Reference Architecture: Microsoft SQL Server

Last update: 13 June 2019
Version 1.1

Describes reference architecture for Microsoft SQL Server using local and shared storage

Contains performance data and sizing recommendations

Includes deployment details and best practices

Contains detailed bill of materials for servers, storage, and network switches

Mike Perks

Dan Ionut Ghervase

Vinay Kulkarni

Lin Xu

Weixu Yang

Table of Contents

1 Introduction

This document describes the reference architecture for the Microsoft SQL Server 2017 using Lenovo® ThinkSystem servers, storage and network switches. The intended audience is IT professionals, technical architects, sales engineers, and consultants to assist in planning, designing, and implementing Microsoft SQL Server.

Microsoft has announced the end of life for Microsoft SQL Server 2008. This document can help address what is needed to migrate existing SQL Server databases to a modern Lenovo infrastructure. It can also be used for other migration efforts or green-field deployments.

This document provides an overview of the business problem and business value that is addressed by Microsoft SQL Server. A description of customer requirements is followed by an architectural overview of the solution and a description of the logical components. The operational model describes the architecture for deploying in small to very large Enterprises. Performance and sizing information is provided with the best practices for deploying Microsoft SQL Server. The last section features detailed Bill of Materials configurations for Lenovo ThinkSystem servers, storage, switches that are used in the solution.

2 Business problem and business value

The following section provides a summary of the business problems that this reference architecture is intended to help address, and the value that this solution can provide.

2.1 Business problem

As one of the fastest growing database platforms, Microsoft SQL Server deployments are becoming increasingly critical to organizations. They are used in everything from departmental databases to business-critical workloads, including enterprise resource planning, customer relationship management and business intelligence. At the same time, enterprises need to consolidate their datacenter footprint with less servers, control costs and accelerate provisioning.

With the announcement of the end of life of Microsoft SQL Server 2008, enterprises are looking for more modern infrastructure to migrate transactional (OLTP) databases and data warehouses (DW). These trends of delivering SQL Server databases as dynamic, consolidated services make it essential to select the right server, storage, and networking architecture.

2.2 Business value

Database performance has long been the primary criteria for selecting infrastructure. Multicore processors and large system memory capacity provide the capability to run large, performance intensive database applications as well as provide the capability to run both multiple databases and multiple instances of Microsoft SQL Server on a single compute server.

Lenovo recommends using the hardware platform presented in this Reference Architecture to migrate off Microsoft SQL Server 2008 to SQL Server 2017. Lenovo ThinkSystem servers offer the best-in-class reliability and performance to provide the largest consolidation of databases, which will both perform well and offer the lowest downtime.

Storage solutions that support Microsoft SQL Server need to handle a dynamic mix of OLTP databases and data warehouses, along with their unique storage I/O profiles and active datasets. This requires efficiently delivering random and sequential read/write at high performance, across sizable amounts of active or hot data. This is achieved using the Lenovo ThinkSystem DM series storage that employs the latest generation of solid state drives to provide high IOPs, low latency, and high storage capacity for SQL Server databases.

3 Requirements

This section descirbes the functional and non-functional requirements for this reference architecture.

3.1 Functional requirements

Table 1 lists the functional requirements of a database management system (database) such as Microsoft SQL Server.

Table 1: Functional requirements

Requirement name	Description
Stores any kind of data	A database management system should be able to store any kind of data including binary such as images and video.
Support ACID Properties	Database must support ACID (Accuracy, Completeness, Isolation, and Durability) properties.
Represents complex relationships between data	Database must represent the complex relationships between data to make the efficient and accurate use of data.
Database schema	Database must provide a method to create and maintain the database schema using both GUI and command line.
Database operations	Database must provide a method to submit SQL queries and return results using GUI, command line, and other interfaces.
Reporting	Database should provide a method to generate formatted reports in various file formats, on-screen or printed
Multiple views	Depending on role, users may see different views of the data.
Concurrent use	Database must be able to respond to multiple requests at a time from multiple sources.

3.2 Non-functional requirements

Table 2 lists the non-functional requirements that are needed for deployment.

Table 2: Non-functional requirements

Requirement name	Description
Data integrity	Integrity ensures the quality and reliability of database system
High availability	This critical part of IT infrastructure must always be available
Disaster recovery	Provide ability for secondary data center to take over if the primary data center suffers a catastrophic failure and all components fail
Scalability	Solution components such as compute and storage capacity and performance scale with an increase in number of concurrent users or transactions
Security	Solution provides ways to secure data based on authorized role
Ease of installation	Reduced complexity of database deployment
Ease of management/operations	Simple management of infrastructure including support for rolling upgrades of hardware and software
Backup/Recovery	Solution support for integrated backup

Reference Architecture: Microsoft SQL Server
Version 1.1

4 Architectural overview

Microsoft SQL Server is a database platform for large-scale online transaction processing (OLTP), data warehouses (DW), and a business intelligence platform for data integration, analysis, and reporting solutions. It uses a common set of tools to deploy and manage databases for in-house and cloud environments.

Figure 1 shows an architectural overview of Microsoft SQL Server on Lenovo ThinkSystem SR650 with Lenovo ThinkSystem DM5000F unified flash storage.
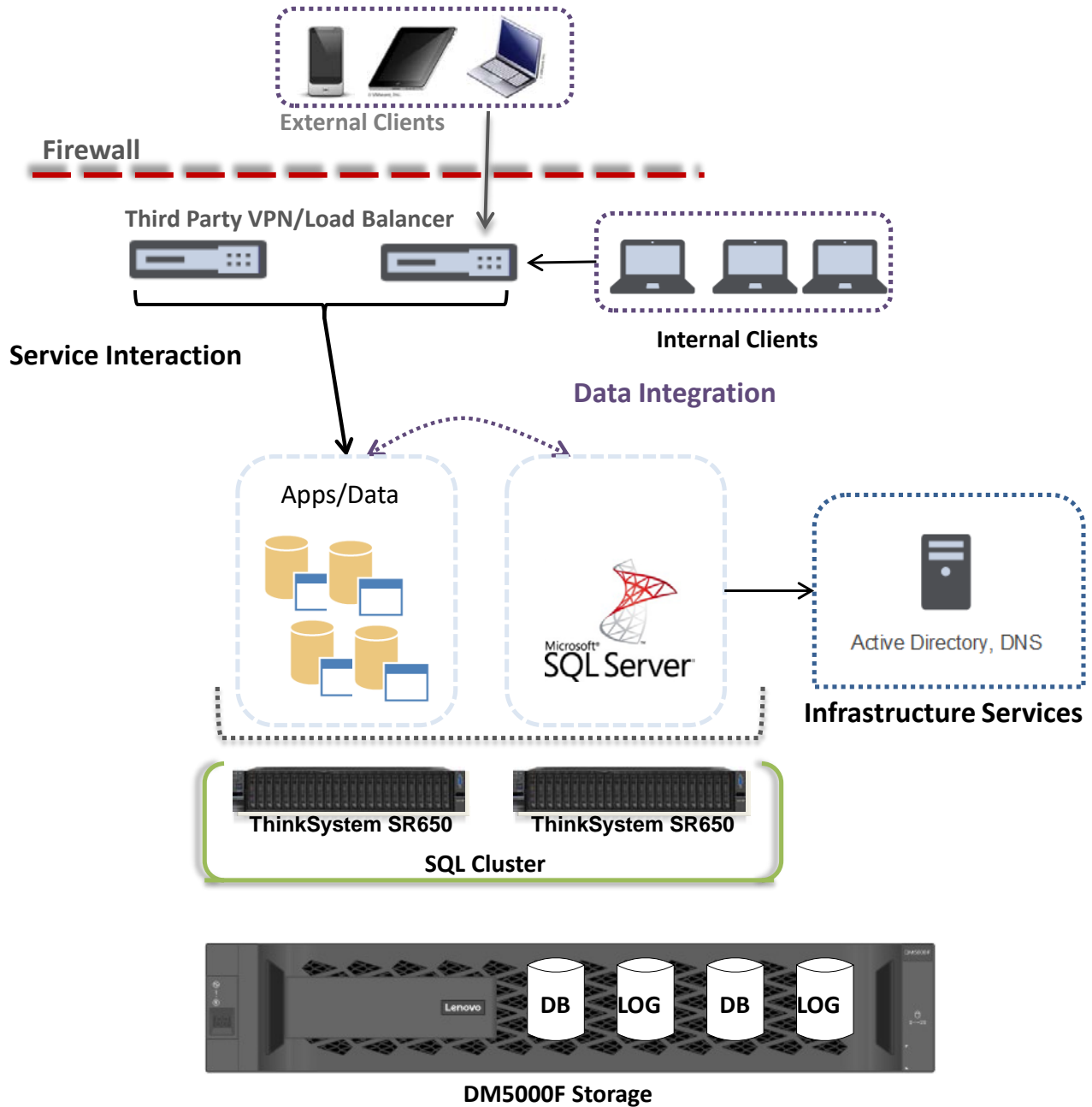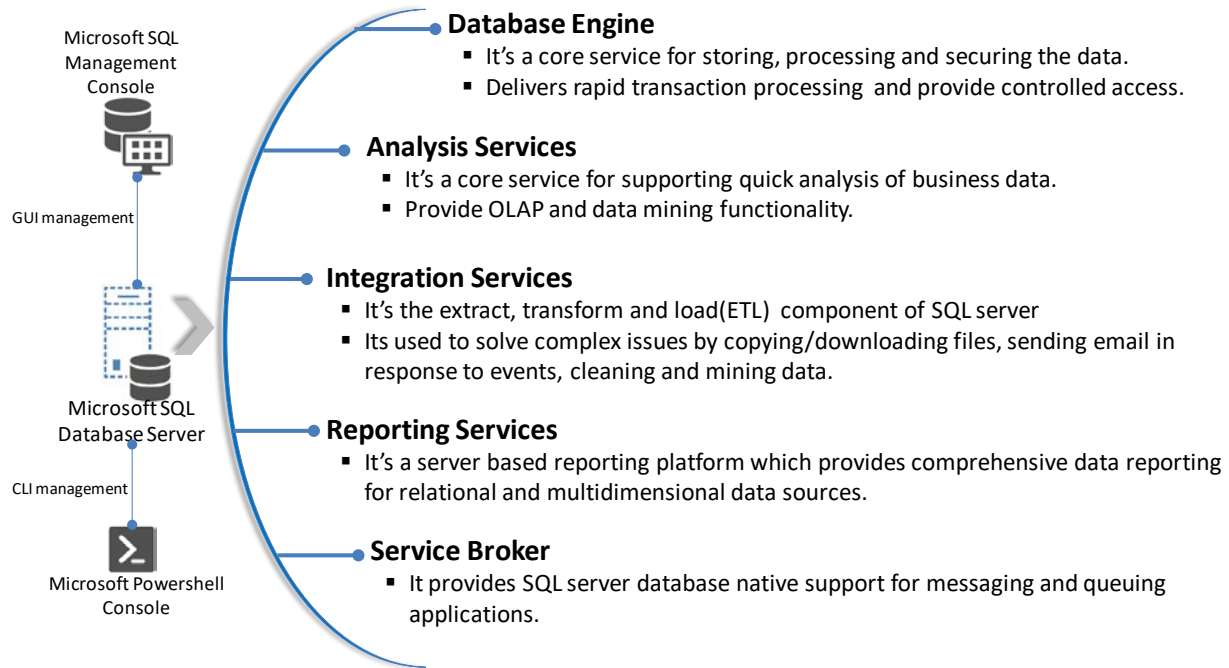

Figure 1: Microsoft SQL Server on Lenovo ThinkSystem SR650 cluster with DM5000F storage

5 Component model

This chapter describes the main software components for the solution. Figure 2 is a layered component view for Microsoft SQL Server.

Database Engine
- It's a core service for storing, processing and securing the data.
- Delivers rapid transaction processing and provide controlled access.

Analysis Services
- It's a core service for supporting quick analysis of business data.
- Provide OLAP and data mining functionality.

Integration Services
- It's the extract, transform and load(ETL) component of SQL server
- Its used to solve complex issues by copying/downloading files, sending email in response to events, cleaning and mining data.

Reporting Services
- It's a server based reporting platform which provides comprehensive data reporting for relational and multidimensional data sources.

Service Broker
- It provides SQL server database native support for messaging and queuing applications.

Microsoft SQL Management Console

GUI management

Microsoft SQL Database Server

CLI management

Microsoft Powershell Console

Figure 2: Component model with Microsoft SQL Server

Microsoft SQL Server features the following main components:

Database Engine	This part of SQL Server actually creates and drives relational databases.
Analysis Services	SQL Server Analysis Services (SSAS) is the data analysis component of SQL Server. It can create OLAP (On Line Analytical Processing) cubes — sophisticated programming objects for organizing data inside a relational database — and do data mining (pulling relevant data out of a database in response to an ad-hoc question).
Integration Services	SQL Server Integration Services (SSIS) performs the extract-transform-load (ETL) process that cleans up and formats raw data from source systems for inclusion in the database as ready-to-use information.
Reporting Services	SQL Server Reporting Services (SSRS) provides reporting regardless of a database's operating system.
Service Broker	SQL Server Service Broker provides native support for messaging and queuing applications which makes it easier to build distributed and reliable applications that use the Database Engine components.

Other software components such as Lenovo XClarity Administrator are not shown.
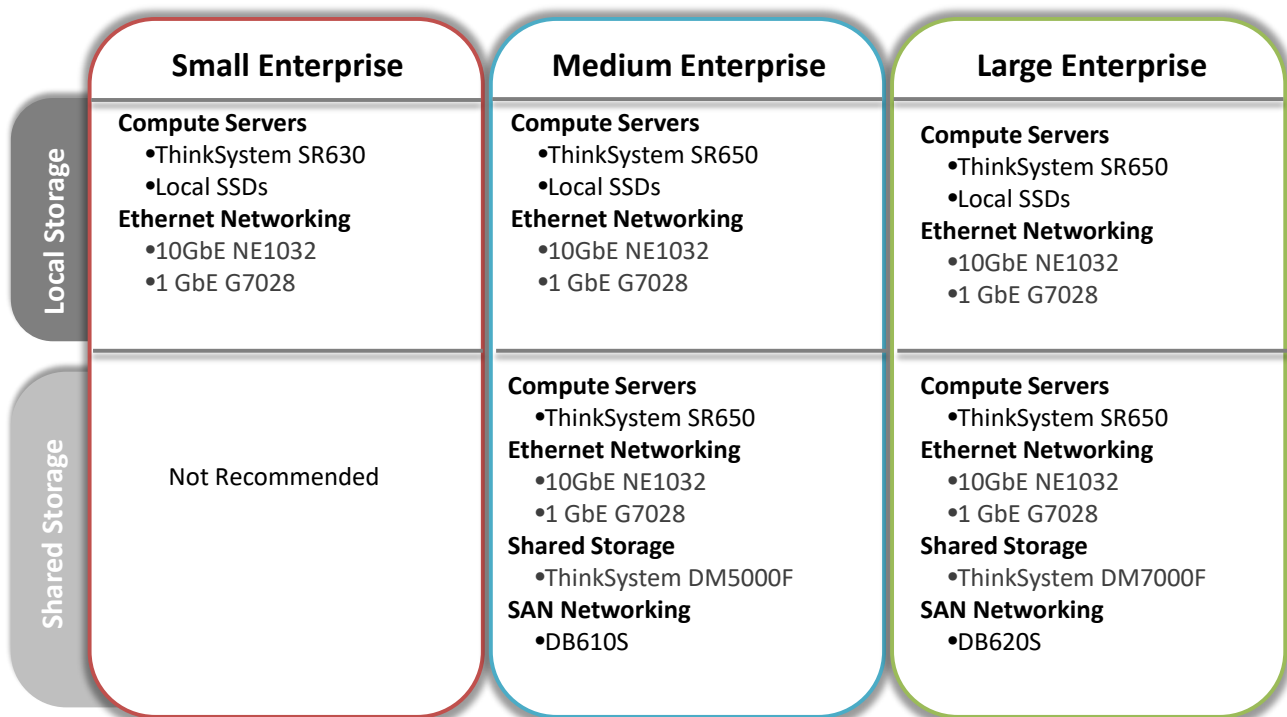
6 Operational model

This section describes the mapping of the logical components of Microsoft SQL Server onto Lenovo servers, storage, and networking. The BOM configurations for the hardware are described in chapter 8 on page 24.

6.1 Operational model scenarios

This reference architecture describes five operational model scenarios that can be grouped as either with local storage or using SAN shared storage. For each type of storage there are variations for small, medium, and large enterprise which equate to increasing SQL loads and performance. Lenovo recommends using local storage only for small enterprise deployments. There are also separate configurations that are tuned to transaction processing (OLTP) versus data warehouses (DW) and analytical processing (OLAP.

Figure 3 summarizes the five configurations.

	Small Enterprise	Medium Enterprise	Large Enterprise
Local Storage	**Compute Servers** •ThinkSystem SR630 •Local SSDs **Ethernet Networking** •10GbE NE1032 •1 GbE G7028	**Compute Servers** •ThinkSystem SR650 •Local SSDs **Ethernet Networking** •10GbE NE1032 •1 GbE G7028	**Compute Servers** •ThinkSystem SR650 •Local SSDs **Ethernet Networking** •10GbE NE1032 •1 GbE G7028
Shared Storage	Not Recommended	**Compute Servers** •ThinkSystem SR650 **Ethernet Networking** •10GbE NE1032 •1 GbE G7028 **Shared Storage** •ThinkSystem DM5000F **SAN Networking** •DB610S	**Compute Servers** •ThinkSystem SR650 **Ethernet Networking** •10GbE NE1032 •1 GbE G7028 **Shared Storage** •ThinkSystem DM7000F **SAN Networking** •DB620S

Figure 3: Summary of operational models for Microsoft SQL Server

Section 6.2 "Hardware components" gives an overview of the Lenovo recommended hardware for deploying Microsoft SQL Server. See the following sections for more information on the recommended configurations:

- 6.3 Compute servers
- 6.3.2 DW/OLAP performance

The performance of the DW/OLAP configurations was tested using a Microsoft proprietary load generation tool called the Fast Track Data Warehouse (FTDW) kit and the results are submitted to Microsoft and validated under the Microsoft Fast Track Data Warehouse program.

The load generation tool was run on the same server as system under test (SUT) making it a host-based workload. The load generation tool uses queries from the TPC-H workload which is a popular decision support benchmark created and maintained by tpc.org.

Each test runs for one hour on Row Store (RS) based database and one hour on Column Store (CS) based database and output metrics are sequential IO table scan rates in terms of Mega Bytes per second (MB/s) for RS database and Queries per hour per Tera Bytes (TB) of data (Queries/hr/TB) for CS database. The tests consist of multiple execution streams running queries in a random order.

Table 7 shows the results of the FTDW tests averaged over the 1-hour periods

Table 7: Results for data warehouse scenarios

Scenario	Users	Database size (TB)	Row Store (MB/s)	Column Store (Queries/Hr/TB)
Small enterprise with local storage	12	9	2,695	1,269
Medium enterprise with shared storage	24	65	5,426	2,272
Medium enterprise with local storage	24	50	9,585	3,727
Large enterprise with local storage	48	90	14,757	4,679

The following conclusions can be drawn from these results:

1. As expected, the I/O bandwidth and query throughput rates increases with the more powerful components in the medium and large enterprise configurations.

2. High availability (HA) scenarios were not tested on the data warehouse configurations. Windows Server failover clustering could be used for HA in the shared storage configuration with minimal performance impact.

The CPU utilization is driven to at least 90% during these tests. SQL Server max server memory setting can be raised up to 90% of server memory to completely cache the database in memory and take advantage of in-memory optimizations in the Microsoft SQL Server product. The results shown above may not be reproducible in your own environment and provide guidance only on what a particular configuration is capable of delivering.

Sizing a data warehouse implementation can be based on size of the database or performance metrics in terms of measured IO throughput for Row Store-based implementations and Query throughput for Column Store-based implementations.

For more information on the Lenovo results for Fast Track Data Warehouse program, see this website: lenovo.com/us/en/data-center/solutions/database/.

- Ethernet networking
- 6.5 SAN storage
- 6.6 SAN networking
- 6.7 Systems management

All configurations utilize Windows Server 2016 Standard Edition and SQL Server 2017 Standard Edition. It is also possible to use SQL Server 2017 Enterprise Edition.

6.2 Hardware components

The following section describes the Lenovo recommended hardware components in a Microsoft SQL Server deployment.

6.2.1 Servers

The following Lenovo ThinkSystem servers are recommended:

- Lenovo ThinkSystem SR650
- Lenovo ThinkSystem SR630

Lenovo ThinkSystem SR650

The Lenovo ThinkSystem SR650 server (as shown in Figure 4) is an enterprise class 2U two-socket versatile server that incorporates outstanding reliability, availability, and serviceability (RAS), security, and high efficiency for business-critical applications and cloud deployments. Unique Lenovo AnyBay technology provides the flexibility to mix-and-match SAS/SATA HDDs/SSDs and NVMe SSDs in the same drive bays. Four direct-connect NVMe ports on the motherboard provide ultra-fast read/writes with NVMe drives and reduce costs by eliminating PCIe switch adapters. Plus, storage can be tiered for greater application performance, to provide the most cost-effective solution. ThinkSystem SR650 server support up to 1.5TB of TruDDR4 Memory currently, and up to 3TB of TruDDR4 Memory in the near future. Its on-board Ethernet solution provides 2/4 standard embedded Gigabit Ethernet ports and 2/4 optional embedded 10 Gigabit Ethernet ports without occupying PCIe slots.

Combined with the Intel® Xeon® Scalable processors product family, the Lenovo ThinkSystem SR650 server offers a high density of workloads and performance that is targeted to lower the total cost of ownership (TCO) per VM. Its flexible, pay-as-you-grow design and great expansion capabilities solidify dependability for any kind of virtualized workload, with minimal downtime. Additionally, it supports two 300W high-performance GPUs and ML2 NIC adapters with shared management.

The Lenovo ThinkSystem SR650 server provides internal storage density of up to 100 TB (with up to 26 x 2.5-inch drives) in a 2U form factor with its impressive array of workload-optimized storage configurations. The ThinkSystem SR650 offers easy management and saves floor space and power consumption for the most demanding storage virtualization use cases by consolidating the storage and server into one system.

Figure 4. Lenovo ThinkSystem SR650 (with 24 x 2.5-inch disk bays)

For more information, see the following websites: ThinkSystem SR650 Product Guide

Lenovo ThinkSystem SR630

The Lenovo ThinkSystem SR630 server (as shown in Figure 5) is an ideal 2-socket 1U rack server for small businesses up to large enterprises that need industry-leading reliability, management, and security, as well as maximizing performance and flexibility for future growth. The SR630 server is designed to handle a wide range of workloads, such as databases, virtualization and cloud computing, virtual desktop infrastructure (VDI), infrastructure security, systems management, enterprise applications, collaboration/email, streaming media, web, and HPC. The ThinkSystem SR630 offers up to twelve 2.5-inch hot-swappable SAS/SATA HDDs or SSDs together with up to four on-board NVMe PCIe ports that allow direct connections to the U.2 NVMe PCIe SSDs.

Figure 5. Lenovo ThinkSystem SR630

For more information, see the following websites: ThinkSystem SR630 Product Guide

6.2.2 10 GbE networking

The Lenovo ThinkSystem NE1032 top of rack (ToR) switch is recommended.

Lenovo ThinkSystem NE1032 RackSwitch

The Lenovo ThinkSystem NE1032 RackSwitch (as shown in Figure 6) is a 1U rack-mount 10 Gb Ethernet switch that delivers lossless, low-latency performance with feature-rich design that supports virtualization, high availability, and enterprise class Layer 2 and Layer 3 functionality. The switch delivers line-rate, high-bandwidth switching, filtering, and traffic queuing without delaying data.

Figure 6: Lenovo ThinkSystem NE1032 RackSwitch

The NE1032 RackSwitch has 32x SFP+ ports that support 1 GbE and 10 GbE optical transceivers, active optical cables (AOCs), and direct attach copper (DAC) cables. The switch helps consolidate server and storage networks into a single fabric, and it is an ideal choice for virtualization, cloud, and enterprise workload solutions. For more information, see this website: lenovopress.com/lp0605.

6.2.3 1 GbE networking

The following Lenovo 1GbE ToR switch is recommended:

- Lenovo RackSwitch G7028

Lenovo RackSwitch G7028

The Lenovo RackSwitch G7028 (as shown in Figure 7) is a 1 Gb top-of-rack switch that delivers line-rate Layer 2 performance at an attractive price. G7028 has 24 10/100/1000BASE-T RJ45 ports and four 10 Gb Ethernet SFP+ ports. It typically uses only 45 W of power, which helps improve energy efficiency.

Figure 7. Lenovo RackSwitch G7028

For more information, see the RackSwitch G7028 Product Guide.

6.2.4 SAN Storage

The following Lenovo SAN storage is recommended:

- Lenovo ThinkSystem DM5000F storage array with optional DM240S Expansion Enclosure(s)
- Lenovo ThinkSystem DM7000F storage array with optional DM240S Expansion Enclosure(s)

Lenovo ThinkSystem DM5000F storage array

Lenovo ThinkSystem DM5000F is a unified, all flash entry-level storage system that is designed to provide performance, simplicity, capacity, security, and high availability for medium to large businesses. Powered by the ONTAP software, ThinkSystem DM5000F delivers enterprise-class storage management capabilities with a wide choice of host connectivity options and enhanced data management features. The ThinkSystem DM5000F is a perfect fit for a wide range of enterprise workloads, including big data and analytics, artificial intelligence, engineering and design, enterprise applications, and other storage I/O-intensive applications.

ThinkSystem DM5000F models (as shown in Figure 8) are 2U rack-mount controller enclosures that include two controllers, 64 GB RAM and 8 GB battery-backed NVRAM (32 GB RAM and 4 GB NVRAM per controller), and 24 SFF hot-swap drive bays (2U24 form factor). Controllers provide universal 1/10 GbE NAS/iSCSI or 8/16 Gb Fibre Channel (FC) ports for host connectivity.

A single ThinkSystem DM5000F Storage Array scales up to 144 solid-state drives (SSDs) with the attachment of Lenovo ThinkSystem DM240S 2U24 SFF Expansion Enclosures. Up to 12 DM5000F Storage Arrays can be combined into a clustered system in a NAS environment, or up to 6 DM5000F Storage Arrays can be combined into a clustered system in a SAN environment.

Figure 8: Lenovo DM5000F storage array

The ThinkSystem DM5000F offers the following key features and benefits:

- All-flash array capabilities to meet the demand for higher speed storage and provide higher IOPs and bandwidth with lower power usage and total cost of ownership than hybrid or HDD-based solutions.
- Improved performance and data protection with RAID-DP and RAID-TEC, as well as support for traditional RAID 4.
- Flexible host connectivity to match diverse client needs with support for unified NAS and SAN storage protocols, including 1/10 GbE NAS and iSCSI, and 8/16 Gb Fibre Channel connectivity.
- 12 Gb SAS drive-side connectivity with multipathing with up to 24x 2.5-inch small form factor (SFF) drives in the 2U24 SFF enclosures.
- Rich set of standard storage management functions available at no extra cost, including snapshots, volume copy, quality of service, thin provisioning, compression, deduplication, encryption, disk-based backup, application- and virtual machine-aware backup, quick data recovery, and asynchronous mirroring.
- Optional WORM (write once, read many) (SnapLock) licensed function to reinforce permanence and integrity of stored data and to ensure compliance with applicable regulations.
- Intuitive, web-based GUI for easy system setup and management.
- Lenovo XClarity support for centralized systems management of Lenovo x86 servers, switches, and storage, which provides automated agent-less discovery, inventory, monitoring, and additional platform-specific functions across multiple systems.
- Designed for high availability with redundant hot-swap components, including controllers and I/O modules, power supplies, and non-disruptive firmware upgrades.

For more information, see this website: lenovopress.com/lp0911.

Lenovo ThinkSystem DM7000F storage array

Lenovo ThinkSystem DM7000F is a scalable, unified, all flash mid-range storage system that is designed to provide high performance, simplicity, capacity, security, and high availability for medium to large businesses. Powered by the ONTAP software, ThinkSystem DM7000F delivers enterprise-class storage management capabilities with a wide choice of host connectivity options, flexible drive configurations, and enhanced data management features, including support for NVMe over Fabrics. The ThinkSystem DM7000F is a perfect fit for a wide range of enterprise workloads, including big data and analytics, artificial intelligence, engineering and design, hybrid clouds, and other storage I/O-intensive applications.

ThinkSystem DM7000F models (as shown in Figure 9) are 3U rack-mount controller enclosures that include two controllers, and 256 GB RAM and 16 GB battery-backed NVRAM (128 GB RAM and 8 GB NVRAM per controller). Universal 1/10 GbE NAS/iSCSI or 4/8/16 Gb Fibre Channel (FC) ports and 1/10 GbE RJ-45 ports provide base host connectivity, with an option for additional 1/10 GbE or 40 GbE NAS/iSCSI, or 8/16/32 Gb FC connections with the adapter cards.

A single ThinkSystem DM7000F Storage Array scales up to 384 SFF solid-state drives (SSDs) with the attachment of Lenovo ThinkSystem DM240S 2U24 SFF Expansion Enclosures. Up to 12 DM7000F Storage Arrays can be combined into a clustered system in a NAS environment, or up to 6 DM7000F Storage Arrays can be combined into a clustered system in a SAN environment.

Figure 9: Lenovo DM7000F storage array

The ThinkSystem DM7000F offers the same key features and benefits as the DM5000F but with the following enhancements:

- Support for NVMe over Fabrics to help achieve up to two times higher performance at a half of the latency.
- Additional support for 40 GbE NAS and iSCSI, and 32 Gb Fibre Channel connectivity.

For more information, see this website: lenovopress.com/lp0912.

6.2.5 SAN switches

The following Lenovo SAN switches are recommended:

- Lenovo ThinkSystem DB610S
- Lenovo ThinkSystem DB620S

Lenovo ThinkSystem DB610S

The Lenovo ThinkSystem DB610S FC SAN Switch provides exceptional price/performance value by delivering market-leading 32 Gb Gen 6 Fibre Channel technology and combining flexibility, simplicity, and enterprise-class functionality to meet the demands of growing flash-based storage environments.

Designed to enable maximum flexibility and reliability, the ThinkSystem DB610S is a compact, 1U rack-mount FC switch that offers low-cost access to industry-leading Storage Area Network (SAN) technology while providing "pay-as-you-grow" scalability to meet the needs of an evolving storage environment.

The DB610S FC SAN Switch (as shown in Figure 10) offers 24x SFP+ ports that support 4/8/16/32 Gbps speeds. The DB610S FC SAN switch provides easy integration into the existing SAN environments while realizing the benefits of Gen 6 Fibre Channel connectivity, and the switch offers a rich set of standard features with the options to expand its capabilities as needed.

The DB610S FC SAN Switch features the EZSwitch Setup wizard and can be configured in Access Gateway Mode to simplify deployment. The switch provides full non-blocking performance with Ports On Demand scalability to support SAN expansion and enable long-term investment protection.

Figure 10: Lenovo ThinkSystem DB610S SAN Switch

For more information, see this website: lenovopress.com/lp0582.

Lenovo ThinkSystem DB620S

The DB620S FC SAN Switch (as shown in Figure 11) is similar to the DB610S but offers 48x SFP+ ports that support 4/8/10/16/32 Gbps speeds and 4x QSFP+ ports that support 128 Gbps (4x 32 Gbps) or 4x 4/8/16/32 Gbps speeds.

Figure 11: Lenovo ThinkSystem DB620S SAN Switch

For more information, see this website: lenovopress.com/lp0580.

Reference Architecture: Microsoft SQL Server
Version 1.1

6.3 Compute servers

The configurations of the compute servers is selected based on the usage requirements. Low core count processors have been chosen to minimize the license cost when using the Microsoft SQL Server core licensing model. Table 3 shows the common configurations for each of the different enterprise sizes.

Table 3: Compute server (common configuration)

Common	Small Enterprise	Medium Enterprise	Large Enterprise
Server Type	SR630 1U	SR650 2U	SR650 2U
Servers for OLTP	2 default, max 12	2 default, max 12	2 default, max 16
Servers for DW/OAP	Optional	1 default	1 default
Processor	2x 6128 (12 cores)	2x 6134 (16 cores)	2x 6136 (24 cores)
Memory	96 GB	192 GB	384 GB
Boot Drives (RAID 1)	2x 480GB boot	2x 480GB boot	2x 480GB boot
10GbE Networking	2 port 10GbE LOM	4 port 10GbE LOM	2x Mellanox CX4 NIC

A minimum of two servers are required for OLTP to provide high availability and business continuity in the event of a single server failure. When using local storage, a Microsoft AlwaysOn availability group should be used so that transactions are executed on both database instances. For shared storage, a Microsoft AlwaysOn failover cluster should be used to ensure that in event of a server failure, incomplete transactions are completed on the other server against the shared storage.

Because data warehouses are not usually mission critical, a minimum of only one server is required and in the case of the Small Enterprise configuration it is completely optional.

Table 4 shows the OLTP local storage configurations for each of the different enterprise sizes. Regular S4610 drives can be substituted for larger capacities (up to 24 drives for data).

Table 4: OLTP local storage

OLTP Local Storage	For all enterprise sizes
Server Type	SR650 2U
RAID adapter	930-16i RAID
Drives for DataDB1 Files (RAID10)	6x 1.92TB
Drives for DataDB2 Files (RAID10)	6x 1.92TB
Drives for DB_Backup (RAID5)	4x 1.92TB
Drives for TempDB1 (RAID1)	2x 800GB
Drives for TempDB2 (RAID1)	2x 800GB
Drives for DB Log (RAID1)	2x 1.92TB
Drives for TempDB Log (RAID1)	2x 1.92TB

Table 5 shows the OLTP shared storage configurations for each of the different enterprise sizes. Because all of the data is shared, a RAID adapter is only used locally for hot swappable boot drives and the focus is on the number and speed of the fibre channel I/O adapters.

Table 5: OLTP shared storage

OLTP Shared Storage	Medium Enterprise	Large Enterprise
RAID adapter	930-16i RAID	930-16i RAID
FC dual-port adapter	1x 32Gb	2x 32Gb
Minimum FC connections	1 per SAN switch	2 per SAN switch

Table 6 shows the DW/OLAP local storage configurations for each of the different enterprise sizes. Because the number of disk writes is much smaller, it is generally recommended to use lower endurance drives for data warehouses. Note that NVMe drives are used for performance.

Table 6: DW/OLAP local storage

DW/OLAP Local Storage	Small Enterprise	Medium Enterprise	Large Enterprise
Server Type	SR630 1U	SR650 2U	SR650 2U
RAID adapter	930-16i RAID	930-8i RAID	930-8i RAID
Drives for Log Files	2x 800GB SS530	2x 800GB SS530	2x 1.6TB SS530
Drives for Data	6x 3.84TB S4510	4x 8TB P4510 (NVMe)	8x 8TB P4510 (NVMe)

Table 7 shows the DW/OLAP shared storage configurations for each of the different enterprise sizes. Because all of the data is shared, a RAID adapter is only used locally for hot swappable boot drives and the focus is on the number and speed of the fibre channel I/O adapters.

Table 7: DW/OLAP shared storage

DW/OLAP Shared Storage	Medium Enterprise	Large Enterprise
RAID adapter	530-8i RAID	530-8i RAID
FC dual-port adapter	2x 32Gb	4x 32Gb
Minimum FC connections	2 per SAN switch	4 per SAN switch

The configurations for the shared storage are described in "SAN storage" on page 19.

6.3.1 OLTP performance

This section provides a high-level summary of the results of executing the HammerDB test suite against various server configurations to ascertain the transaction performance. HammerDB is a graphical open source database load testing and benchmarking tool for Linux and Windows to test databases running on any operating system. HammerDB is automated, multi-threaded and extensible with dynamic scripting support. See this website for more details: hammerdb.com.

Table 8 shows the various scenarios that were tested using one instance of Microsoft SQL Server 2017 Enterprise Edition per physical server. Although Lenovo recommends using an AlwaysOn Availability Group (AAG) to provide high availability, performance tests were executed without an AAG to provide a performance comparison.

Table 8: OLTP Test Scenarios

Storage Type	Medium Enterprise 2 x 6134 CPUs (16 cores)	Large Enterprise 2 x 6136 CPUs (24 cores)	Unconstrained licensing 2 x 8276 CPUs (56 cores)
Local SSDs (no AAG)	✓	✓	
Local SSDs (with AAG)	✓	✓	✓
With shared storage in a cluster	✓	✓	✓

Separate load servers were used to simulate the user load on the database. Each OLTP database workload was simulated by running a varying number of users from 200 to 600 users. The most consistent results were with 400 users. Each user was configured to run a varying number of transactions from 300,000 to 400,000 transactions per user.

Each test scenario had a 5 minute ramp up phase, ran in a steady state for 15 minutes and was repeated 3 times. Runs which did not show a very consistent transaction rate during the steady state period were discarded. Not every data point was measured but a sufficient number were tested to be able to find the maximum number of transactions per minute (TPM) and new orders per minute (NOPM).

Table 9 shows the best results from HammerDB for the eight scenarios described in Table 8.

Table 9: HammerDB results for OLTP

Scenario	Users	Transactions per User	Transactions per minute (TPM)	New Orders per minute (NOPM)
Medium enterprise and no AAG	600	320k	3,250,577	705,528
Medium enterprise with AAG	600	320k	2,976,350	648,062
Medium enterprise with shared storage	400	320k	3,127,585	680,273
Large enterprise and no AAG	600	320k	5,144,206	1,108,722
Large enterprise with AAG	600	320k	4,629,881	1,007,119
Large enterprise with shared storage	400	320k	4,869,142	1,057,951
Unconstrained with AAG	200	400k	8,341,185	1,812,939
Uncontained with shared storage	400	400k	9,092,498	1,975,865

In all cases the CPU utilization was between 93% and 95% and less than 384GB of memory was used. The results shown above may not be reproducible in your own environment and provide guidance only on what a particular configuration is capable of.

The following conclusions can be drawn from these results:

- The overhead of using AlwaysOn Availability Groups (AAG) with local storage is approximately 10%. This is significant but required to provide high availability for configurations with local storage.

- The performance with using DM5000F shared storage is slightly better than using local SSDs with AAG because the server has less work. The shared storage provides the high availability instead of the server.

- In the unconstrained case with AAG, the server is using all 56 cores (112 threads) and a steady transaction rate was only achievable by decreasing the number of users to 200. Under normal circumstances, no customer should be executing with SQL Server using 93-98% of the CPU resource.

- The performance scales linearly as the processing power of the system increases. Figure 12 shows the relative performance of the different storage scenarios. The horizontal axis represent the power of the processors in the server which is calculated from multiplying the number of cores by the clock speed. The vertical axis is the number of transactions per minute.
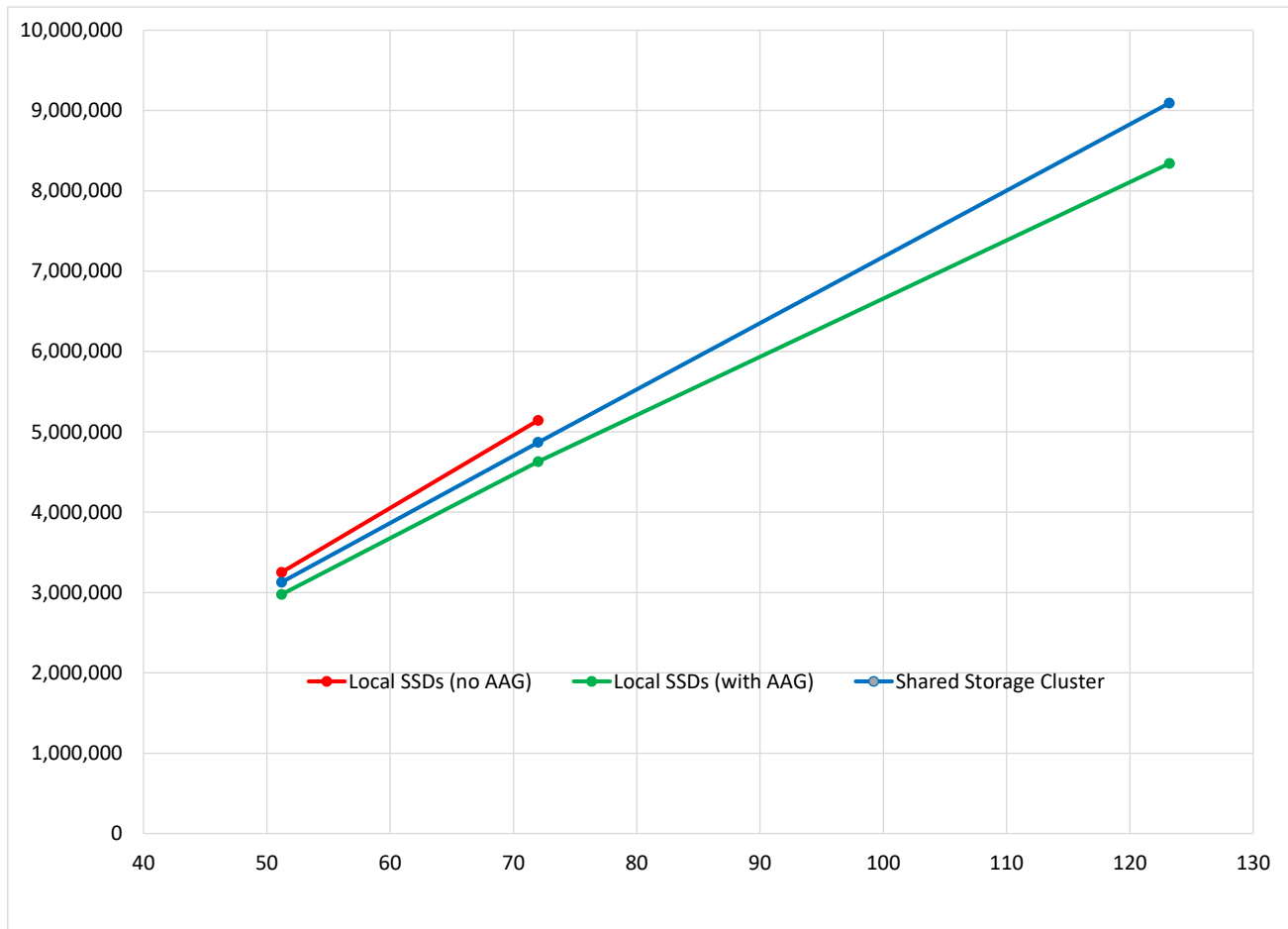
Figure 12: Microsoft SQL Server performance on ThinkSystem SR650 2 socket server

6.3.2 DW/OLAP performance

The performance of the DW/OLAP configurations was tested using a Microsoft proprietary load generation tool called the Fast Track Data Warehouse (FTDW) kit and the results are submitted to Microsoft and validated under the Microsoft Fast Track Data Warehouse program.

The load generation tool was run on the same server as system under test (SUT) making it a host-based workload. The load generation tool uses queries from the TPC-H workload which is a popular decision support benchmark created and maintained by tpc.org.

Each test runs for one hour on Row Store (RS) based database and one hour on Column Store (CS) based database and output metrics are sequential IO table scan rates in terms of Mega Bytes per second (MB/s) for RS database and Queries per hour per Tera Bytes (TB) of data (Queries/hr/TB) for CS database. The tests consist of multiple execution streams running queries in a random order.

Table 7 shows the results of the FTDW tests averaged over the 1-hour periods

Table 7: Results for data warehouse scenarios

Scenario	Users	Database size (TB)	Row Store (MB/s)	Column Store (Queries/Hr/TB)
Small enterprise with local storage	12	9	2,695	1,269
Medium enterprise with shared storage	24	65	5,426	2,272
Medium enterprise with local storage	24	50	9,585	3,727
Large enterprise with local storage	48	90	14,757	4,679

The following conclusions can be drawn from these results:

- As expected, the I/O bandwidth and query throughput rates increases with the more powerful components in the medium and large enterprise configurations.
- High availability (HA) scenarios were not tested on the data warehouse configurations. Windows Server failover clustering could be used for HA in the shared storage configuration with minimal performance impact.

The CPU utilization is driven to at least 90% during these tests. SQL Server max server memory setting can be raised up to 90% of server memory to completely cache the database in memory and take advantage of in-memory optimizations in the Microsoft SQL Server product. The results shown above may not be reproducible in your own environment and provide guidance only on what a particular configuration is capable of delivering.

Sizing a data warehouse implementation can be based on size of the database or performance metrics in terms of measured IO throughput for Row Store-based implementations and Query throughput for Column Store-based implementations.

For more information on the Lenovo results for Fast Track Data Warehouse program, see this website: lenovo.com/us/en/data-center/solutions/database/.

6.4 Ethernet networking

The 10GbE data network is the fabric that carries all inter-node I/O traffic for the Microsoft cluster, in addition to the user data traffic. It is recommended that two top of rack (ToR) switches are used for redundancy. In order to support the logical pairing of the network adapter ports and to provide automatic failover of the switches, the Lenovo ThinkSystem NE1032 supports virtual link aggregation groups (VLAGs). When VLAG is enabled over the inter-switch link (ISL) trunk, it enables logical grouping of these switches. When one of the switches is lost, or the uplink from the host to the switch is lost, the connectivity is automatically maintained over the other switch.

The 1GbE network is used for management of the Lenovo servers and storage. The Lenovo RackSwitch G7028 can be used for this network.

Table 10 shows the networking configurations recommended for each of the different enterprise sizes.

Table 10: Configurations for Ethernet networking

Ethernet Networking	Small Enterprise	Medium Enterprise	Large Enterprise
10 GbE data	1 x NE1032	2 x NE1032	2 x NE1032
1 GbE management	1 x G7028	1 x G7028	1 x G7028

6.5 SAN storage

Using SAN shared storage for SQL Server requires high availability of OLTP compute servers so that incomplete transactions can be fully executed. For high availability, the log files, tmp files, and database instances are all stored on the SAN storage.

Table 12 shows the SAN storage configurations recommended for each of the different enterprise sizes.

Table 11: Configurations for Fibre Channel networking

SAN Storage	Medium Enterprise	Large Enterprise
Storage Array	ThinkSystem DM5000F	ThinkSystem DM7000F
Storage Expansion Units	Optional	Minimum of 2 x DM240S
Drives for OLTP storage	12 x 3.84TB SSD	24 x 3.84TB SSD
Drives for DW/OLAP storage	12 x 3.84TB SSD	24 x 3.84TB SSD
Minimum FC connections	2 per SAN switch	4 per SAN switch

The number and capacity of the SSDs used in shared storage is dependent on the number of SQL instances and backups that are required. The configurations above provide "average" capacities. More DM240S storage expansion units can be added to give additional capacity. The best method to increase performance is to cluster more than one DM storage array which is outside the scope of this Reference Architecture.

6.6 SAN networking

The Fibre Channel SAN network carries all of the data traffic in and out of the SAN shared storage. Lenovo recommends using two SAN switches for high availability. Table 12 shows the SAN networking configurations recommended for each of the different enterprise sizes.

Table 12: Configurations for Fibre Channel networking

SAN Networking	Medium Enterprise	Large Enterprise
Fibre Channel switch	2 x ThinkSystem DB610S	2 x ThinkSystem DB620S

The choice of switch is somewhat dependent on the number of Fiber Channel connections that are needed which in turn depends upon the number and type of servers that are being used.

6.7 Systems management

Lenovo XClarity™ Administrator is a centralized resource management solution that reduces complexity, speeds up response, and enhances the availability of Lenovo® server systems and solutions.

The Lenovo XClarity Administrator provides agent-free hardware management for Lenovo's ThinkSystem® rack servers, System x® rack servers, and Flex System™ compute nodes and components, including the Chassis Management Module (CMM) and Flex System I/O modules. Figure 13 shows the Lenovo XClarity administrator interface, in which Flex System components and rack servers are managed and are seen on the dashboard. Lenovo XClarity Administrator is a virtual appliance that is quickly imported into a virtualized environment server configuration.
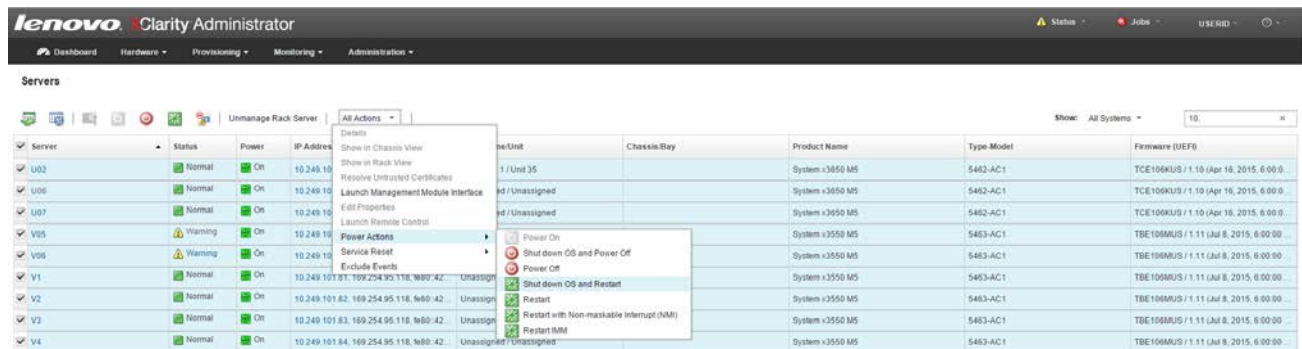
Figure 13: XClarity Administrator interface

7 Deployment considerations

This chapter contains additional considerations for deploying Microsoft SQL Server.

7.1 Sizing considerations

There are several considerations for sizing Microsoft SQL Server. The processor cores and clock speed depend on the number of transactions per minute required from the database. The performance results in chapter 6 show the capabilities of a server using two high-end Intel Xeon Scalable processors. If more processing power is required, then the 4 and 8 processor capable Lenovo ThinkSystem SR850 or ThinkSystem SR950 servers could be used. Lenovo recommends leaving 20-30% spare processor capacity to allow for failover when a server fails in the cluster.

Because Microsoft SQL Server is designed to utilize all available memory, Lenovo recommends not to over-provision memory as it leads to paging and thus significantly slowing performance.

The amount of storage to allocate for the database depends on a number of factors including the type of application that is being served by SQL Server. There are many different sizing tools available from Microsoft and other sources. Additional storage capacity should be built in both for growth in the database size and in case of a server failure when the data needs to be rebalanced across the remaining servers in the cluster. Lenovo recommends at least 30-40% extra capacity in the storage.

7.2 Best practices and limitations

This section describes recommended best practices to provide data optimization and high availability of Microsoft SQL Server.

7.2.1 Hardware and operating system configuration

Lenovo recommends the following best practices for configuring the hardware and Windows Server 2016 operating system:

- Update to the latest firmware and driver levels on storage, switches, servers and all components
- Configure the server UEFI settings to set operating mode to maximum performance
- Configure high availability for the operating system boot drives by using RAID 1
- Configure high availability for data files using RAID 5 for the local drives. The ThinkSystem DM storage array automatically provides redundancy.
- Configure high availability for log file local drives using RAID 1 or RAID 10.
- Set the power plan in Windows Server 2016 to high performance: Control Panel > System & Security > Power Options > High Performance
- Enable lock pages in memory option using Windows Group policy tool to prevent paging of data files.
- If the server is dedicated to a single SQL instance then set processor affinity for SQL Server to use all the processors in the system otherwise set the affinity so an instance does not span more than one processor.
- All SQL database and log drives should be formatted with 64KB NTFS cluster size as it enhances the I/O performance without adding any overhead. The operating system and SQL database drives should be formatted with the standard 4KB NTFS cluster size. Drives space utilization should not be above 80% to achieve optimal performance.

7.2.2 Cluster high availability

Microsoft SQL Server AlwaysOn architecture has two methods to provide high availability:

- Always On Failover Cluster instance (FCI) using Windows Server Failover Clustering (WSFC) as the base technology and requires shared storage.
- Always On Availability Groups (AAG) where one or more databases can be configured to fail over together.

Table 13 shows some of the major features and differences between the two methods:

Table 13: Always On Features and Differences

	Always On Failover Clustering	**Always On Availability Groups**
Unit of Failover	Instance	Database or groups of databases
Storage	Shared storage required	No Shared storage
Query and reporting from secondary databases	No	Yes

The Microsoft AlwaysOn is a recommended high availability mechanism for servers with local storage. It uses a shared-nothing approach where transactions are replicated to other nodes so each node has a full copy of the database and transaction logs. This provides a very fast failover mechanism.

Data replication across data centers can also be done using Always On Availability Groups (AAGs) assuming there is sufficient band-width between the sites. The scenarios for active-active and active-passive DR sites using AAGs are outside the scope of this document.

7.2.3 SQL Server

For maximum IO performance, data files holding the SQL Server data and TempDB should be evenly distributed across all data volumes.

On high core count systems, it is beneficial to have multiple data files. For example a 16 core system can have 16 data files. Lenovo recommends to separate the data files into multiple filegroups based on logical separation in the database. Heavily accessed tables can be put in multiple files using a separate filegroup that is distributed evenly on high performance drives. Eight data files distributed across two RAID 10 volumes was used for the Lenovo performance testing.

TempDB is used as scratch space by the applications and is one of the most important factors of SQL performance. The number of TempDB files to be used is based on the core count. If the core count is less than 8, then configure the same number of TempDB files. If the number of cores is higher than 8, then start with 8 TempDB files and monitor the contention for in-memory allocation (PAGELATCH_XX). The number of TempDB file should be increased in increments of four until contention is eliminated. It is recommended to create all TempDBs file with the same size and not allow for autogrowth. The TempDB file sizing should be based on the application and is usually 1-10% of the database size.

The SQL log files for the database and TempDB are written sequentially, so using multiple log files does not improve their performance. Using single log file per database is recommended.

Lenovo recommends adding the "–T834" trace flag to SQL Server startup parameters. When this trace flag is set, SQL Server uses Windows large-page memory allocations for the buffer, pool which improves performance by increasing the efficiency of the translation look-aside buffer (TLB) in the CPU.

7.3 Database consolidation considerations

One of the advantages of more modern infrastructure as presented in this Reference Architecture is the ability to consolidate multiple databases into fewer servers. There are four main approaches:

- Single instance per server
- Multiple instances on a single server
- Using virtualization
- Using containers

A single instance of Microsoft SQL Server per server is the most familiar in that the whole of a server's resources are dedicated to SQL Server. That also means that the environment is fixed for the server which might mean that certain databases need to be in a different instance for regulatory, compliance, or security reasons.

Using multiple instances of Microsoft SQL Server per server provides a balance between isolation of difference database environments and potentially higher utilization of a single server. However multiple instance are competing for server resources which may cause problems under heavy load where there is no enough processing power. The worst case scenario is that RAM is overcommitted and paged to disk. These problems can be ameliorated by configuring the SQL Server instances using CPU affinity and establishing minimum and maximum settings to reduce memory overcommitment and conflicts.

The third method is to virtualize the SQL Server environment in a virtual machine (VM). This allows VMs to be mixed on a single server as well as provide high availability and resource balancing across a cluster of servers. Using virtualization may result in the highest consolidation ratio. Non-production environments such as Development, Test, and Pre-production are particularly suited to the virtual machine approach. Note that it is still necessary to configure CPU affinity for VMs to provide the best performance.

A new feature in Windows Server 2016 is support for containers. In some ways containers provide all of the advantages of both virtual machines and multiple SQL Server instances without any of the drawbacks. This is a new technology that customers are just starting to explore.

Microsoft have published several whitepapers on database consolidation, which should be consulted for more information and best practices:

- [SQL Consolidation Planning and recommended practices](#)
- [SQL Server Consolidation Guidance](#)

8 Appendix: Bill of Materials

This appendix features the Bill of Materials (BOMs) for different configurations of hardware for Microsoft SQL Server deployments on Lenovo hardware. There are sections for servers, storage and networking switches that are orderable from Lenovo.

The BOM lists in this appendix are not meant to be exhaustive and must always be confirmed with the configuration tools. For connections between ToR switches and devices (servers, storage, and chassis), the connector cables are configured with the device. The ToR switch configuration includes only transceivers or other cabling that is needed for failover or redundancy.

8.1 Server BOMs for OLTP

The following tables list the server BOMs for online transaction processing.

Table 14: ThinkSystem SR630 (OLTP small enterprise with local storage)

Code	Description	Quantity
7X02CTO1WW	SELO : ThinkSystem SR630 - 3yr Warranty	1
AUW1	ThinkSystem SR630 2.5" Chassis with 10 Bays	1
AWEB	Intel Xeon Gold 6128 6C 115W 3.4GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AUW9	ThinkSystem SR630/SR570 2.5" AnyBay 10-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
B31E	ThinkSystem RAID 930-16i 8GB Flash PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
A2KL	Secondary Array - RAID 10	1
B49Q	ThinkSystem 2.5" Intel S4610 3.84TB Mainstream SATA 6Gb Hot Swap SSD	6
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
B11V	ThinkSystem M.2 5100 480GB SATA 6Gbps Non-Hot Swap SSD	2
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AVWA	ThinkSystem 750W (230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCB	ThinkSystem Toolless Slide Rail Kit with 1U CMA	1
A1PJ	3m Passive DAC SFP+ Cable	2
AUW3	Lenovo ThinkSystem Mainstream MB - 1U	1
AUNP	FBU345 SuperCap	1
2302	RAID Configuration	1

Table 15: ThinkSystem SR650 (OLTP medium enterprise with local storage)

Code	Description	Quantity
7X06CTO1WW	MELO : ThinkSystem SR650 - 3yr Warranty	1
AUVV	ThinkSystem SR650 2.5" Chassis with 8, 16 or 24 bays	1
AWE9	Intel Xeon Gold 6134 8C 130W 3.2GHz Processor	2
AUNB	ThinkSystem 16GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	2
AUR5	ThinkSystem 2U/Twr 2.5" AnyBay 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNJ	ThinkSystem RAID 930-8i 2GB Flash PCIe 12Gb Adapter	1
B31E	ThinkSystem RAID 930-16i 8GB Flash PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
A2KL	Secondary Array - RAID 10	1
B49Q	ThinkSystem 2.5" Intel S4610 3.84TB Mainstream SATA 6Gb Hot Swap SSD	16
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
B11V	ThinkSystem M.2 5100 480GB SATA 6Gbps Non-Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
A1PJ	3m Passive DAC SFP+ Cable	4
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
2302	RAID Configuration	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUNP	FBU345 SuperCap	2
B31F	ThinkSystem M.2 480GB SSD Thermal Kit	1

Table 16: ThinkSystem SR650 (OLTP large enterprise with local storage)

Code	Description	Quantity
7X06CTO1WW	LELO : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE3	Intel Xeon Gold 6136 12C 150W 3.0GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUR5	ThinkSystem 2U/Twr 2.5" AnyBay 8-Bay Backplane	2
5978	Select Storage devices - configured RAID	1
B31E	ThinkSystem RAID 930-16i 8GB Flash PCIe 12Gb Adapter	1
AUV2	ThinkSystem 1610-4P NVMe Switch Adapter	1
B58B	ThinkSystem U.2 Intel P4610 6.4TB Mainstream NVMe PCIe3.0 x4 HS SSD	8
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
A2KG	Secondary Array - RAID 1	1
B4Y6	ThinkSystem 2.5" SS530 1.6TB Performance SAS 12Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AUKX	ThinkSystem Intel X710-DA2 PCIe 10Gb 2-Port SFP+ Ethernet Adapter	2
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
7015	Secondary Array 2 HDDs	1
7008	Primary Array 2 HDDs	1
B0ML	Feature Enable TPM on MB	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AUSZ	ThinkSystem SR650 Service Label LI	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUNP	FBU345 SuperCap	1

Table 17: ThinkSystem SR650 (OLTP medium enterprise with shared storage)

Code	Description	Quantity
7X06CTO1WW	MESO : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE9	Intel Xeon Gold 6134 8C 130W 3.2GHz Processor	2
AUNB	ThinkSystem 16GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNG	ThinkSystem RAID 530-8i PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AUNV	ThinkSystem Emulex LPe32002-M2-L PCIe 32Gb 2-Port SFP+ Fibre Channel	1
AVWD	ThinkSystem 750W(230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
A1PJ	3m Passive DAC SFP+ Cable	2
ASR7	Lenovo 3m LC-LC OM3 MMF Cable	2
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
7008	Primary Array 2 HDDs	1
B0ML	Feature Enable TPM on MB	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1

Reference Architecture: Microsoft SQL Server
Version 1.1

Table 18: ThinkSystem SR650 (OLTP large enterprise with shared storage)

Code	Description	Quantity
7X06CTO1WW	LESO : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE3	Intel Xeon Gold 6136 12C 150W 3.0GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNG	ThinkSystem RAID 530-8i PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AUNV	ThinkSystem Emulex LPe32002-M2-L PCIe 32Gb 2-Port SFP+ Fibre Channel	2
AUKX	ThinkSystem Intel X710-DA2 PCIe 10Gb 2-Port SFP+ Ethernet Adapter	2
AVWD	ThinkSystem 750W(230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
ASR7	Lenovo 3m LC-LC OM3 MMF Cable	4
A1PJ	3m Passive DAC SFP+ Cable	4
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
7008	Primary Array 2 HDDs	1
B0ML	Feature Enable TPM on MB	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1

8.2 Server BOMs for DW/OLAP

The following tables list the server BOMs for online analytical processing.

Table 19: ThinkSystem SR630 (OLAP small enterprise with local storage)

Code	Description	Quantity
7X02CTO1WW	SELDW : ThinkSystem SR630 - 3yr Warranty	1
AUW1	ThinkSystem SR630 2.5" Chassis with 10 Bays	1
AWEB	Intel Xeon Gold 6128 6C 115W 3.4GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AUW9	ThinkSystem SR630/SR570 2.5" AnyBay 10-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
B31E	ThinkSystem RAID 930-16i 8GB Flash PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
A2KJ	Secondary Array - RAID 5	1
B49C	ThinkSystem 2.5" Intel S4510 3.84TB Entry SATA 6Gb Hot Swap SSD	6
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
B11V	ThinkSystem M.2 5100 480GB SATA 6Gbps Non-Hot Swap SSD	2
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AVWA	ThinkSystem 750W (230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCB	ThinkSystem Toolless Slide Rail Kit with 1U CMA	1
A1PJ	3m Passive DAC SFP+ Cable	2
AUW3	Lenovo ThinkSystem Mainstream MB - 1U	1
AUNP	FBU345 SuperCap	1
2302	RAID Configuration	1

Table 20: ThinkSystem SR650 (OLAP medium enterprise with local storage)

Code	Description	Quantity
7X06CTO1WW	MELDW : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE9	Intel Xeon Gold 6134 8C 130W 3.2GHz Processor	2
AUNB	ThinkSystem 16GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM	12
AUR5	ThinkSystem 2U/Twr 2.5" AnyBay 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNJ	ThinkSystem RAID 930-8i 2GB Flash PCIe 12Gb Adapter	1
B58J	ThinkSystem U.2 Intel P4510 8.0TB Entry NVMe PCIe3.0 x4 Hot Swap SSD	4
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
B11V	ThinkSystem M.2 5100 480GB SATA 6Gbps Non-Hot Swap SSD	2
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
A1PJ	3m Passive DAC SFP+ Cable	4
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
2302	RAID Configuration	1
AUNP	FBU345 SuperCap	1
B31F	ThinkSystem M.2 480GB SSD Thermal Kit	1

Table 21: ThinkSystem SR650 (OLAP large enterprise with local storage)

Code	Description	Quantity
7X06CTO1WW	LELDW : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE3	Intel Xeon Gold 6136 12C 150W 3.0GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUR5	ThinkSystem 2U/Twr 2.5" AnyBay 8-Bay Backplane	2
5978	Select Storage devices - configured RAID	1
B31E	ThinkSystem RAID 930-16i 8GB Flash PCIe 12Gb Adapter	1
AUV2	ThinkSystem 1610-4P NVMe Switch Adapter	1
B58J	ThinkSystem U.2 Intel P4510 8.0TB Entry NVMe PCIe3.0 x4 Hot Swap SSD	8
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
A2KG	Secondary Array - RAID 1	1
B4Y6	ThinkSystem 2.5" SS530 1.6TB Performance SAS 12Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AUKX	ThinkSystem Intel X710-DA2 PCIe 10Gb 2-Port SFP+ Ethernet Adapter	2
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUNP	FBU345 SuperCap	1

Table 22: ThinkSystem SR650 (OLAP medium enterprise with shared storage)

Code	Description	Quantity
7X06CTO1WW	MESDW : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE9	Intel Xeon Gold 6134 8C 130W 3.2GHz Processor	2
AUNB	ThinkSystem 16GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNG	ThinkSystem RAID 530-8i PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AUNV	ThinkSystem Emulex LPe32002-M2-L PCIe 32Gb 2-Port SFP+ Fibre Channel Adapter	2
AVWD	ThinkSystem 750W(230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
A1PJ	3m Passive DAC SFP+ Cable	2
ASR7	Lenovo 3m LC-LC OM3 MMF Cable	4
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
7008	Primary Array 2 HDDs	1
B0ML	Feature Enable TPM on MB	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1

Table 23: ThinkSystem SR650 (OLAP large enterprise with shared storage)

Code	Description	Quantity
7X06CTO1WW	LESDW : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWE3	Intel Xeon Gold 6136 12C 150W 3.0GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
5978	Select Storage devices - configured RAID	1
AUNG	ThinkSystem RAID 530-8i PCIe 12Gb Adapter	1
2499	Install smallest capacity, slower drives starting in Array 1	1
A2K7	Primary Array - RAID 1	1
B49M	ThinkSystem 2.5" Intel S4610 480GB Mainstream SATA 6Gb Hot Swap SSD	2
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUR4	ThinkSystem 2U x8/x8/x8 PCIE FH Riser 1	1
AUNV	ThinkSystem Emulex LPe32002-M2-L PCIe 32Gb 2-Port SFP+ Fibre Channel	4
AUKX	ThinkSystem Intel X710-DA2 PCIe 10Gb 2-Port SFP+ Ethernet Adapter	2
AVWD	ThinkSystem 750W(230/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
B0MJ	Feature Enable TPM 1.2	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
ASR7	Lenovo 3m LC-LC OM3 MMF Cable	8
A1PJ	3m Passive DAC SFP+ Cable	4
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
7008	Primary Array 2 HDDs	1
B0ML	Feature Enable TPM on MB	1
2302	RAID Configuration	1
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1

8.3 SAN storage BOM

Table 24 and Table 25 list the storage array BOMs. Table 26 lists the storage expansion unit BOM (at least two are needed for the DM7000F large enterprise configuration).

Table 24: ThinkSystem DM5000F (medium enterprise)

Code	Description	Quantity
7Y41CTO1WW	MES_DM : Lenovo ThinkSystem DM5000F All Flash Array	1
B38L	Lenovo ThinkSystem Storage 2U24 Chassis	1
B5RJ	DM Series Premium Offering	1
B39F	Lenovo ThinkSystem DM Series DM3000/DM5000 Cntr, 16Gb FC/10Gb	2
B3WK	Lenovo ThinkSystem DM5000F 92TB SSD Pack (24x 3.8TB SSD)	1
A3RG	0.5m Passive DAC SFP+ Cable	2
B4BP	Lenovo ThinkSystem Storage USB Cable, Micro-USB	1
6311	2.8m, 10A/100-250V, C13 to C14 Jumper Cord	2
B4D0	Lenovo ThinkSystem DM Series ONTAP Software, Encryption Capable	1
B39L	Lenovo ThinkSystem DM Series 2U24 Bezel	1

Table 25: ThinkSystem DM7000F (large enterprise)

Code	Description	Quantity
7Y40CTO1WW	LES_DM : Lenovo ThinkSystem DM7000F All Flash Array	1
B38K	Lenovo ThinkSystem DM Series 3U Chassis	1
B5RJ	DM Series Premium Offering	1
B39H	Lenovo ThinkSystem DM Series DM7000 Controller	2
B38U	Lenovo ThinkSystem DM Series 32Gb 2 port Fibre Channel Card	4
B2PB	Lenovo 3m LC-LC OM4 MMF Cable	8
B4BP	Lenovo ThinkSystem Storage USB Cable, Micro-USB	1
6311	2.8m, 10A/100-250V, C13 to C14 Jumper Cord	2
B4D0	Lenovo ThinkSystem DM Series ONTAP Software, Encryption Capable	1
B39M	Lenovo ThinkSystem DM Series 3U Bezel	1
B4CW	Lenovo ThinkSystem DM7000 Accessory	1

Table 26: ThinkSystem DM240S

Code	Description	Quantity
7Y58CTO1WW	Lenovo ThinkSystem DM240S 2U24 SFF Expansion Enclosure	1
B38L	Lenovo ThinkSystem Storage 2U24 Chassis	1
B39J	Lenovo ThinkSystem DM Series DM120S/240S/600S IOM	2
B3WH	Lenovo ThinkSystem DM7000F 92TB SSD Pack (24x 3.8TB SSD)	1
AU17	1m External MiniSAS HD 8644/MiniSAS HD 8644 Cable	2
6311	2.8m, 10A/100-250V, C13 to C14 Jumper Cord	2
B39L	Lenovo ThinkSystem DM Series 2U24 Bezel	1

8.4 Networking BOM

Table 27 and Table 28 list the BOMs for the network switches for 1 GbE and 10 GbE connectivity respectively

Table 27: RackSwitch G7028

Code	Description	Quantity
7159HCP	1G : Lenovo RackSwitch G7028 (Rear to Front)	1
AT09	Lenovo RackSwitch G7028 (Rear to Front)	1
A3KP	Adjustable 19" 4 Post Rail Kit	1
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
A4C0	RackSwitch G7000 1.8m DC to DC RPS Power Cable	1
A4BZ	RackSwitch G7000 Redundant Power Supply	1

Table 28: ThinkSystem NE1032 RackSwitch

Code	Description	Quantity
7159HD1	10G : Lenovo ThinkSystem NE1032 RackSwitch (Rear to Front)	1
AU3A	Lenovo ThinkSystem NE1032 RackSwitch (Rear to Front)	1
A3KP	Adjustable 19" 4 Post Rail Kit	1
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2

8.5 SAN networking BOM

Table 29 and Table 30 list the BOMs for the SAN network switches.

Table 29: ThinkSystem DB610S

Code	Description	Quantity
6559HC1	Lenovo ThinkSystem DB610S 24-p 24x16Gb SWL SFP, Enterprise SW	2
AVPQ	Lenovo DB610S 32Gb FC SAN Switch 24-port 24x16Gb SWL SFP	2
6311	2.8m, 10A/100-250V, C13 to C14 Jumper Cord	2

Table 30: ThinkSystem DB620S

Code	Description	Quantity
6415HC1	Lenovo ThinkSystem DB620S FC SAN Switch 24x32Gb SWL SFP	2
AVG2	Lenovo DB620S FC SAN Switch (Entry)	2
6204	2.8m, 10A/100-250V, C13 to IEC 320-C20 Rack Power Cable	2

8.6 Rack BOM

Table 31 lists the BOM for a rack assuming a partly filled rack that has up to 24 power connections. Up to 4 additional PDUs can be added.

Table 31: Rack BOM

Code	Description	Quantity
9363RC4	Rack : 42U 1100mm Deep Dynamic Rack	1
A1RC	42U 1100mm Enterprise V2 Dynamic Rack	1
5895	1U 12 C13 Switched and Monitored 60A 3 Phase PDU	2

Resources

For more information about the topics that are described in this document, see the following resources:

- Microsoft SQL Server 2017:
  microsoft.com/en-us/sql-server/sql-server-2017

- Microsoft SQL Server 2017 data warehousing
  microsoft.com/en-us/sql-server/data-warehousing

- Microsoft Always On Availability Groups
  docs.microsoft.com/en-us/sql/database-engine/availability-groups/windows/overview-of-always-on-availability-groups-sql-server

- Microsoft Windows Server Failover Clustering Instance
  docs.microsoft.com/en-us/sql/sql-server/failover-clusters/windows/windows-server-failover-clustering-wsfc-with-sql-server

- Microsoft Windows Server 2016
  docs.microsoft.com/en-us/windows-server/get-started/server-basics

Document History

Version 1.0   21 February 2019   • Initial version

Version 1.1   13 June 2019   • Updated with improved performance results and better coverage of both local drive and shared storage configurations including use of Intel Xeon Scalable Processor generation 2.

Trademarks and special notices