides a link-level flow control mechanism that can be controlled independently for each frame priority.

Enhanced Transmission Selection (ETS): ETS provides a common management for assignment of bandwidths to the different frame priority.

Congestion Notification (CN): CN provides end-to-end congestion management for protocols without built-in congestion-control mechanisms.

Data Center Bridging Exchange protocol (DCBX): DCBX is a discovery and capability exchange protocol to convey capabilities and configurations of the above three DCB features between neighbors to make sure of a consistent configuration across the network.

With the help of the above four functionalities, DCB converts the Ethernet into a loss-free transport protocol enabling a reliable, robust Ethernet SAN fabric. This enables us to use a single fabric for all types of traffic.

Dell offers a converged infrastructure platform, called the Active Infrastructure family, which leverages the latest technical advancement in virtualization. The Active Infrastructure family uses the latest Dell PowerEdge 12th Generation servers, 10GbE iSCSI EqualLogic storage and 10GbE networking using Dell Force10 S4810 TOR switches and PowerEdge IO Aggregators.

This white paper discusses Dell Active System 800, a member of the Active Infrastructure family platform for consolidating Microsoft SQL server databases on to a converged virtual platform using Hyper- V3.0. For more details on Active Infrastructure and the Active System family, see: http://www.dell.com/Learn/us/en/555/converged-infrastructure

Importance of database workload classification on converged virtual platforms

There are usually two types of database workloads: Online Transactional Processing (OLTP) and Decision Support System (DSS).

Operational or Online Transactional Processing (OLTP): OLTP workloads are characterized by numerous small interactive transactions that generally require sub-second response time. The OLTP systems are required to support many small transactions in a concurrent manner. The OLTP workloads are characterized by small IO random access patterns. It is not unusual to have more writes than reads in heavy OLTP systems.

Decision Support System (DSS): DSS workloads are different from OLTP workloads. DSS workloads are read intensive that generally have fewer users connected to the system. These workloads are characterized by large sequential data access patterns.

Because the transaction nature and IO behavior of OLTP and DSS workloads are different, the virtual machine configurations must be carefully designed to the meet the specific database requirements in the converged virtual platforms, such as Dell Active System 800. Exclusive sharing of virtual infrastructure by different workloads imposes the risk of unpredicted performance implications during heavy load scenarios.

This white paper focuses on deriving the database characteristics for OLTP workloads in a converged virtualized infrastructure.

Evaluating a standard set of database workloads is helpful in determining how a specific virtualization platform behaves at different load scenarios. This exercise also helps in easing the database virtual machine deployments, as it presents a common set of maintenance and configuration requirements. This activity also helps get an idea of the expected database performance using a specific configuration and workload.

The following sections discuss the process of designing various database classifications of Online Transaction Processing (OLTP) database workloads based on a set of assumptions. These assumptions represent the important performance parameters being evaluated when accessing the performance of any database configuration. The resulting database virtual machine configurations, along with the observed performance characteristics, are used as a baseline in designing virtual databases for the specific customer requirements. The configuration may need to be fine-tuned to achieve the optimal performance for the different workloads.

What makes Dell Active Infrastructure Systems ideal platforms for database solutions

The four key elements of Dell Active Infrastructure Systems that make it the ideal platform for consolidation and different from other competitive solutions are:

Modular infrastructure: Dell Active Systems was built with a modular M1000e blade enclosure in which full height, half height and quarter height blades can be combined, installed and serviced together. It provides great flexibility and scalability to customers because they can add compute nodes as they grow. Up to 16 Dell PowerEdge M620 blades can be installed and serviced together within a single M1000e blade enclosure.

Flexible IO Networking Model: Dell Active System platforms use simplified blade IO modules that provide flexible networking options with greater networking bandwidth. The Dell PowerEdge M I/O Aggregators are used in the Dell Active System 800 platform. It extends the consolidation benefits of virtualization by replacing multiple GbE NICs, mezzanine cards and switches with fewer, higher-bandwidth 10GbE NICs and I/O modules for each blade chassis. It pushes all the advanced switch management tasks such as VLAN management, Spanning Tree protocol configuration and Quality of Services (QoS) to the Top-of-Rack switches there by reducing extra configuration efforts on blade IO modules. Dell Force10 S4810 switches used as ToR switches are low-latency and high-density 10GbE/40GbE switches. The compact Dell Force10 S4810 design provides industry-leading density of 48 dual speed (1/10GbE) ports (SFP+) as well as four 40GbE QSP+ uplink ports to conserve the valuable rack space and simplify the migration to 40Gbps in highly virtualized data centers.

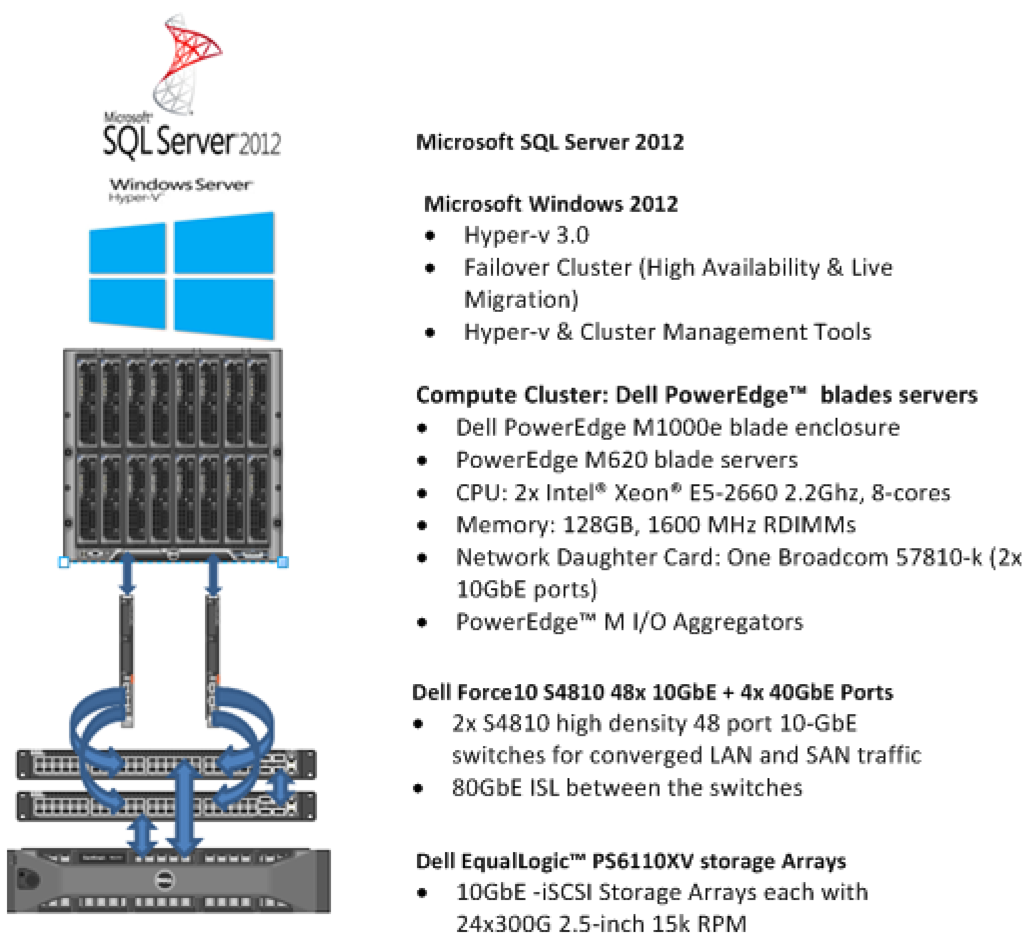
Converged Fabric: Dell believes that the converged fabrics enable customers to get the greatest simplicity and efficiency. Dell Active Infrastructure Systems are converged platforms where the LAN and SAN traffic is combined with in a single fabric, which simplifies and improves the efficiency of customer networking devices. The Dell Active System 800 platform uses Dell Force10 switches and Dell EqualLogic PS6110 series storage arrays, which supports converged fabrics by using the DCB protocol. The Dell Active Infrastructure system also provides the flexibility of using Top of Rack (TOR) switches from different vendors.

Converged Infrastructure Management: Dell offers end-to-end management of the complete stack using Dell Active System Manager. It focuses on reducing the operational costs and complexity of managing the complete infrastructure. Through capabilities such as template-based provisioning, automated configuration and infrastructure lifecycle management, Dell Active System Manager enables IT to respond rapidly to business needs, maximize data center efficiency and strengthen quality of IT service delivery.

Dell Active System 800 reference architecture for Microsoft SQL Server databases

The Dell Active System 800 solution combines the latest technology advancements in the virtualization era. The Dell Active System 800 solution offers rapid and simple scalability. Customers can grow their computer and storage capacity in a modular and non-disruptive scaled-out fashion. This pre-integrated solution includes Dell PowerEdge M620 blade servers, PowerEdge I/O Aggregators and Dell Force10 for 10GbE networking and Dell EqualLogic iSCSI storage in a converged iSCSI fabric. The Figure 1 illustrates the major hardware and software components in the Dell Active System 800 platform.

Figure 1. Dell Active System 800 reference architecture for SQL Server database deployment.



The following sections provide more details on the hardware and software components used in the proposed reference architecture.

Dell PowerEdge™ M620 Blade servers: Dell PowerEdge™ M620 servers are two-socket half-height blade servers, which can fit inside the Dell PowerEdge M1000e blade chassis. Up to 16 Dell PowerEdge M620 blade servers can be installed in single M1000e blade enclosure and each blade server is populated with two-Intel® Xeon E5-2660 2.2GHz, 8-Core processors and 128GB memory. Each blade server is configured with one Broadcom BCM 57810-k 10GbE dual port NDC card in slot A. These two ports are internally connected to PowerEdge M I/O Aggregators in Fabric A.

Dell PowerEdge™ M I/O Aggregator: The Dell PowerEdge M I/O Aggregator is a flexible 1/10GbE aggregation device that is preconfigured for easy *plug and play* deployment into converged networks. It provides connectivity to the Converged Network Adaptor (CAN)/network adaptors internally and externally to upstream network devices. The Dell PowerEdge M I/O Aggregator has two integrated 40 Gbps Ethernet ports on the base module. These ports are connected to the Dell Force10 S4810 TOR switches with 4 x 10GB breakout cables.

Dell Force10™ S4810 Switches: Dell Force10™ S4810 is a 10/40GbE TOR switch built for high performing applications. It supports up to 48 x 1/10 GbE (SFP+) ports as well as 4 x 40 GbE QSFP+ uplink ports. It also supports Powerful Quality of Service (QoS) features coupled with Data Center Bridging support. In this configuration, two S4810 switches are used and lagged with each other using 40 GbE uplink ports. A Virtual Link Trunk (VLT) domain is created between switches for efficient usage of 40 GbE uplink ports.

Dell EqualLogic PS6110XV: The Dell EqualLogic PS6110XV is a virtualized high-performing 10GbE iSCSI storage array with dual 10GbE controllers and dual power suppliers. With 300G 15,000 RPM SAS disk drives, the Dell EqualLogic PS6110XV array provides high transactional sequential performance with up to 7.2TB of storage capacity in a 2U form factor. It supports SFP+ and 10GBASE–T port options with automatic vertical port failover. The 10GbE ports of Dell EqualLogic PS6110 are connected to the Dell Force10™ S4810 TOR switches. Storage capacity can be scaled up by simply adding Dell EqualLogic PS6110XV storage arrays to the existing group of storage arrays.

Building the proposed Active System 800 converged infrastructure for SQL Server deployments

This section provides information on how we built the converged infrastructure.

The two blade hosts were clustered using Microsoft Windows Server 2012 with Hyper-V role enabled on both the hosts. Using the native network Load Balancing Failover (LBFO) teaming feature of Windows 2012, the two ports of the BCM 57810 NDC card were teamed. Many virtual NICs were created for different networks such as the public, cluster private, Live Migration and VM network, and so on. Each network is isolated by using specific VLAN IDs.

In addition to the standard LAN device, the BCM 57810 NDC card on each host was enumerated as an iSCSI initiator device for a storage network. This enabled us to use a single BCM 57810 NDC card for both LAN and SAN traffic using network convergence.

On the Dell Force10™ S4810 TOR switches, the DCB was enabled and configured appropriately. We allocated 50 percent bandwidth for LAN traffic and 50 percent for the SAN traffic. This bandwidth allocation is configurable and may be altered based on the customer requirements.

For more details on the Dell Active System 800 reference architecture design and deployment guides, see the Active System 800 documentation.

Test methodology

The objective behind the test exercise was to come up with a set of virtual database configurations designed to meet the small/medium/large workload classifications. These classifications were derived from a set of performance assumptions.

Performance assumptions for the database categories

The following sections describe how to classify the OLTP database requirements for different ranges of applications based on various characterization parameters.

Different applications may have different OLTP database requirements. Typically, the OLTP database requirements are mostly characterized by the following parameters:

- **Database Size:** Is the size of the OLTP database. The sizes of OLTP databases may range from one to some hundreds of Gigabytes.
- **Number of Concurrent Users:** Indicates the number of concurrent database users. In OLTP systems, we typically observe 10 to hundreds of concurrent database users querying the database.
- **Number of Transactions per Second (TPS):** Indicates the number of transactions that a database is able to deliver in a second.
- **Response Time (RT):** Indicates the time taken by the database engine to run and return the result of the query.

The parameters help to determine the amount of hardware resources such CPU, memory and storage required by a virtual machine.

The following table briefs the baseline assumptions to classify the virtual database configurations into small/medium/large categories.

Table 1. Baseline assumptions for OLTP database workload classification

Parameters	Small Workload	Medium Workload	Large Workload
Database Size	< 20GB	< 100GB	> 250GB
Concurrent Users/connections	< 50	< 500	> 1000
Transactions / Sec	< 50	< 500	> 1000
Response Time	< 50 ms	< 50 ms	< 50 ms

The above OLTP database workload classification is subjective and derived based on our assumptions. If customers have different requirements, the classifications must be re defined based on workload characteristics.

Using the classifications in Table 1 as a baseline, the goal is to design and optimize the virtual machine configuration to meet the small, medium and large database workloads.

Guidelines for configuring the virtual machines for small, medium and large workloads on the Dell Active System 800 platform

The following sections provide guidelines on hardware selection while configuring the Virtual machine to meet the workloads defined in the Table1 under converged virtual platform.

For each category of workload defined previously, the goal is to determine and optimize the virtual machine configuration to achieve the corresponding workload characteristics.

One virtual machine was created for each small, medium and large workload. To derive and optimize virtual machine configuration, each virtual machine was stressed with different combinations of hardware resources to achieve the corresponding workload characteristics defined in Table 1.

The following section provides recommendations on hardware selection.

Processors: The selection of processors for a given workload plays an important role because you do not want to over provision the processors for any virtual machine. In the process of evaluating number of vCPUs required for a given a VM, the VM was stressed, making sure that the CPU usage is at least in the range of 40 to 50 percent making sure at the same time that their corresponding workload characteristics are achieved.

Memory: Knowing the number of vCPUs assigned to the virtual machine, size of the underlying databases, user load, workload type and query response time helps us in deciding the memory requirements of a given virtual machine. To begin, the memory assigned is roughly 10 percent of database size to each virtual machine. The memory settings were adjusted to improve the vCPU usage to at least 40 to 50 percent and reduce the IO stress (%idle time to around 20 to 30 percent) on disks. For each of the derived virtual configurations, the memory configuration was fixed at the point beyond which we did not observe any significant improvement in performance (TPS and Response Time).

Storage: In our test configuration, pass-through disks were used to store the databases on each guest machine as the pass-through disks allows the guest operating system to bypass the Hyper-V host file system and access the disk directly. For each virtual machine, two pass-through disks were used; one for storing user databases and other for a Tempdb. As Dell EqualLogic PS6110XV is virtualized storage, the IO for a given LUN was served by all 24 drives.

Converged Network: The ETS bandwidth allocations were appropriately configured based on the LAN vs. SAN bandwidth requirements. In the default configuration, we allocated 50-50 percent network bandwidth for LAN and SAN traffic. Customers may have to observe closely the LAN vs. SAN traffic and then allocate bandwidth percentages. The LAN traffic includes public, private, Live Migration traffics, and so on. Customers should reconsider network band allocations if they have different networking requirements. Make sure that Jumbo frames (MTU) are set on servers, switches and storage.

Results and observations from the virtual machine configuration for each classification

In this section we discuss the resulted virtual machine configurations for each classified workload.

The detailed test configuration that was used to deploy and analyze the virtual machines for each of the workloads is listed in Table 2.

Table 2. Test configuration

Hardware/Software Component	Details
Servers	2* PowerEdge M620 blades
CPUs	2* Intel® Xeon® E5-2660 @2.2GHz, 8-Core per Server
Memory	128G @1600MHz(RDIMMs)
Network Adaptors	1* Broadcom BCM57810-k dual port 10GbE network Adapter for Fabric A
Network IO Modules	2* PowerEdge M I/O Aggregators in Fabrics A1 & A2
Top-of-Row switches	2* Dell Force10 S4810 with a VLT domain
Storage	1* Dell EqualLogic PS6110XV with 24* 300G 15k RPM drives
Hypervisor	Windows 2012 Hyper-V
Operating System	Windows 2012 Data Center Edition
Database	SQL Server 2012 Enterprise Edition
DCB settings	50 to 50 percent bandwidth allocation for two ETS groups (One for LAN and other for SAN traffic)
Test Tool Used	Dell Quest Benchmark factory
Test Workload	OLTP (TPC-C standard workload)

Based on the database size requirements for each class of workload described in Table1, different-sized LUNs were created for storing the databases of workload categories described in the following Table 3.

Table 3. Database LUNs used for each workload

Parameters	Small Workload	Medium Workload	Large Workload
Drive size for User Database	40GB	150GB	500GB
Drive size for TempDB	15GB	70GB	150GB

The virtual machine configurations are derived from our tests based on assumptions discussed in Table1. Consider the future growth of the databases when capacity planning for your databases.

Observations

The following section provides details on the resultant virtual machine configuration and the resulted performance numbers during our test exercises.

Table 4 describes the virtual machine configurations, database sizes and actual user loads that we used to stress each virtual machine. Table 4 also shows the resulted TPS, response time, CPU and disk usage.

Table 4. Observed virtual machine configurations and resulted behavior

Parameters	Small Workload	Medium Workload	Large Workload
vProcs	#1	#2	#6
Memory	2GB	10GB	20GB
Operating System	Windows Server 2012	Windows Server 2012	Windows Server 2012
Database	SQL Server 2012	SQL Server 2012	SQL Server 2012
User Database Size	20GB	80GB	300GB
Concurrent Users/connections	45	450	1500
Transactions / Sec (TPS)	46	460	1580
Response Time (RT)	18 ms	25 ms	18 ms
vCPU usage	10 to 15 percent	40 to 50 percent	45 to 50 percent

By using the virtual machine configurations, the TPS required by each workload with in the accepted range of response time (<50 ms) was achieved. Attempts were made using the static and dynamic memory allocation feature of Hyper- V3.0. No significant performance difference was observed while using the Dynamic memory setting.

Table 4 provides the baseline numbers of small, medium and large virtual machines. Based on the workload requirements (DB size, TPS, RT, and so on), customers can leverage the baseline virtual machine configurations to design and deploy their database virtual machines quickly using the Dell Active System 800 platform.

Conclusion

Data center virtualization is undoubtedly one of the efficient ways to prevent the under-usage of processing resources in the environment. Converged architecture provides ease and flexibility on infrastructure deployments.

Establishing a set of standard baseline database virtual configurations helps with understanding the performance tuning guidelines for the specific workload. This also helps in formulating a set of standard deployment processes and maintenance procedures to cater to different workload categories.

The baseline performance numbers for each of the categories is helpful for understanding the virtualization capability of the infrastructure and to get an understanding of the behavior of the entire system.

The Dell Active system 800 platform provides flexibility and ease in achieving highly-efficient virtualized solutions. The major architectural components include Dell PowerEdge Servers, Dell EqualLogic Storage arrays, Dell Force10 10GbE switches, Hyper- V3.0, Microsoft SQL server 2012 Database and Dell Active System Manager. These components blend into fully converged reference architecture to provide the best of advantages for a fully virtualized database environment.

Table 5 describes the workload classification that we derived based on our assumptions for the proposed converged virtualized platform and it also shows the results of the virtual machine configurations for each class of workload.

Table 5. Workload Classification derived on our assumptions

Parameters	Small Workload	Medium Workload	Large Workload
vProcs	#1	#2	#6
Memory	2GB	10GB	20GB
Database Size	20GB	80GB	300GB
Concurrent Users/connections	45	450	1500
Transactions / Sec	46	460	1580
Response Time	18 ms	25 ms	18 ms

The virtual machine configurations and the corresponding performance characteristics help establish a baseline in designing the virtual database deployments.

To summarize, the Dell Converged Architecture solutions provides the latest technologies in the industry to have a fully converged virtualized platform for high efficiency. For any type of database workload, it is important to devise a set of design guidelines and recommendations to design an optimally performing database virtual machine.