HPE Reference Architecture for SQL Server 2017 on Red Hat Enterprise Linux with HPE Synergy 660 Compute Module and HPE Nimble Storage AF60

Contents

Executive summary

As organizations continue to grow, demand for faster database transaction processing escalates. Many businesses also are searching for solutions that are scalable, highly available, and provide increased flexibility to meet their needs. To address these issues, customers often refresh their platforms as a way to improve performance and need to ensure their database systems are resilient to achieve their business goals.

This Reference Architecture is intended to provide expected performance implications associated with Microsoft® SQL Server 2017 on Red Hat Enterprise Linux® with the HPE Synergy 660 Gen10 Compute Module server and All-Flash HPE Nimble Storage AF60. We will also highlight SQL Server 2017 Always On Failover Cluster Instance (FCI) using HPE Serviceguard.

The Reference Architecture is ideal for customers:

- Looking to improve transaction rate on general purpose and Small/Mid-level SQL Server OLTP database deployments.
- To refresh their current infrastructure and environment (e.g. SQL Server 2008 or SQL Server 2008 R2) to upgrade to new SQL Server features and seeking better transaction performance.
- Customers using Oracle OLTP databases who wish to stay on Linux and migrate to a lower cost solution.
- Looking to configure highly available SQL Server Failover Cluster Instance on Linux.

HPE Synergy Composable Infrastructure speeds up deployment as well as simplifies scaling up and scaling out to meet SQL Server database workload demands and reduces management complexity using HPE OneView. SQL Server resources can be deployed rapidly through the software-defined intelligence embedded in the HPE Synergy Composer and HPE Synergy Image Streamer. An administrator can utilize HPE Synergy Image Streamer to develop a deployment plan to install and configure both the operating system and application software.

This Reference Architecture (RA) demonstrates scale up performance of HPE Synergy 660 Gen10 Compute from 2-processors to 4-processors running SQL Server 2017 on Linux configured with storage on HPE Nimble Storage AF60. In this solution, high performance HPE Synergy 660 Gen10 Compute Module was chosen to demostratate scale-up performace. The performance testing on this Reference Architecture configuration demonstrated that the solution scales from ~33K batch request per second with 2-processors configuration and up to 65K batch requests per second with 4-processors configuration in the same hardware configuration. Additionally, showcases instance level protection for SQL Server Always On Failover Cluster Instance on Linux using HPE Serviceguard.

Target audience: This white paper is for IT architects, IT managers, database engineers, and administrators. A working knowledge of server architecture, networking architecture, and storage design is recommended.

Document purpose: The purpose of this document is to describe a Reference Architecture, highlighting key implementation details and benefits to technical audiences.

This Reference Architecture describes solution testing performed in February 2019.

Introduction

Today's businesses face a constant challenge keeping pace with the huge data processing and storage requirements generated by all aspects of their business, and as well as keeping the environment resilent for their users. To meet the growing and evolving demands of your business, you need a software-defined infrastructure that provides industry-leading performance, reliability, and faster deployment.

HPE Synergy Composable Infrastructure, the platform built from the ground up, offers a software-defined experience that empowers IT to create and deliver new value instantly and continuously. HPE Synergy supports both 2-socket and 4-socket compute modules which provide the performance, scalability, density optimization, and configuration flexibility to power a variety of workloads. Microsoft's SQL Server 2017 running on Linux deployed on HPE Synergy 660 Gen10 Compute Module with HPE Nimble Storage AF60 can provide your business a great solution to meet your database needs and performance needs. HPE Nimble arrays powered by HPE InfoSight predictive analysis technology, are perfectly suited for database OLTP environments with the ability to group arrays together for management and aggregation. HPE Serviceguard provides high availability on Linux and supports both Always On Failover Cluster Instance (FCI) and Always On Availability Groups (AG). This white paper highlights configuration of FCI with shared storage as HPE Nimble Storage AF60.

The solution is ideal for customers, who are looking to refresh their current infrastruture to gain faster performance, upgrade to new SQL Server versions and effectivetly address future storage requirement.

Note

See "Comparing Microsoft SQL Server version" in Resources and additional links for more information

Solution overview

This solution focuses on online transaction processing (OLTP) use cases. The solution is deployed with SQL Server 2017 on Red Hat Enterprise Linux running on HPE Synergy 660 Gen10 Compute Module and compares database performances between a 2-processors and a scale-up 4-processors configuration. The HPE Nimble Storage AF60 is used to provision the SQL Server databases. Additionally, two HPE Synergy 660 Gen10 Compute Module are configured as cluster nodes of highly available SQL Server FCI using HPE Serviceguard.

The key components of the solution are:

- The HPE Synergy 660 Gen10 Compute Module can be scaled-up from 2 to 4 processors based on business needs
- The HPE Nimble Storage AF60 provides storage infrastructure, with its modular design, provides unlimited scale-up capability by simply adding and connecting arrays to the existing SAN
- HPE Serviceguard which provides high availability and business continuity options for SQL Server on Linux platform

Customers can start with a 2-socket configuration then later add CPU's and memory to scale up as business needs grow without adding another hardware footprint.

Figure 1 shows the overview of the solution.

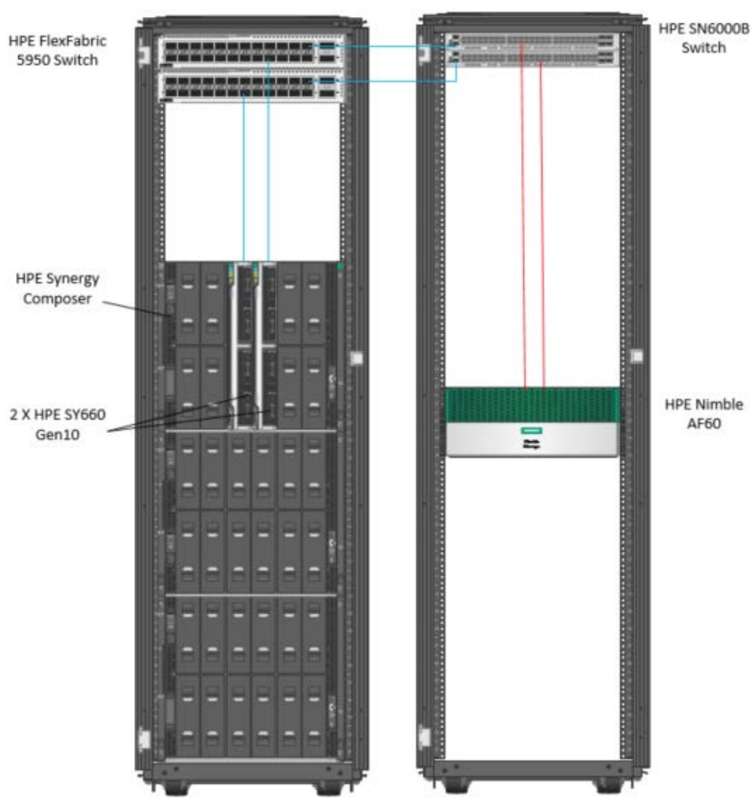


Figure 1. Physical solution diagram

HPE Synergy 660 Gen10 Compute Module

The HPE Synergy 660 Gen10 Compute Module delivers agility, control and security in a 4-socket, full-height form factor to support demanding workloads and virtualization density. Powered by 2 or 4 Intel® Xeon® Scalable family processors, 48 slots for HPE DDR4 SmartMemory supporting up to 6TB, flexible storage controller options with up to 4 SFF drives (8 uFF drives) and/or up to 4 internal M.2 drives, and six (6) I/O mezzanine slots. This compute module is designed to create a pool of flexible compute capacity within a composable infrastructure. This makes the HPE Synergy 660 Gen10 Compute Module the ideal platform for virtualization density, high availability, and scale-up enterprise workloads.

Figure 2 shows the diagram of HPE Synergy 660 Gen10 Compute Module.

Figure 2. HPE Synergy 660 Gen10 Compute

HPE Nimble All-Flash Array

HPE Nimble Storage All-Flash Arrays combine a flash-efficient architecture with HPE InfoSight predictive analytics to achieve fast, reliable access to data and 99.9999% guaranteed availability. Radically simple to deploy and use, the arrays are cloud-ready, providing data mobility to the cloud through HPE Cloud Volumes. HPE Nimble Storage All-Flash Arrays include all-inclusive licensing, easy upgrades, and flexible payment options – while also being future-proofed for new technologies, such as NVMe and Storage Class Memory (SCM). The HPE Nimble controllers are fully redundant and are configured in an Active/Standby configuration. Redundant management ports are also included in the base chassis.

HPE Nimble Storage All-Flash Arrays provide exceptional expandability through the use of additional drive shelves and expansion slots for adding more I/O cards. Each HPE Nimble array can support up to 144 drives and six I/O cards. The additional drive shelves do not include controllers, and simply add capacity to the drive array. The HPE Nimble All-Flash Arrays now scale up from 128TB in the AF20Q to 4PB in the AF80 using four expansion shelves.

- HPE InfoSight Predictive Analytics – Array metrics are loaded into HPE InfoSight, where they are continuously monitored for possible issues. Any issues are then resolved before they can incur downtime.
- 33-66% lower Total cost of ownership (TCO) compared to disk or hybrid solutions – All-flash storage features much higher performance and reliability.
- Flexibility – HPE Nimble Adaptive Flash Arrays provide a lower cost solution for backup/DR/DevTest and archival needs.
- Up to 5x or more data reduction from variable block inline deduplication and compression and provides data reduction, snapshots, and Triple+ Parity RAID with no performance impact.
- App-granular, Federal Information Processing Standards (FIPS) certified encryption provides data-at-rest and over-the-wire protection. Data shredding is built-in.

- Scale-up capacity – HPE Nimble Storage All-Flash Arrays can support over 1000TB of raw storage using optional expansion shelves.
- Ease of management – HPE Nimble arrays can be grouped together in up to groups of 4, so that they can be managed as one unit. As arrays are added they can simply be added to existing or new groups.

Figure 3, shows front view of HPE Nimble All-Flash array.

Figure 3. HPE Nimble All-Flash Array

While basic management is configured through each array's web interface (up to 4 arrays can be grouped together for management), each array can be configured to send health, configuration, and performance data back to HPE InfoSight. Sending data back to HPE InfoSight connects your storage environment to the predictive analytics engine that is the cornerstone of InfoSight's unique proactive management system. Problems can be automatically detected, corrected, and resolved before they impact operations or performance. Figure 4, depicts the dashboard screen in HPE InfoSight, being displayed is an overall status of the storage environment, including capacity, health, and performance.

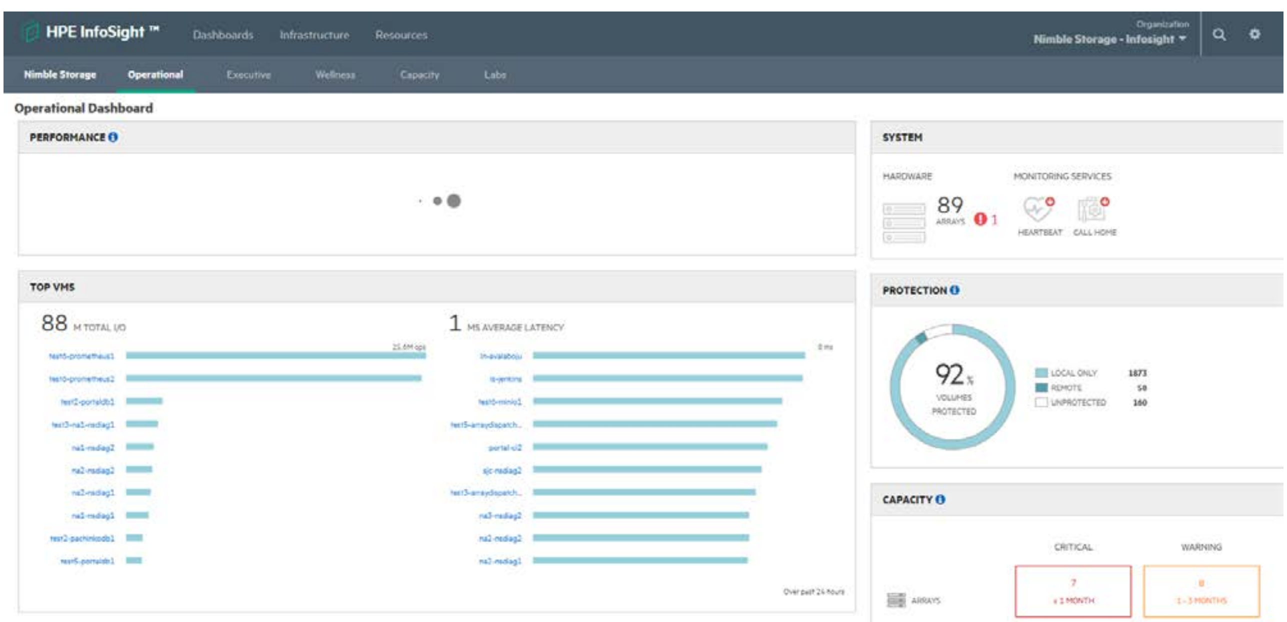


Figure 4. HPE InfoSight web portal

HPE Serviceguard for Linux (SGLX)

HPE Serviceguard for Linux is a high availability and disaster recovery clustering solution that proactively monitors and protects critical applications across virtual and physical environments. HPE Serviceguard for Linux provides fast failover in the event of failure. The HPE Serviceguard for Linux cluster monitors the availability of applications, services, databases, and platforms over distances, while protecting against outages and reducing downtime during planned maintenance windows.

SQL Server 2017 supports the following high availability and disaster recovery deployment models:

- Failover Cluster Instance (FCI) -Based on shared storage architecture, an FCI contains two or more nodes with only one node active at a time and secondary node(s) are available but passive, ready to take an active role during failover. SQL Server FCI provides instance level protection.
- Always On Availability Groups (AG) - In this architecture, primary and standby databases are configured with SQL Server Always On Availability Groups (AG). The databases can be located on the same premises or in the geographically dispersed data centers, the data can be replicated by SQL Server. SQL Server Always On Availability Groups (AG) provides database level protection.

In order to achieve high availability and disaster recovery, both of these architectures would require a cluster manager to perform various critical tasks like monitoring the health of databases, server, network, storage, virtualization layer, virtual machine guests, OS and take automatic actions to minimize application downtimes. HPE Serviceguard for Linux (SGLX) introduces the support of SQL Server to provide mission critical class of robustness to SQL Server deployments and will ensure robust monitoring, reliable actions, and protection against data loss. This market-leading high availability and disaster recovery clustering solution protects your applications from a multitude of infrastructure and application faults across physical or virtual environments over any distance. It reduces the impact of unplanned downtime without compromising data integrity and performance, and helps you achieve near zero planned downtime for maintenance.

When SGLX detects a failure in case of Failover Cluster Instance deployments, it follows shutdown procedures and recovers the database by restarting SQL Server on the adoptive cluster node. In case of Availability Groups (AG) deployments, SGLX will also monitor and administer the replication between primary and standby databases. In case of failures SGLX will perform automatic role management to recover from failures by promotion of the standby database instance to primary. With Availability Groups support, SGLX provides database level protection with faster recovery time.

HPE Serviceguard FCI workload

HPE Serviceguard for Linux provides Instance level protection for SQL Server on Linux deployed in a Failover Cluster Instance model. In the event of failures, HPE Serviceguard for Linux with SQL Server Failover Instance ensures that failures at any level can be automatically mitigated within seconds.

Following are some of the salient features of protection by HPE Serviceguard Solutions for Microsoft SQL Server on Linux Failover Instance:

- Monitoring of diagnostic data and health information of the SQL Server, and various other components as required by the Microsoft SQL Server
- Out of the Box integration
- Instance level protection for SQL Server
- Automatic and Automated failure detection and failover
- Easy Deployment and management from Serviceguard Manager Graphical user interface

---

Note

See "Managing HPE Serviceguard for Linux A. 12.30.00" in Resources and additional for more information

---

Solution Architecture

This solution demonstrates three use cases:

- First use case: SQL Server Scale-up performance from 2-processors to 4-processors on HPE Synergy 660 Gen10 Compute Module
- Second use case: SQL Server Failover Instance using HPE Serviceguard for Linux on HPE Synergy 660 Gen10 Compute Modules
- Third use case: Failover test

First use case

In the first use case, as shown in Figure 5, a preconfigured HPE Synergy 12000 Frame having two HPE Synergy 660 Gen10 Compute Module is connected using HPE Virtual Connect 16Gb FC Module to uplink fiber channel switches. Controller A and Controller B of HPE Nimble Storage AF60 are connected to fiber channel switches. This is done to establish connection between HPE Nimble All-Flash Array and HPE Synergy

Frame 12000. HPE Synergy 660 Gen10 Compute Module is managed using HPE OneView, while HPE Nimble All-Flash Array is managed using HPE Nimble Storage portal.

A single HPE Synergy 660 Gen10 Compute Module was chosen for scale-up performance. Red Hat Enterprise Linux version 7.4 followed by SQL Server 2017 is installed on a local disk. HPE Nimble volumes were created through Nimble GUI Administration to provision files for SQL data, SQL transaction log, and tempdb. Installed Nimble Linux Kit to configure and validate recommended settings. SQL Server Resource Governor is enabled for CPU Affinity to direct the workload to the desired NUMA Nodes. Steps to create Resource Pool, Workload group and Resource Governor Classifier functions are mentioned in Appendix B in section SQL Server Resource Governor Configuration.

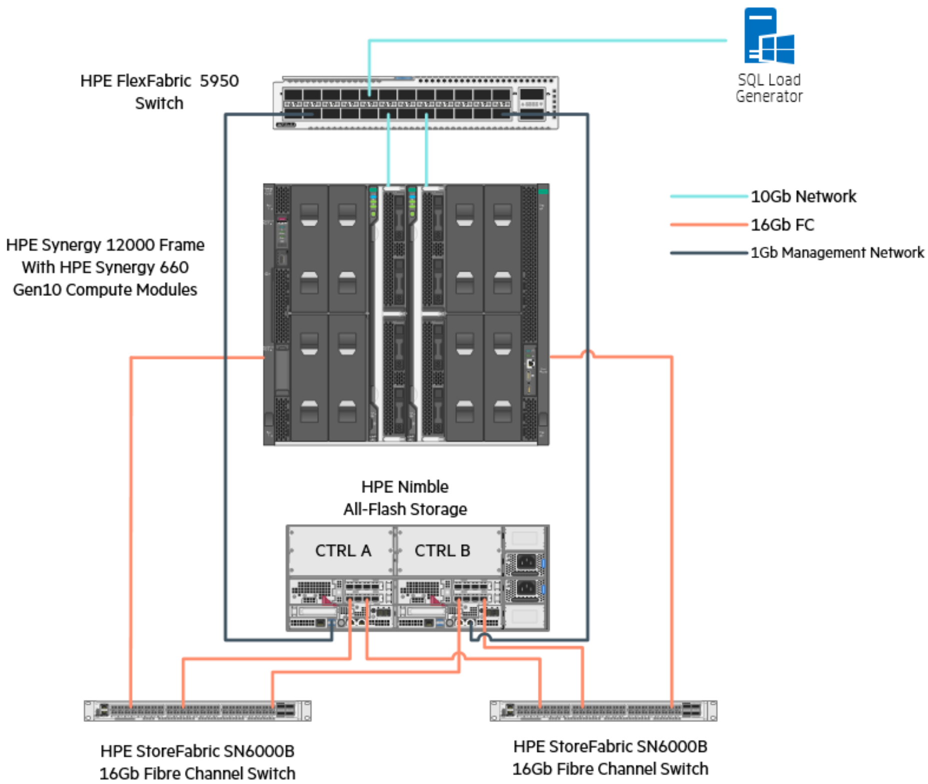


Figure 5. Solution Architecture diagram with network connection

Second use case and Third use case

For second use case and third use case, two HPE Synergy 660 Gen10 were used and installed with Red Hat Enterprise Linux version 7.4. HPE Serviceguard was used to implement SQL Server Failover Cluster Instance. Before installing HPE Serviceguard, the SQL Server configuration had to be configured, as mentioned in Appendix B. The HPE Nimble volumes were created and presented as shared storage to cluster for provisioning SQL Server system databases and user databases. Figure 6 depicts SQL Server FCI configured.

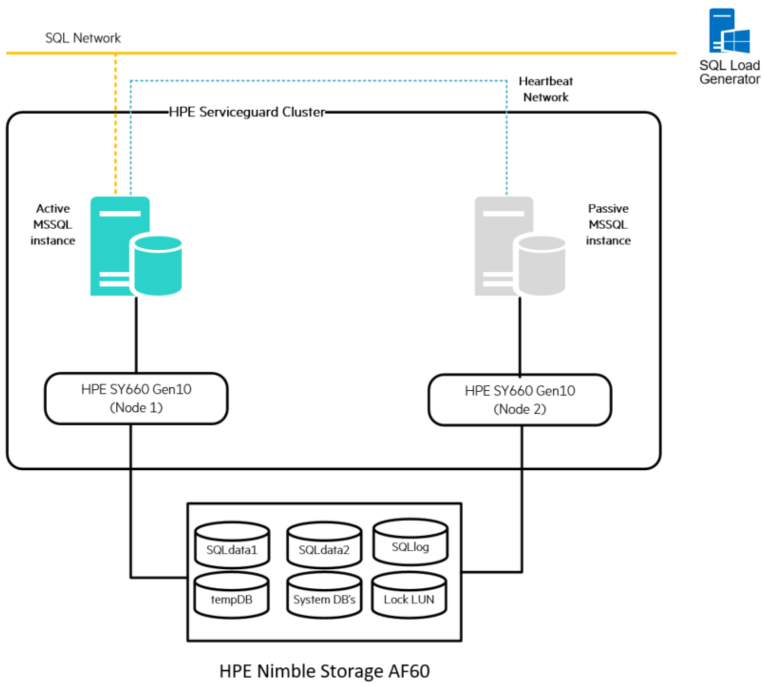


Figure 6. SQL Server Always On Failover Cluster Instance with HPE Serviceguard

HPE Serviceguard provides instance level protection for SQL Server and provides automatic failure detection and failover. When SGLX detects a failure in case of Failover Cluster Instance deployments, it follows shutdown procedures, and recovers the database by restarting SQL Server on the adoptive cluster node.

During a failover scenario, the cluster could encounter a split-brain scenario. To avoid split-brain syndrome, HPE Serviceguard handles this scenario using cluster lock and it's implemented in two ways, either through Lock LUN or Quorum Server. The Lock LUN is a special piece of storage (known as a partition) that is shareable by all nodes in the cluster. When a node obtains the cluster lock, this partition is marked so that other nodes will recognize the lock as "taken". In this solution, Lock LUN was chosen to build the SQL Server FCI, and while Lock LUN is also suppored for Fibre channel protocol only.

The Lock LUN provides arbitration services for HPE Serviceguard clusters, when a cluster partition is discovered, when equal-sized groups of nodes become separated from each other, the Lock LUN allows one group to achieve quorum and form the cluster, while the other group is denied quorum and cluster services on them. In this solution, Lock LUN is created on the HPE Nimble Storage AF60 along with the SQL Server system and user databases. For details about deployment and configuration see Appendix B under HPE Serviceguard deployment and SQL Server FCI workload configuration.

Note

See "Managing HPE Serviceguard for Linux A. 12.30.00" in Resources and additional links for more information about Cluster Lock, Lock LUN and Quorum Server

Solution components

The following components were used for this solution, and the configuration can be varied based on needs.

Hardware

The single HPE Synergy 12000 Frame is configured with:

- 1x HPE Synergy Composer
- 1x Virtual Connect SE 40 Gb F8 Module for Synergy
- 2x HPE Synergy 660 Gen10 Compute Module. This is a four-socket compute module and was populated with 28-core processors for a total of 112 cores. Each HPE Synergy 660 Gen10 Compute module consists of following:
  - 4x Intel Xeon Platinum 8180 CPU @ 2.50GHz processors
  - 24x HPE 64GB Quad Rank 64GB DDR4 Memory
  - 1x HPE Synergy 3820C 10/20Gb CNA
  - 2x HPE Synergy 3830C 16Gb FC HBA

Note

While a single frame HPE Synergy was used for this testing, the recommended production configuration for high availability a pair of Virtual Connect SE 40Gb F8 Modules for Synergy. In production environment, it requires a minimum of three HPE Synergy Frames with two Virtual Connect SE 40Gb F8 Modules for Synergy, two HPE Synergy Composers and two HPE Synergy Image Streamer (based on requirement).

The HPE Nimble Storage AF60 was configured as follows:

- 1x HPE Nimble Storage AF60 all-flash array
  - 48x 480GB SSDs, with total capacity of up to 22TB
  - 4x 16Gb Fibre Channel ports

Software

- Red Hat Enterprise Linux version 7.4
- HPE OneView
- HPE Nimble Storage Toolkit for Linux version 2.3 application software
- Microsoft SQL Server 2017, CU12
- HPE Serviceguard 12.30.00

Additionally it is recommended to install the HPE Nimble Storage Toolkit for Linux version 2.3. The toolkit works with multipathing software and provide services to identify, verify settings and manage HPE Nimble volumes. It also provides an interface to create/update and remove HPE Nimble volumes from the Linux host.

Best practices and configuration guidance for the solution

HPE Synergy 660 Gen10 Compute Module

The basic input/output system (BIOS) settings were configured as below. BIOS settings by leveraging on HPE Synergy Server profiles created through HPE OneView.

- Workload Profile – Transactional Application Processing
- Hyper-Threading – Disabled
- Virtualization Technology – Disabled

- Intel VT-d – Disabled
- SR-IOV – Disabled

Note
See HPE OneView in Resources and additional links for more information Server profiles.

SQL Server configuration

- Trace Flag T834- Large Pages enabled
- Max degree of parallelism set to 1
- SQL PAL was configured to use 93% of available memory through BASH environment variable $PAL_MEMORY_SIZE
- Soft-NUMA disabled
- Multiple tempdb data files were created after installation
- Resource Governor was enabled for CPU affinity

Red Hat 7.4 Enterprise

Performance best practices and configuration guidelines were applied as recommended by Microsoft

- CPU frequency governor set to performance
- Energy_PERF_BIAS set to performance
- min_perf_pct set to 100
- C-States set to C1 only
- Transparent Huge Page (THP) enabled

Disk settings

- Disk readahead set to 4096

Sysctl settings

- kernel.sched_min_granularity_ns = 10000000
- kernel.sched_wakeup_granularity_ns = 15000000
- vm.dirty_ratio = 40
- vm.dirty_background_ratio = 10
- vm.swappiness=10
- kernel.numa_balancing=0
- vm.max_map_count=262144

Note
See Resources and additional links for more information about performance best practices and configuration guidelines for SQL Server on Linux.

HPE Nimble All-Flash Array

HPE Nimble All-Flash Array was configured as per best practices guide when using Fibre Channel HPE Nimble Storage with Linux and HPE Nimble deployment considerations for SQL Server.

Nimble Linux kit was used to validate and configure the recommended settings for volumes, which were created for provisioning SQL Server database files.

Note

See Resources and additional links for more information about HPE Nimble Storage Deployment Considerations for Microsoft SQL Server, and HPE Nimble Storage Deployment Considerations for Linux on Fibre Channel.

Capacity and sizing

Workload description and test methodology

The workload consists of OLTP databases which are representative of a stock trading application emulator, in which the clients connect to the OLTP databases and perform transactional operations of buy, sell, and market orders. The workload is comprised of approximately 70 percent disk reads and 30 percent disk writes.

Scale-up tests were focused on two use cases:

- Performance of HPE Synergy 660 Gen10 Compute Module with 2-processor
- Performance of HPE Synergy 660 Gen10 Compute Module with 4-processor

For both scenarios, databases were provisioned on HPE Nimble Storage AF60.

Note

See "GUI Administration Guide" in Resources and additional links for more information about creating and managing volumes.

SQL Server Resource Governor was enabled and configured such that each database had its CPU/Memory affinity isolated at CPU socket level. In other words, each database had its affinity set to use resources from a dedicated NUMA node on the server. For the most part the default resource pool settings were used to create the resource pools, except the AFFINITY NUMANODE value was uniquely set for each database. For the OLTP workload groups, in order to have consistent and reproducible results across multiple workload runs, we set MAXDOP value to 1. Unlike DSS/Warehouse queries, OLTP transactions do not yield a performance benefit with increased parallelism.

Performance of HPE Synergy 660 Gen10 Compute Module with 2-processor

In this test case, eight database size of approximately 200GB each was used. Using SQL Server Resource Governor, Resource Pool were affinitized to NUMA 0 and NUMA 2 to receive workload from two SQL Logins DB DB01User and DB02User. A total of database size of 1.6TB was used for this testing. Table 1, shows the Resource Governor configuration for 2-processors.

Table 1. Resource Governor configuration for 2-processors

Username	Workload Group	Resource Pool	Affinity NUMA Node	Databases	Database size
DB01User	WGDB01	RSPOOLDB01	0	TPCE01 to TPCE_04	4x 200GB
DB02User	WGDB02	RSPOOLDB02	2	TPCE05 to TPCE_08	4x 200GB

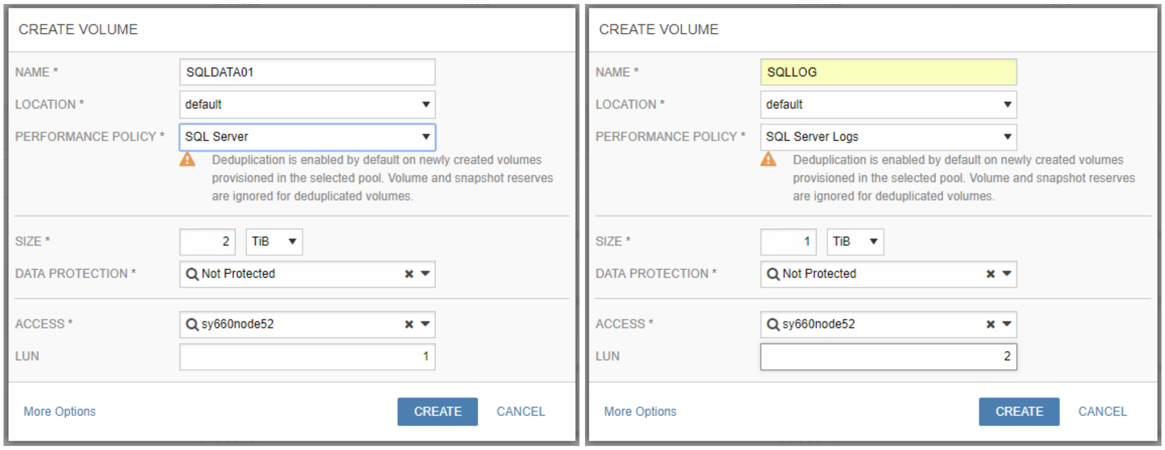


Figure 7. Creating volume using HPE Nimble Storage Portal

HPE NimbleOS GUI was used to provision three volumes SQLDATA1, SQLDATA2 and SQLLOG to create SQL databases. Two volumes for data were created to split the data files into separate volumes to enhance the write operations. During volume creation, for data files and tempdb Performance Policy was set to predefined "SQL Server" and for log files it was set as "SQL Server Logs". Table 2, shows mapping of volume, mount points and database files.

Table 2. HPE Nimble All-Flash Storage Volume provisioning

Volume Name	Size	Device mapping	OS Mount path	Performance Profile
SQLDATA1	1TB	/dev/mapper/mpathk	/sql/sqldata1	SQL Server
SQLDATA2	1TB	/dev/mapper/mpathl	/sql/sqldata2	SQL Server
tempdb	512GB	/dev/mapper/mpathg	/sql/tempdb	SQL Server
SQLLOG	750GB	/dev/mapper/mpathm	/sql/sqllog	SQL Server Logs

The Resource Governor was configured to route workload sessions with unique logins to use the corresponding workload group, which in turn, used the corresponding resource pool. Resource Governor Classifier function was configured to map username workload group to access database workloads. A script was created for ease of configuration. The same script was used for the 2-Socket and 4-Socket testing. For sample scripts to configure the Resource Governor, see Appendix B, for more details.

With 2-processor HPE Synergy 660 Gen10 Compute Module, 768GB of physical memory was used. The SQL Server maximum memory was set to 90% of physical memory. For each testing, the database was restored to ensure results are normalized. The workload was generated from workload server and the total CPU utilization reached about 90%. The test was run for 30 minutes and batch requests were captured through Activity Monitor from SQL Server. CPU usage, IO and network data care was captured using vmstat of Linux.

Performance of HPE Synergy 660 Gen10 Compute Module with 4-socket

In this test case, 16 databases of approximately 200GB each were used. Using SQL Server Resource Governor, Resource Pool was affinitized to NUMA 0, NUMA 1, NUMA 2, and NUMA 3 to receive workload from four SQL Logins: DB01User, DB02User, DB03User, and DB04User. A total of database size of 3.2TB was used for this testing. Table 3, shows the Resource Governor configuration for 4-socket testing.

Table 3. Resource Governor configuration for 4-socket

Username	Workload Group	Resource Pool	Affinity NUMA Node	Databases	Database size
DB01User	WGDB01	RSPOOLDB01	0	TPCE_01 to TPCE_04	4x 200GB
DB02User	WGDB02	RSPOOLDB02	1	TPCE_05 to TPCE_08	4x 200GB
DB03User	WGDB03	RSPOOLDB03	2	TPCE_09 to TPCE_12	4x 200GB

DB04User	WGDB04	RSPOOLDB04	3	TPCE_13 to TPCE_16	4x 200GB

To provision databases, from HPE Nimble web portal three volumes SQLDATA1, SQLDATA2 and SQLLOG volumes were created. Two volumes for data were created to split the data files into separate volumes to enhance the write operations. During volume creation, for data files and tempdb Performance Policy was set to predefined “SQL Server” and for log files it was set as “SQL Server Logs”. Table 4, shows mapping of volume, mount points and database files.

Table 4. HPE Nimble All-Flash Storage Volume provisioning

Volume Name	Size	Device mapping	OS Mount path	Performance Profile
SQLDATA1	2TB	/dev/mapper/mpathk	/sql/sqldata1	SQL Server
SQLDATA2	2TB	/dev/mapper/mpathl	/sql/sqldata2	SQL Server
Tempdb	512GB	/dev/mapper/mpathg	/sql/tempdb	SQL Server
SQLLOG	1.5TB	/dev/mapper/mpathm	/sql/sqllog	SQL Server Logs

With 4-processor HPE Synergy 660 Gen10 Compute Module, 1.5TB of physical memory was used. The SQL Server maximum memory was set to 90% of physical memory. For each testing, the database was restored to ensure results are normalized. The workload was generated from the workload server and after total CPU utilization reached about 90%, the test was run for 30 minutes. Batch requests were captured through Activity Monitor from SQL Server. CPU usage, IO and network data care was captured using vmstat of Linux.

Workload test results

In 2-processor testing, eight databases were used and on each databases 25 user’s workload were driven. A total of 200 user’s workload were driven and it was observed that the CPU reached about 84% from 2-processor (considering only 2 NUMA’s), and batch requests were recorded at 33k per second. The average IOPS recorded was 27K. Figure 8, shows the CPU usage of two NUMA 0 and NUMA 1 captured using the htop utility in Linux. Figure 9, shows recorded batch requests per second.

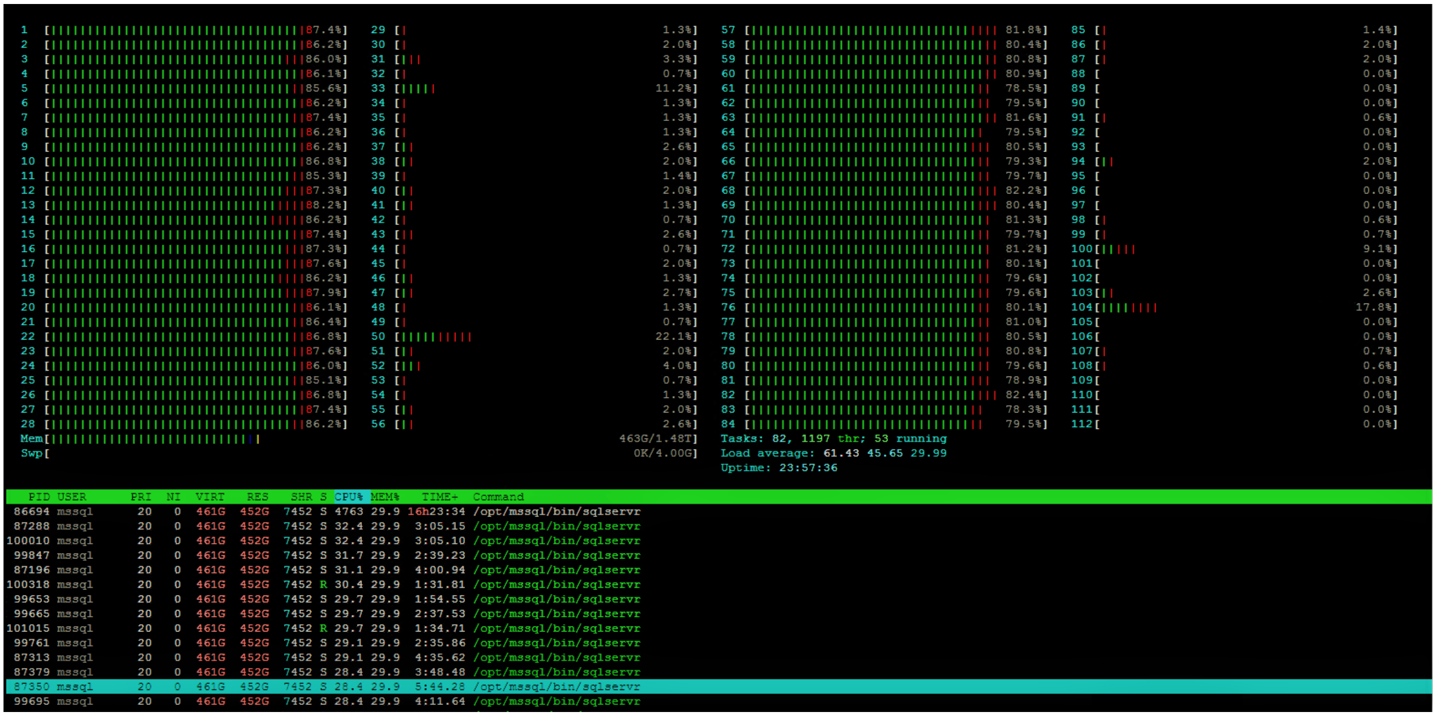


Figure 8. CPU usage for 2-processors

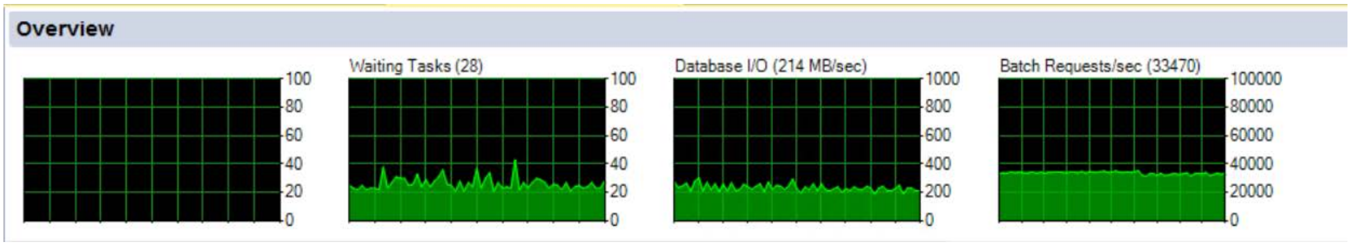


Figure 9. Batch Requests per second for 2-processors, captured using SQL Activity Monitor

In 4-processor testing, sixteen databases were used and on each databases 20 user's workload were driven. A total of 320 user's workload were driven and it was observed that the CPU reached about 85% and batch requests were recorded at 65k per second. The average IOPS recorded was 56K. Figure 10, shows the CPU usage of all the NUMAs captured using the htop utility in Linux. Figure 11, shows recorded batch requests per second.

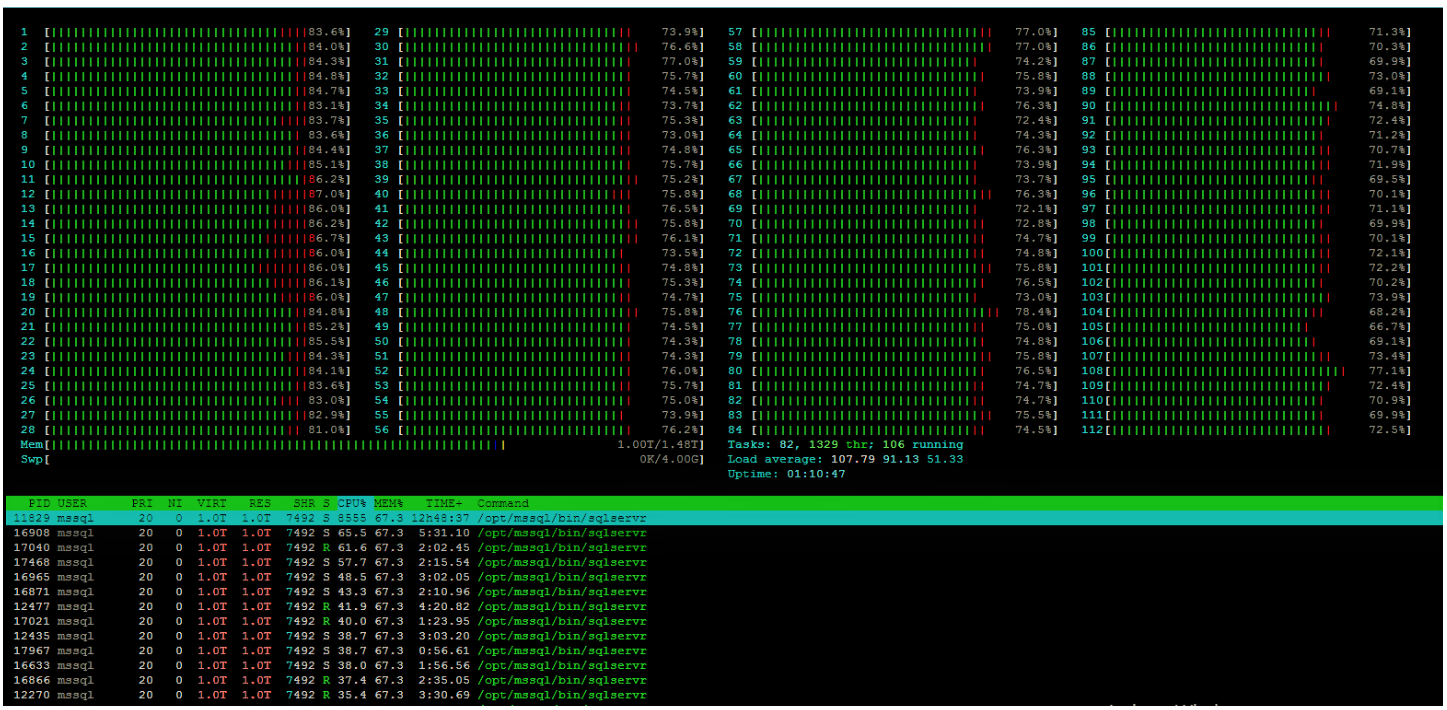


Figure 10. CPU usage for 4-processors

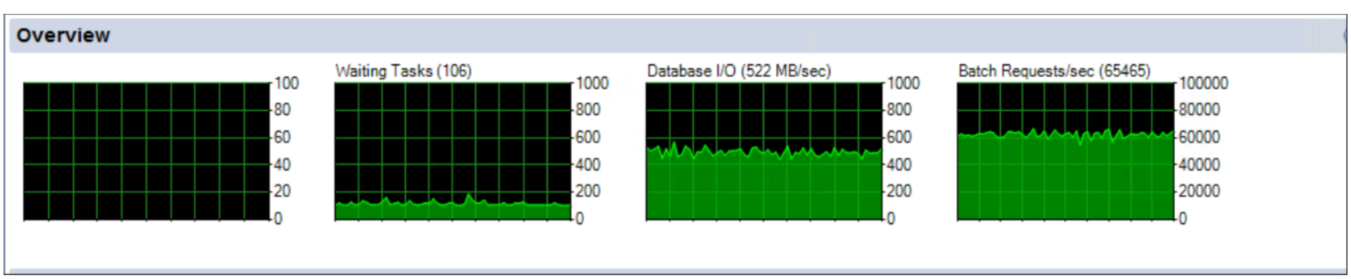


Figure 11. Batch requests per second for 4-processor, captured using SQL Activity Monitor

The results shows when the processor is scaled from 2-processors to 4-processors, the batch requests per second increased by 96%, doubling the database capacity and increasing the workload by 60%. Figure 12, depicts the scale-up results when running SQL Server 2017 with Linux on HPE Synergy 660 Gen10 Compute Module with HPE Nimble Storage AF60. The solution is able to improve server performance and capacity by scaling up computing power and storage at a lower cost than the additional hardware footprint.

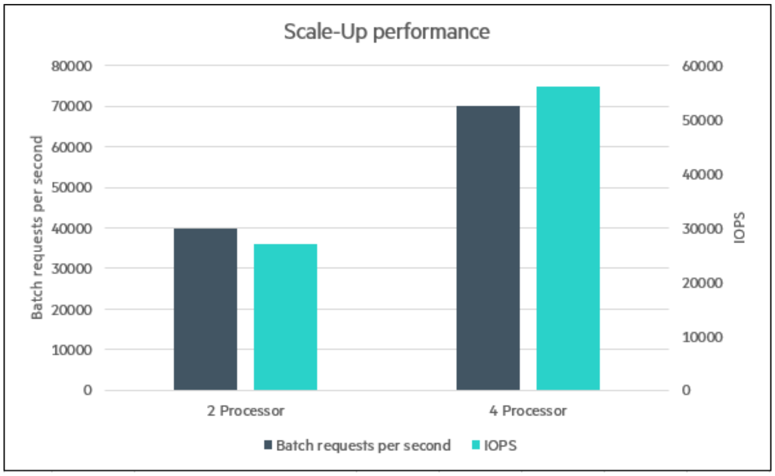


Figure 12. Scale-up performance of HPE Synergy 660 Gen10 with HPE Nimble Storage AF60

Failover test results

To demonstrate the solution's high availability capability using HPE Serviceguard, failover scenario was tested when high transaction workload was driven across 16 user databases with approximately 200GB each, in total 3.2TB size of databases. HPE Serviceguard MEMBER_TIMEOUT was set to 3 seconds. There were 320 user connections to these databases, each of these clients have 200 transaction per second, totaling to 64000 transactions per second.

The primary node was abruptly bought down to simulate failure of hardware. The HPE Serviceguard was able to detect the failure, reform the cluster and start the SQL Server application on the adoptive node in 3 Seconds. The database recovery took about 30 seconds. SQL Server instance was accessible to clients in a span of 34 seconds from the start of failure. Table 5, shows the details of failover scenario handled by HPE Serviceguard. Figure 13 shows represents the timelines of failure and recovery of the database.

Table 5. HPE Serviceguard database recovery timelines.

Sequence Number	Event	Timestamp	Time Elapsed between each event
1	Primary Node was stopped abruptly	5:04:17	
2	Identified by Serviceguard	5:04:17	0:00:00
3	Cluster Reformation by Serviceguard	5:04:20	0:00:03
4	Package Start on passive Node	5:04:20	0:00:00
5	SQL Server DB recovery complete	5:04:50	0:00:30
6	Package Start complete	5:04:51	0:00:01
		Total time take from failure to recovery (1-6)	0:00:34

Figure 13, represents the sequence of failover handled by HPE Serviceguard.

1. Failure of Primary node
2. Complete Reformation of cluster, start of SQL Server on Adoptive Node, Start of DB recovery
3. Database recovery complete
4. Ready for client connection

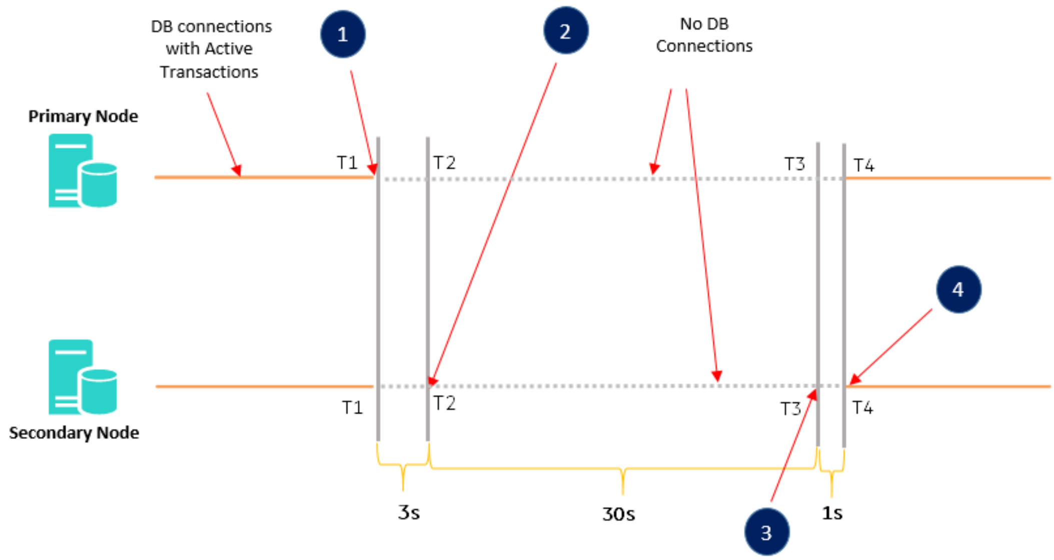


Figure 13. SQL Server Failover scenario with HPE Serviceguard

During testing, using HPE Nimble Storage AF60 16 user databases were provisioned with approximately 200GB each, for a total of 3.2TB. However, as shown in Figure 14, HPE Nimble Storage reduces the storage foot print by utilizing its in-line de-duplication and compression technology.

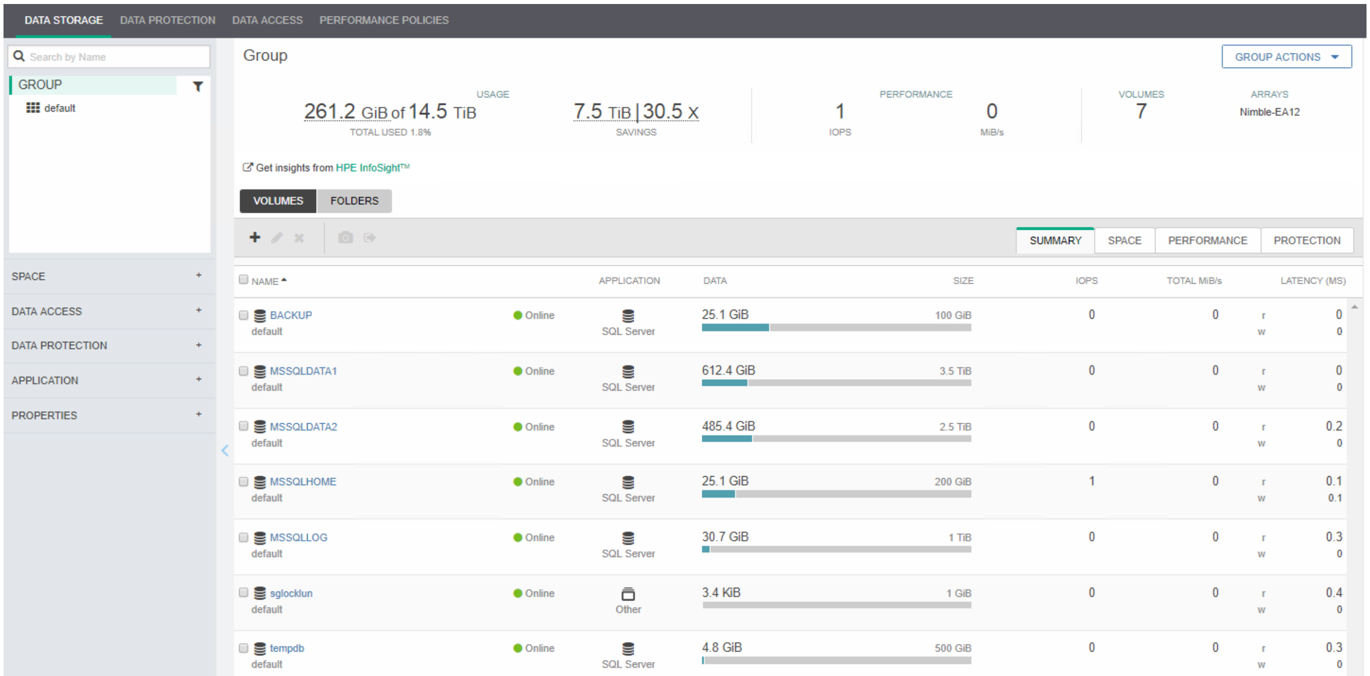


Figure 14. HPE Nimble Storage AF60 capacity

This Reference Architecture utilizies a pre-configured HPE Synergy Composble Infrastructure, with software-defined experience that empowers IT to create and deliver new value instantly and continuously. The result demonstrates, the HPE Synergy 660 Gen10 Compute Module combined with HPE Nimble Storage able to improve server performance, and capacity by scaling up computing power and storage. The solution can be cost effective by gaining scale-up performance in the same hardware by increasing the compute and memory capacity.

Analysis and Recommendations

This Reference Architecture demonstrates SQL Server 2017 scale-up performance with Red Hat Enterprise Linux running on HPE Synergy 660 Gen10 Compute Module using HPE Nimble Storage AF60. By leveraging on HPE Synergy's software-defined experience, HPE Synergy simplifies by scaling-up to meet SQL Server database workload demands and reduces management complexity using HPE OneView. The solution is ideal for customers who are looking to increase transaction performance, and anticipate the need to accommodate future workloads. The solution is cost effective helps in reducing hardware footprint by gaining the performance on the same hardware.

The solution demonstrated that administrators can start off with 2-processor SQL Server solution, and as need arises, the solution can easily scale-up to 4-processors. HPE Synergy 660 Gen10 Compute Module with 2-processors, SQL Server achieved 33k batch requests per second and by scaling up to 4-processors on the same hardware, SQL Server achieved 65k batch requests per second. The solution can handle almost double the workload with support from HPE Nimble All-Flash Storage and deliver sustained performance. Further, HPE Nimble All-Flash array helps in reducing storage foot print by leveraging on inline-deduplication and data compression technology.

For customers looking to build highly available SQL Server on Linux platform, customers can significantly minimize unplanned downtime of their business workloads and protect data integrity through rich capabilities of HPE Serviceguard and its deep integration with SQL Server application. Serviceguard quickly detects the problems with infrastructure and applications, and take required actions such that the workload automatically gets failed over to HA or DR node and business continuity is not compromised. For planned maintenance, Serviceguard eliminates the downtime by providing strong ability to seamlessly move the workload to standby node.

Appendix A: Bill of materials (BOMs)

The following BOMs contain electronic license to use (E-LTU) parts. Electronic software license delivery is now available in most countries. HPE recommends purchasing electronic products over physical products (when available) for faster delivery and for the convenience of not tracking and managing confidential paper licenses. For more information, please contact your reseller or a Hewlett Packard Enterprise representative.

Note

Part numbers are at time of publication/testing and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your Hewlett Packard Enterprise Reseller or Hewlett Packard Enterprise Sales Representative for more details. hpe.com/us/en/services/consulting.html

Table 1a. HPE Synergy 12000 Frame

Qty	Part number	Description
1	797739-B21	HPE Synergy 12000 2FLM6PS10F Frame
1	804943-B21	HPE Synergy 12000 Frame 4x Lift Handle
1	804353-B21	HPE Synergy Composer
1	804942-B21	HPE Synergy Frame Link Module
1	804938-B21	HPE Synergy 12000 Frame Rack Rail Kit
2	779227-B21	HPE Virtual Connect SE 16Gb FC Module for HPE Synergy
2	779218-B21	HPE Synergy 20Gb Interconnect Link Module

Table 1b. HPE Synergy 660 Gen10 Compute Module

Qty	Part number	Description
2	872119-L21	HPE Synergy 660 Gen10 Intel Xeon-Platinum 8180 (2.5GHz/28-core/205W) FIO Processor Kit
2	872119-B21	HPE Synergy 660 Gen10 Intel Xeon-Platinum 8180 (2.5GHz/28-core/205W) Processor Kit
2	777430-B21	HPE Synergy 3820C 10/20Gb Converged Network Adapter
2	777452-B21	HPE Synergy 3830C 16Gb Fibre Channel Host Bus Adapter
48	815101-B21	HPE 64GB (1x64GB) Quad Rank x4 DDR4-2666 CAS-19-19-19 Load Reduced Smart Memory Kit

Table 1c. HPE Nimble Storage AF60

Qty	Part number	Description
1	Q8H42A	HPE NS AF60 All-Flash CTO Base Array
1	Q8H46A	HPE NS AF60/80 11.52TB FIO Flash Bndl
1	Q8B83B	HPE Nimble Storage Embedded 10GBASE-T 2-port FIO Adapter
1	HT6Z0A3	HPE NS 3 FC 4H Parts Exchange Support
1	Q8H42A	HPE NS AF60 All-Flash CTO Base Array

Table 1d. Switches

Qty	Part number	Description
2	QR480B	HPE SN6000B 16Gb 48/48 FC Switch
1	Q8H46A	HPE NS AF60/80 11.52TB FIO Flash Bndl

Table 1e. HPE Serviceguard

Qty	Part number	Description
8	BB095ACE	HPE Serviceguard for Linux x86 Advanced (Per core based license)

Appendix B: Configuration adjustments

This section covers configuration aspects implemented for this Reference Architecture solution.

SQL Server Resource Governor Configuration

The HPE Synergy 660 Gen10 Compute Module is a 4-socket Compute Module. In this solution, Resource Governor was enabled to direct the workload based on SQL Server username from which the workload was initiated. The steps involved in configuring Resource Governor are:

1. Create Resource pool and configure CPU affinity by specifying respective NUMA
2. Create Workload Group
3. Create username to intiate the workload from this username.
4. Create Resource Governor Classifier function to direct the workload based on username to respective Workload Group.

Following TSQL scripts were used to create and configure the Resource Governor:

```
IF EXISTS (SELECT name FROM sys.resource_governor_resource_pools WHERE name= 'RSPOOLDB01')
        ALTER RESOURCE POOL RSPOOLDB01 WITH(min_cpu_percent=0,
                                        max_cpu_percent=100,
                                        min_memory_percent=0,
                                        max_memory_percent=100,
                                        cap_cpu_percent=100,
        AFFINITY NUMANODE = (0),
                                        min_iops_per_volume=0,
                                        max_iops_per_volume=0)
ELSE

                        CREATE RESOURCE POOL RSPOOLDB01 WITH(min_cpu_percent=0,
                                        max_cpu_percent=100,
                                        min_memory_percent=0,
                                        max_memory_percent=100,
                                        cap_cpu_percent=100,
        AFFINITY NUMANODE = (0),
                                        min_iops_per_volume=0,
                                        max_iops_per_volume=0)
GO
|
ALTER RESOURCE GOVERNOR RECONFIGURE;

IF EXISTS (SELECT name FROM sys.dm_resource_governor_workload_groups WHERE name= 'WGDB01')
                ALTER WORKLOAD GROUP WGDB01 WITH(group_max_requests=0,
                                importance=High,
                                request_max_cpu_time_sec=0,
                                request_max_memory_grant_percent=25,
                                request_memory_grant_timeout_sec=0,
         MAX_DOP = 1) USING RSPOOLDB01
ELSE
                CREATE WORKLOAD GROUP WGDB01 WITH(group_max_requests=0,
                                importance=High,
                                request_max_cpu_time_sec=0,
                                request_max_memory_grant_percent=25,
                                request_memory_grant_timeout_sec=0,
         MAX_DOP = 1) USING RSPOOLDB01
GO
ALTER RESOURCE GOVERNOR RECONFIGURE;
```

```
USE MASTER
GO
IF NOT EXISTS(SELECT loginname FROM master.dbo.syslogins WHERE name ='DB01user')
        CREATE LOGIN DB01user
        WITH PASSWORD ='**********',
        DEFAULT_DATABASE=master,
        DEFAULT_LANGUAGE=us_english,
        CHECK_EXPIRATION=OFF,
        CHECK_POLICY=OFF
ELSE
        ALTER LOGIN DB01user WITH PASSWORD = '**********'
GO
ALTER LOGIN DB01user ENABLE
GO
ALTER SERVER ROLE sysadmin ADD MEMBER DB01user
GO
```

```
CREATE FUNCTION [dbo].[OLTPClassifier]() RETURNS sysname
        WITH SCHEMABINDING
AS
BEGIN
        DECLARE @WorkloadGroup AS SYSNAME
        IF(SUSER_NAME()= 'OLTP\db01user')
                SET @WorkloadGroup='WGDB01'
        ELSE IF (SUSER_NAME()= 'OLTP\db02user')
                SET @WorkloadGroup='WGDB02'
        ELSE IF (SUSER_NAME()= 'OLTP\db03user')
                SET @WorkloadGroup='WGDB03'
        ELSE IF (SUSER_NAME()= 'OLTP\db04user')
                SET @WorkloadGroup='WGDB04'
RETURN @WorkloadGroup
END
GO

ALTER RESOURCE GOVERNOR WITH (CLASSIFIER_FUNCTION = dbo.OLTPClassifier);
GO
ALTER RESOURCE GOVERNOR RECONFIGURE
GO
```

Using these scripts, Resource Governor can be configured for other Resource Pools and Workload Groups.

HPE Serviceguard deployment and SQL Server FCI workload configuration

For this solution, we used two HPE Synergy 600 Gen10 to build SQL Server Always On Failover cluster using HPE Nimble Storage AF60 as shared storage.

Below steps to be completed before starting installation of HPE Serviceguard.

1. Create network teaming on both the nodes participating in forming the cluster. HPE Serviceguard recommends creating two team consisting of two interfaces in each team. Figure 15 shows, image volumes, were created in HPE Nimble Storage AF60

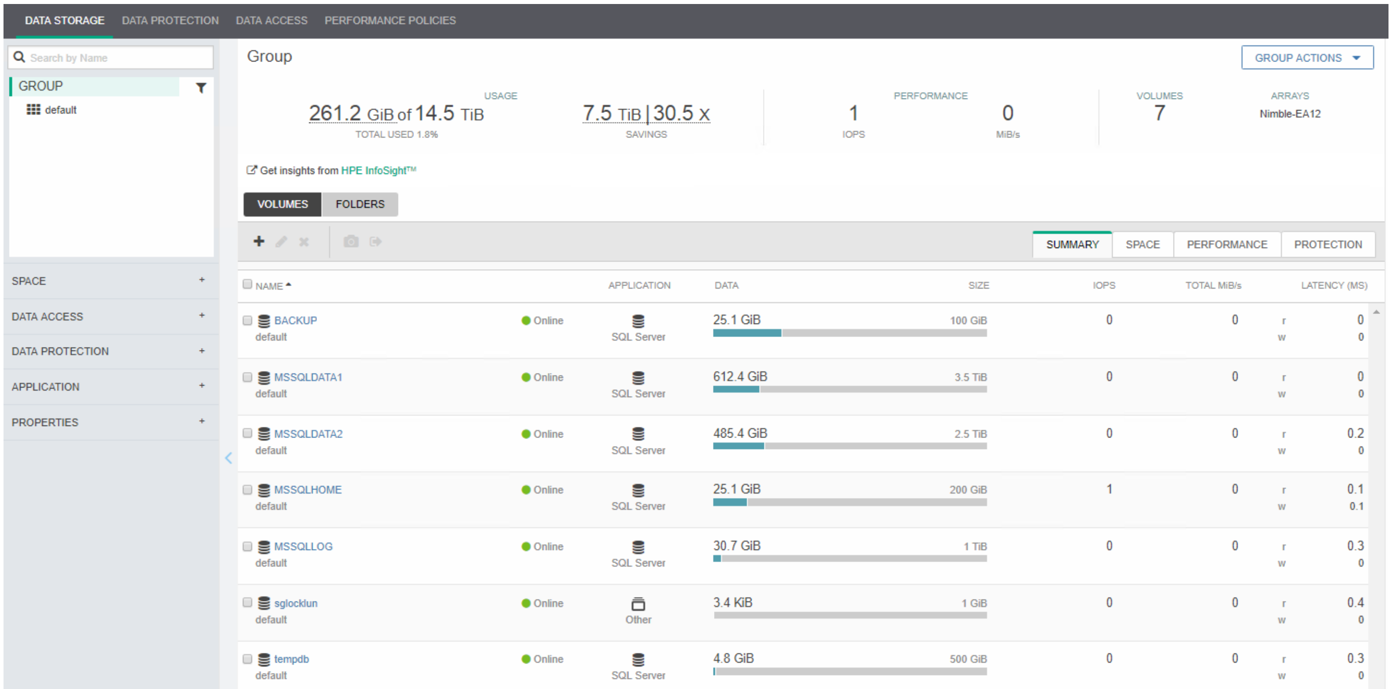


Figure 15. HPE Nimble Storage AF60 Volumes

2. Volumes groups are created and mounted in Linux host, Figure 16, shows mount points.

```
/dev/mapper/vg--sqllog-lvol0   1009G   18G  950G   2% /var/opt/sqllog
/dev/mapper/vg--sqldata-lvol0  3.5T  3.4T     0 100% /var/opt/sqldata
/dev/mapper/vg--sqlhome-lvol0  2.9G   80M  2.6G   3% /var/opt/mssql
/dev/mapper/vg--tempdb-lvol0   4.7G   92M  4.4G   3% /var/opt/tempdb
/dev/mapper/vg--sysdbs-lvol0   1.9G   52M  1.8G   3% /var/opt/sysdbs
```

Figure 16. Mount points in RHEL

3. Run mssql-conf setup to complete SQL Server configuration. On the second node, folder structure has to be created.

Installing HPE Serviceguard and cluster configuration

1. Obtain and install packages for Serviceguard. Mount the Serviceguard for Linux DVD or ISO image in /mnt.
2. Create a Yum Repository file in /etc/yum.repos.d/sglx.repo with the following contents:

   [sglx]

   name=Serviceguard

   baseurl=file:///mnt/RedHat/RedHat7

   gpgcheck=0

   enabled=1

3. Ensure that the user sgmgr exists in the system. Export the SGMGR_ENV environment and run the RPM command.

   export SGMGR_ENV=<password>

4. Install Serviceguard Solutions for Microsoft SQL Server on Linux by running the following command:

   yum install serviceguard-manager.noarch serviceguard-extension-for-mssql.noarch

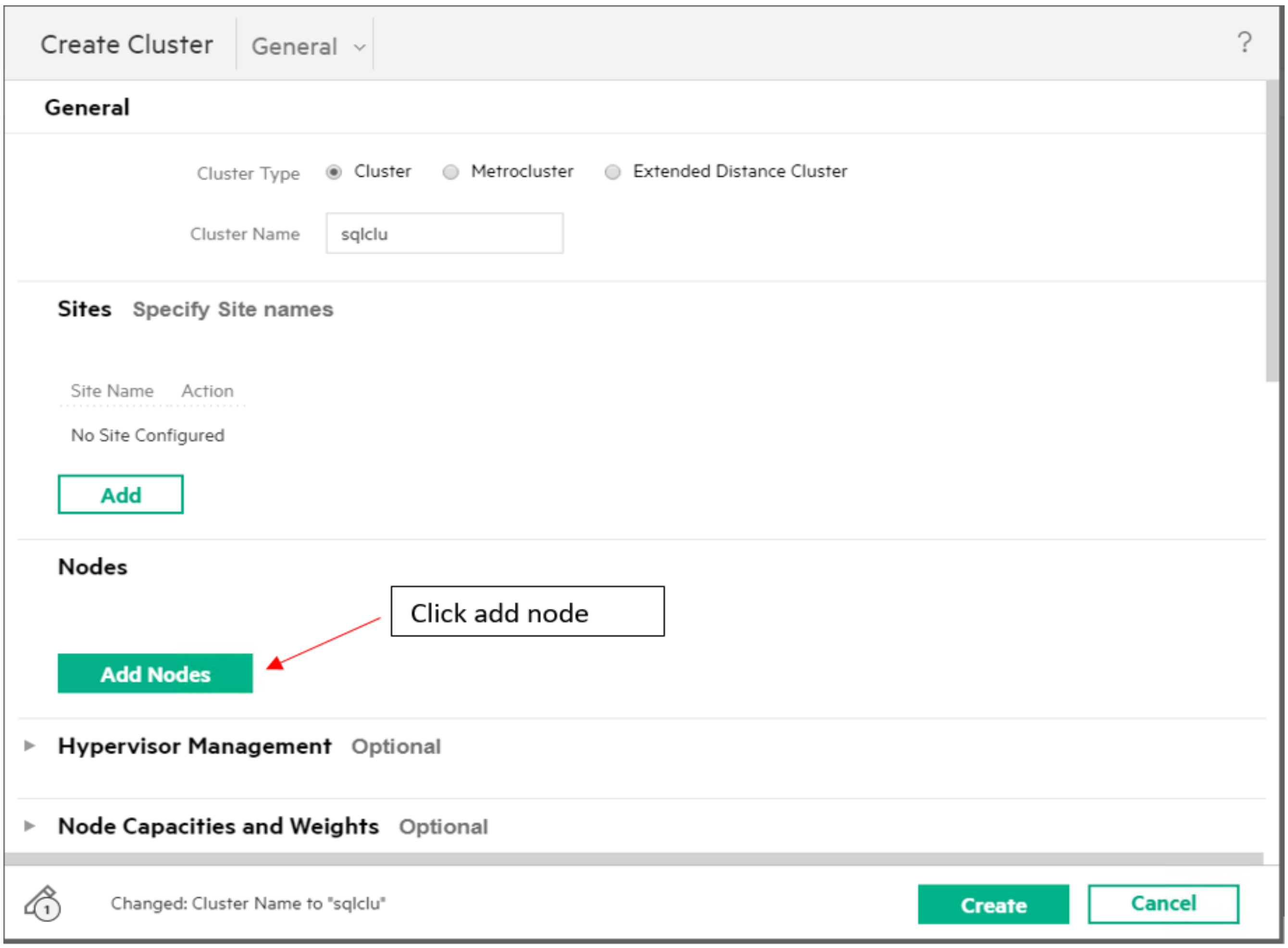
5. Prepare nodes and quorum server for creating cluster. Quorum Server installed with Red Hat Enterprise Linux (RHEL) in the same network.

   cmpreparecl –n node1 –n node2

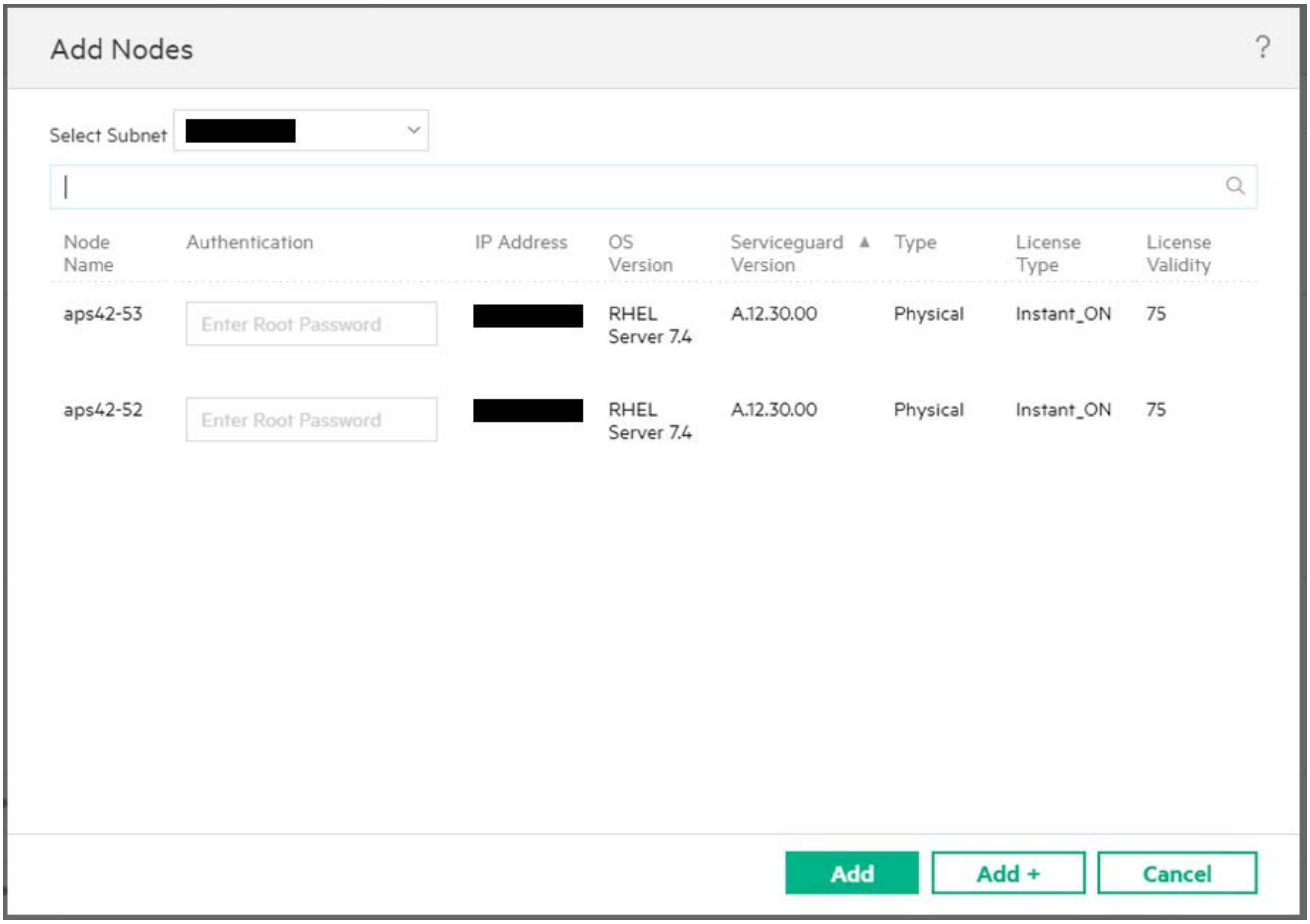
6. Login to HPE Serviceguard Manager portal.

   https://<servername>:5522, and login with root credentials

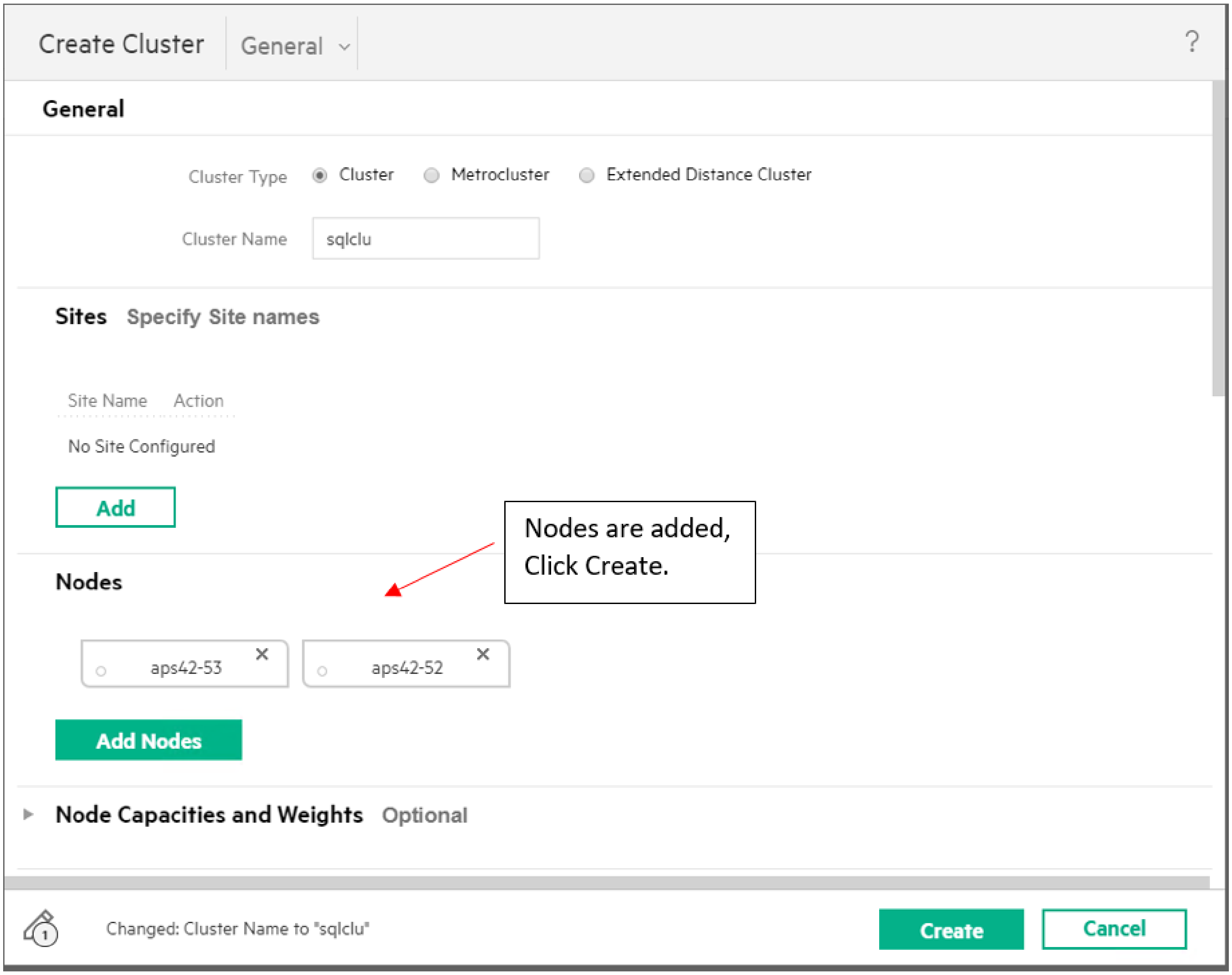
7. Create the Cluster using HPE Serviceguard Manager.



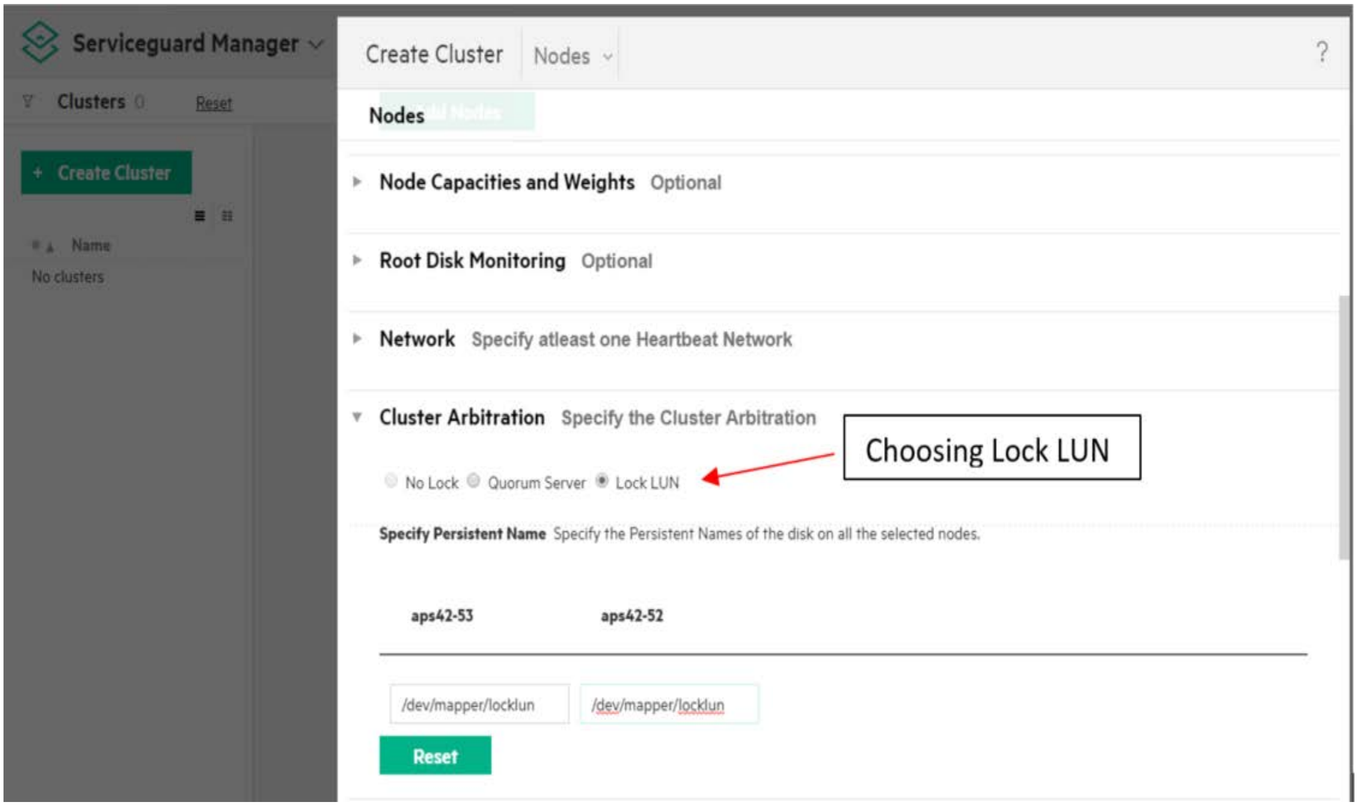
8. Add nodes to Cluster Creation Wizard.



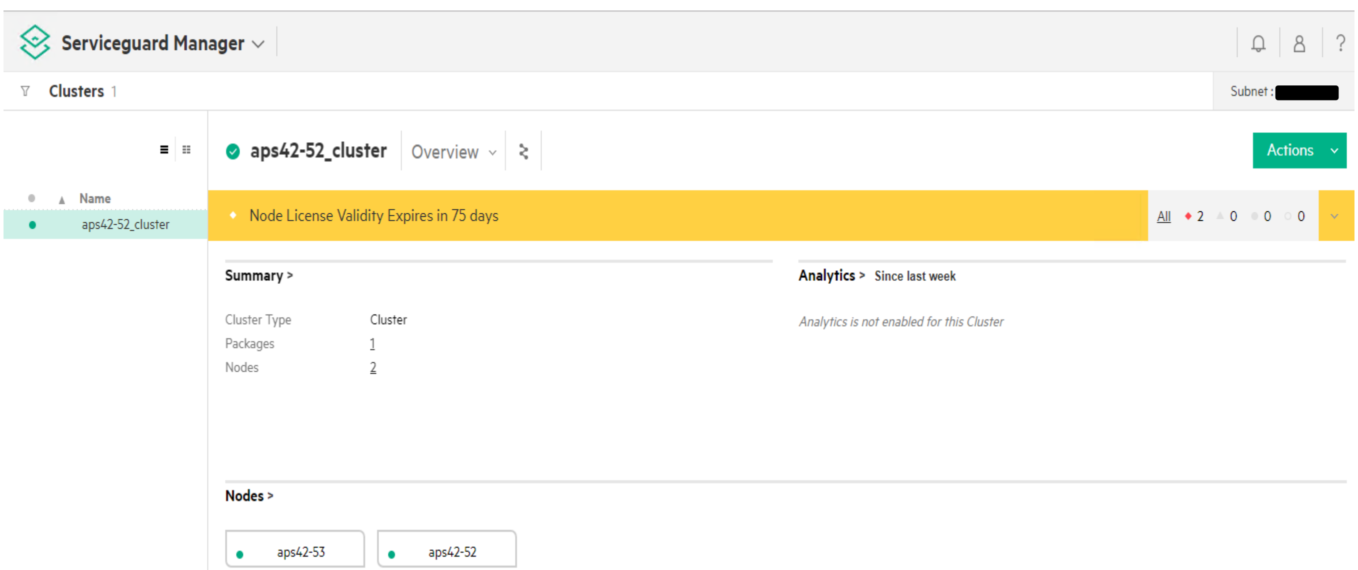
9. Create the new Cluster.



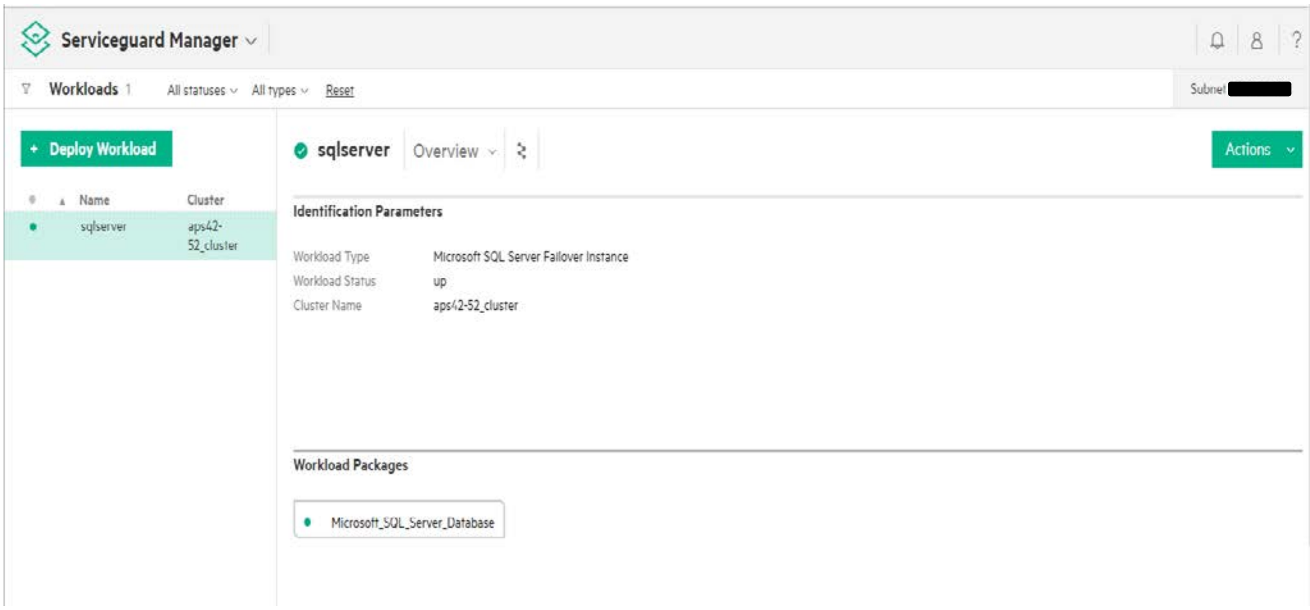
10. Add Lock LUN as Cluster arbitrator.



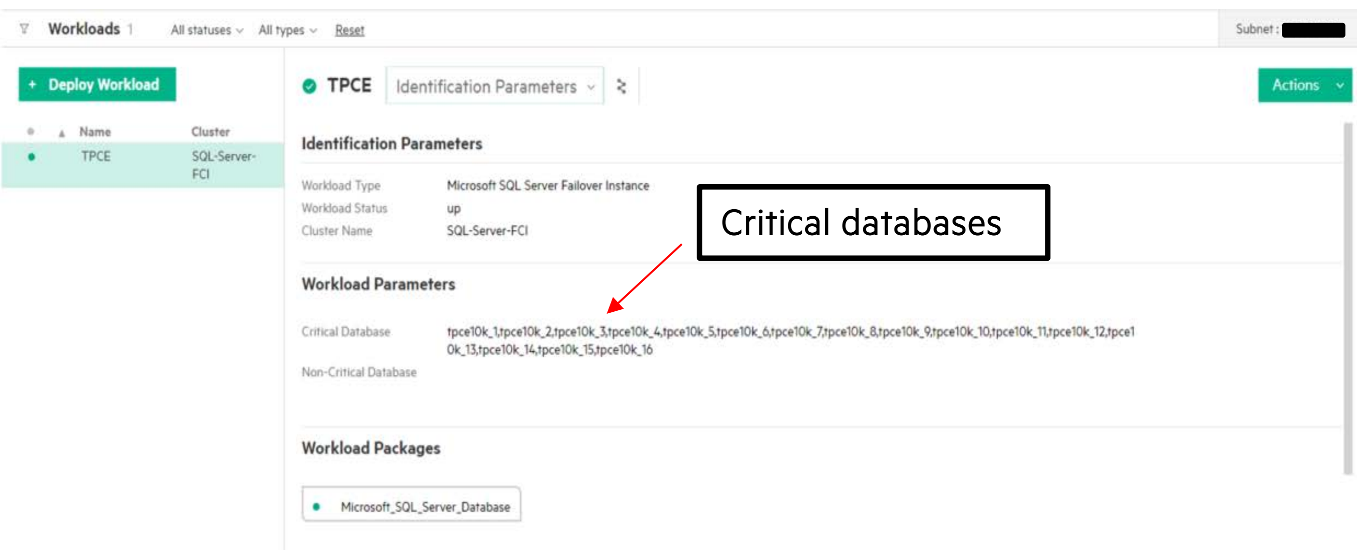
11. Create Cluster creation completed and running.



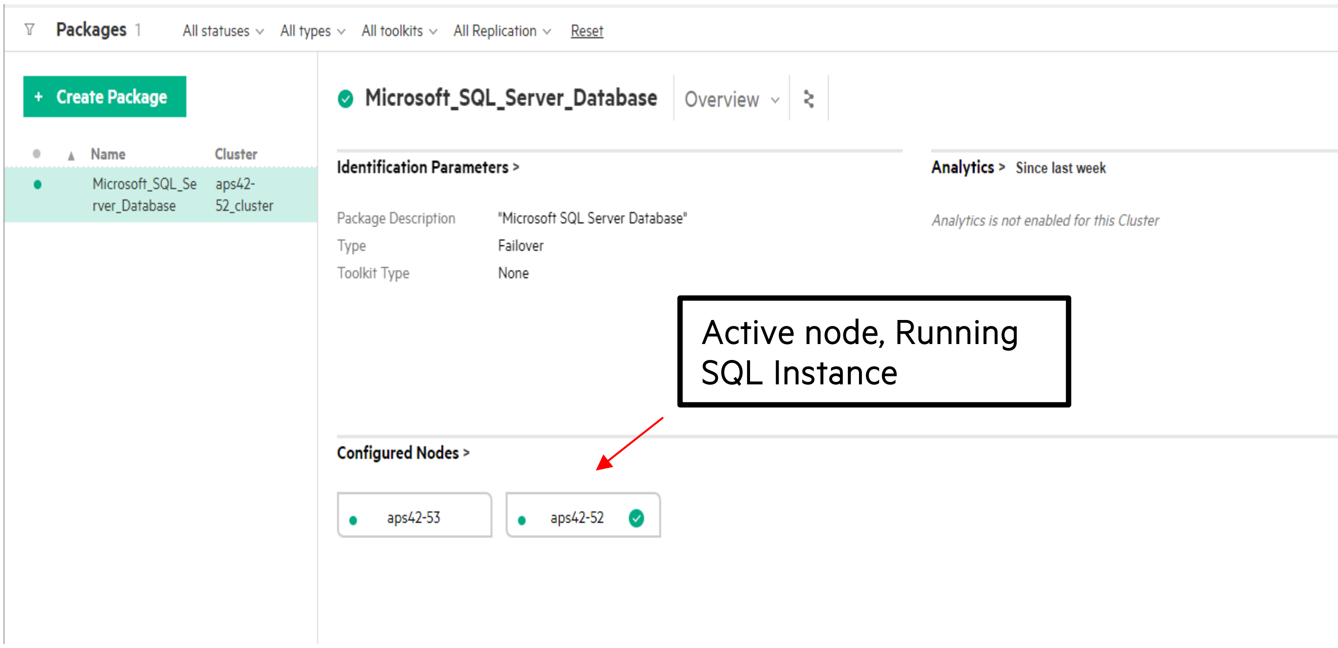
12. Deploy SQL Server Failover workload



Here is a list of critical databases which are part of the Cluster now.



13. Deploy SQL Server Failover workload



Resources and additional links

HPE Reference Architectures, hpe.com/info/ra

HPE Synergy 660 Gen10 Compute Module- QuickSpecs, https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00008522enw

HPE OneView, https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-a00048164en_us&docLocale=en_US

Performance best practices and configuration guidelines for SQL Server on Linux, https://docs.microsoft.com/en-us/sql/linux/sql-server-linux-performance-best-practices?view=sql-server-2017

HPE Nimble Storage Deployment Considerations for Microsoft SQL Server, https://infosight.hpe.com/InfoSight/media/cms/active/public/tmg_HPE_Nimble_Storage_Deployment_Considerations_for_Microsoft_SQL_Server_doc_version_family.pdf

HPE Nimble Storage Deployment Considerations for Linux on Fibre Channel, https://infosight.hpe.com/InfoSight/media/cms/active/pubs_fc_linux_dcg_1.0.0.0.pdf

HPE Serviceguard Solutions for Microsoft SQL Server for Linux User Guide, https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-a00039040en_us&docLocale=en_US

Managing HPE Serviceguard for Linux A. 12.30.00, https://h50146.www5.hpe.com/products/software/oe/linux/mainstream/support/doc/other/ha_cluster/pdfs/emr_201807na-a00052276en_us.pdf

HPE InfoSight GUI Administration Guide, https://infosight.hpe.com/InfoSight/media/cms/active/public/pubs_GUI_Administration_Guide_NOS_50x.whz/index.html

Compare Microsoft SQL Server versions, https://www.microsoft.com/en-us/sql-server/sql-server-2017-comparison#CP_StickyNav_1

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

a00069875enw, April 2019