



# VOSS Automate Architecture and Hardware Specification Guide

Release 25.1

June 05, 2025

## Legal Information

- Copyright © 2025 VisionOSS Limited.  
All rights reserved.
- This information is confidential. If received in error, it must be returned to VisionOSS ("VOSS"). Copyright in all documents originated by VOSS rests in VOSS. No portion may be reproduced by any process without prior written permission. VOSS does not guarantee that this document is technically correct or complete. VOSS accepts no liability for any loss (however caused) sustained as a result of any error or omission in the document.

DOCUMENT ID: 20250605105221

# Contents

<b>1</b>	<b>VOSS Automate</b>	<b>1</b>
1.1	Automate deployment topologies . . . . .	1
1.2	Automate hardware specifications . . . . .	18
1.3	Scale and performance . . . . .	25
	<b>Index</b>	<b>28</b>

# 1. VOSS Automate

## 1.1. Automate deployment topologies

### 1.1.1. Overview

Automate offers two main deployment topologies:

- *Unified node cluster topology*
- *Modular node cluster deployment topology*

Two additional deployment options are available:

- *Cloud deployments*
- *VOSS Automate Cloudv1 (SaaS)*

### 1.1.2. Node types

Automate deployment topologies are comprised of a configuration of the following types of nodes, each performing specific functions within the topologies:

- Web proxy node
- Unified/single node
- Application node
- Database node

Each node type is comprised of one or more of the following components (software subsystems):

Component	Description
Operating system	Ubuntu, stripped down / hardened
Platform	Docker, isolated components
Web server	Nginx, receives and forwards HTTP requests <ul style="list-style-type: none"> <li>• Hosts static files: CSS, JS and images</li> <li>• Load balance between unified nodes (UNs): round robin, configurable, for example, two data centres</li> <li>• Detects inactive UN: removes from round robin</li> </ul>
Database	MongoDB (scalable, distributed), PostgreSQL (scalable)
Application	JavaScript, Python, REST API, device drivers, workflow engine, transactions/queue engine, RBAC, search, bulk loader, and more ...

The matrix outlined in the table describes the set of components in each node type:

Node type	Components				
	Operating system	Platform	Web server	Database	Application
Web proxy	X	X	X		
Unified/single node	X	X	X	X	X
Application	X	X			X
Database	X	X		X	

### 1.1.3. Unified node cluster topology

Automate's **Unified Node Cluster** topology provides the following options:

- *Single-node cluster (cluster-of-one/standalone) (testing-only)*
- *Single-node cluster (cluster-of-one/standalone) with VMWare HA*
- Two node with web proxies
- Four node with web proxies
- Six node with web proxies

**Important:** Choose between a Unified Node deployment or a Modular Architecture deployment.

In a *Unified Node Cluster* deployment, Automate is deployed as *one* of the following:

- A single unified node cluster
- Two unified nodes
- A cluster of multiple nodes with High Availability (HA) and Disaster Recovery (DR) qualities

Each node can be assigned one or more of the following functional roles:

Functional role	Description
Web proxy	Load balances incoming HTTP requests across unified nodes.
Single unified node	Combines the Application and Database roles for use in a non-multi-clustered test environment.
Unified	Similar to the <i>Single unified node</i> role Application and Database roles, but clustered with other nodes to provide HA and DR capabilities.

The nginx web server is installed on the web proxy, *Single Unified Node*, and the *Unified Node Cluster*, but is configured differently for each role.

In a clustered environment containing multiple *Unified Node Clusters*, a load balancing function is required to offer HA (High Availability providing failover between redundant roles).

Automate supports deployment of either the web proxy node or a DNS load balancer. Consider the following when deciding whether to select a web proxy node or a DNS:

- The web proxy node takes load off the *Unified Node Cluster* to deliver static content (HTML/JAVA scripts). When using DNS or a third-party load balancer, the *Unified Node Cluster* must process this information.
- DNS is unaware of the state of the *Unified Node Cluster*.
- The web proxy detects if a *Unified Node Cluster* is down or corrupt. In this case, the web proxy will select the next *Unified Node Cluster* in a round robin scheme.

---

**Important:** It is recommended that you run no more than two *Unified Node Clusters* and one web proxy node on a physical (VMware) server.

Additionally, it is recommended that the disk sub-systems are unique for each *Unified Node Cluster*.

---

The table describes the defined deployment topologies for test and production:

Deployment topology	Description
Test	<p>A standalone, <i>Single Unified Node</i>, with Application and Database roles combined.</p> <p>No high availability or disaster recovery (HA/DR) is available.</p> <hr/> <p><b>Important:</b> A test deployment must be used only for test purposes.</p> <hr/>
Production with unified nodes	<p>In a clustered system, comprising:</p> <ul style="list-style-type: none"><li>• Two, three, four, or six unified nodes (each with combined Application and Database roles)</li><li>• Zero to four (maximum two if two unified nodes) web proxy nodes offering load balancing.</li></ul> <p>The web proxy nodes can be omitted if an external load balancer is available.</p>

#### Single-node cluster (cluster-of-one/standalone) (testing-only)

**Note:** A *Single-node cluster (cluster-of-one/standalone)* deployment should be used *only* for test purposes.



The table describes the advantages and disadvantages of a *Single-node cluster (cluster-of-one/standalone)* deployment topology:

Advantages	Disadvantages
<ul style="list-style-type: none"> <li>• Smallest hardware footprint</li> </ul>	<ul style="list-style-type: none"> <li>• No high availability or disaster recovery</li> <li>• Less throughput than clusters</li> </ul>

### Single-node cluster (cluster-of-one/standalone) with VMWare HA

The table describes the advantages and disadvantages of a *Single-node cluster (cluster-of-one/standalone)* with VMWare HA deployment topology:

Advantages	Disadvantages
<ul style="list-style-type: none"> <li>• Smallest hardware footprint</li> <li>• Disaster recovery available</li> </ul>	<ul style="list-style-type: none"> <li>• Less throughput than clusters</li> </ul>

### Multi-node cluster with unified nodes

To achieve geo-redundancy using the unified nodes, consider the following:

- Either four or six unified nodes (each node combining Application and Database roles), are clustered and split over two geographically disparate locations.
- Two web proxy nodes to provide high availability, ensuring that an Application role failure is gracefully handled. More may be added if web proxy nodes are required in a DMZ.

---

**Important:** It is strongly recommended *not* to allow customer end-users the same level of administrator access as the restricted groups of Provider- and Customer administrators. For this reason, Self-service web proxies as well as Administrator web proxies should be used.

Systems with Self-service-only web proxies are *only* recommended where the system is customer facing, but where the customer does not administer the system themselves.

---

- Web proxy and unified nodes can be contained in separate, firewalled networks.
- Database synchronization takes places between all database roles, thus offering disaster recovery and high availability.
- For six unified nodes, all nodes in the cluster are active. For an eight node cluster (with latency between data centers greater than 10ms), the two nodes in the disaster recovery node are passive; that is, the `voss workers @ command` has been run on the disaster recovery nodes.

---

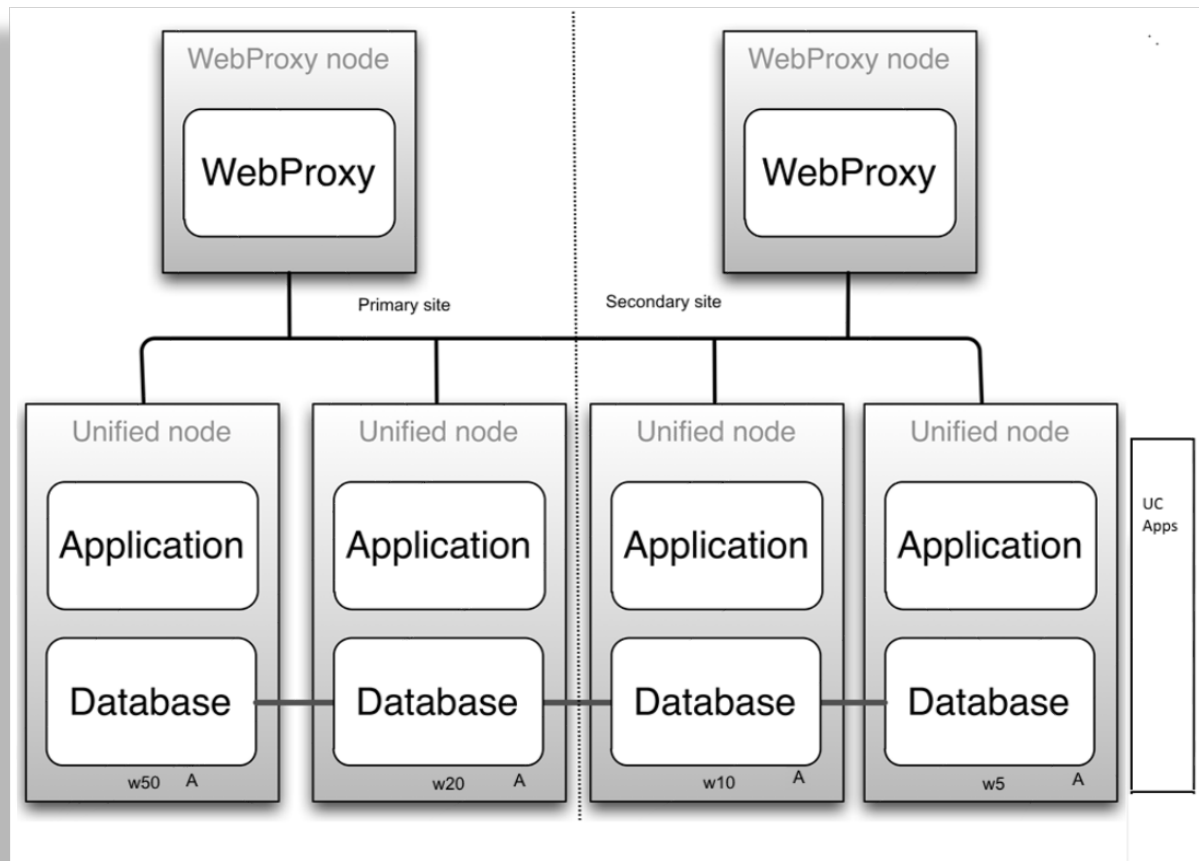
**Note:** Primary and fall-back secondary database servers can be configured manually. Refer to the *Automate Platform Guide* for details.

---

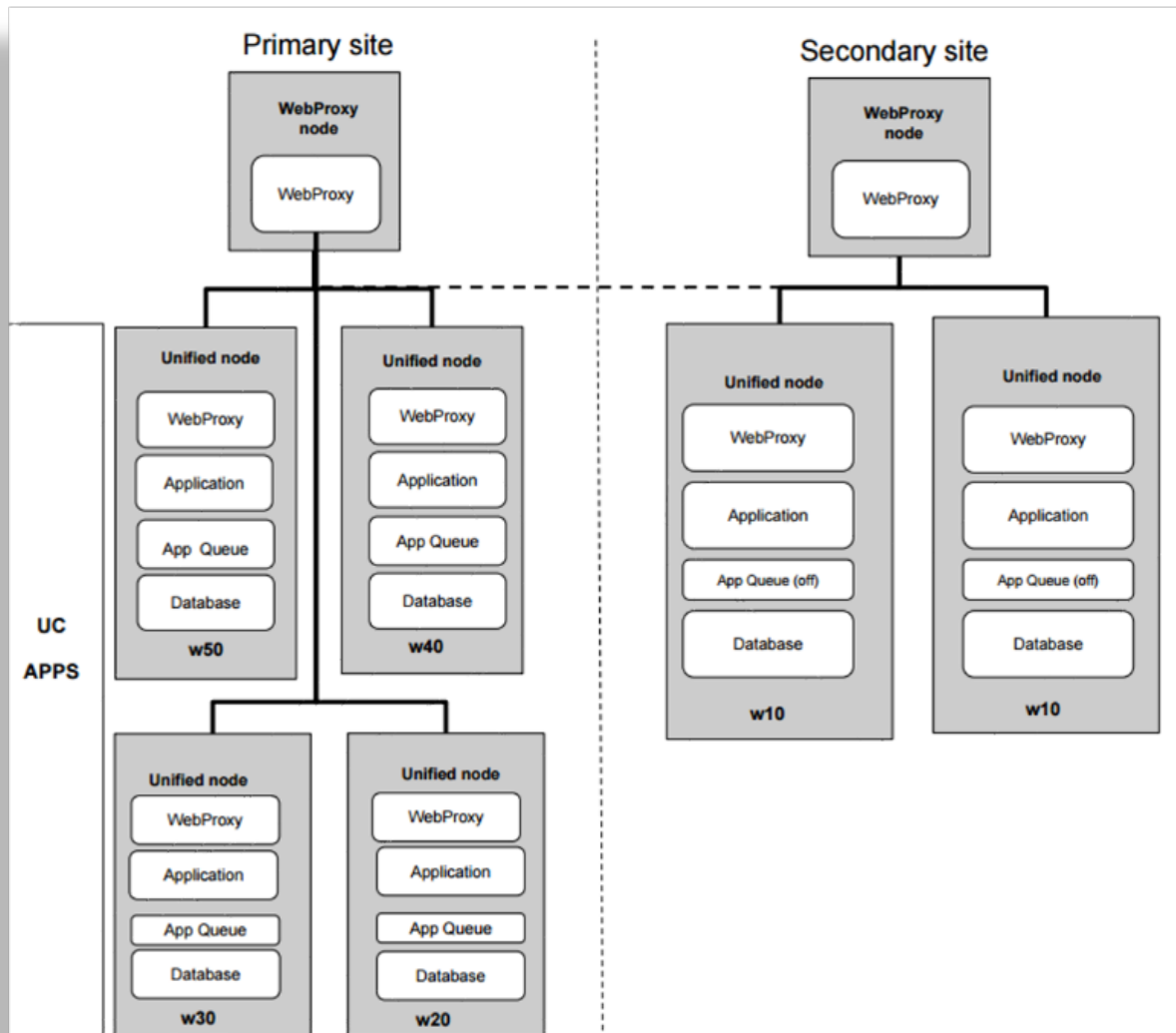


**Example: Six node cluster**

The diagram illustrates an example of a *six node cluster*:

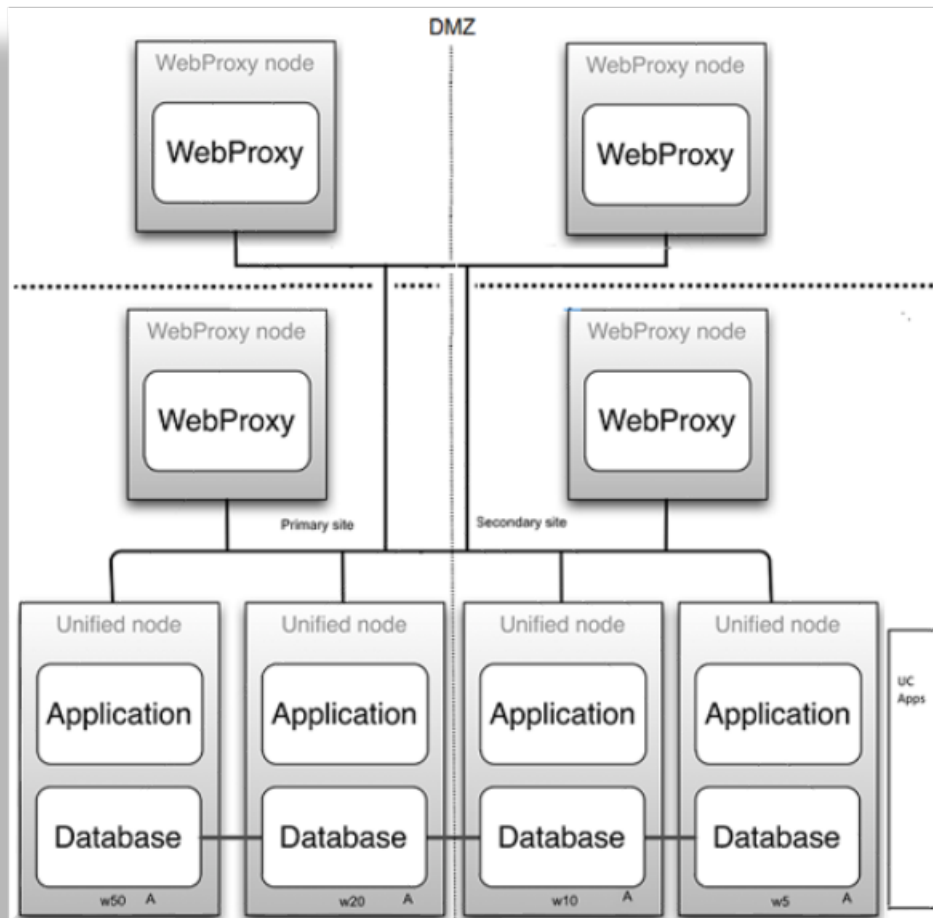
**Example: Eight node cluster**

The diagram illustrates an example of an *eight node cluster*:



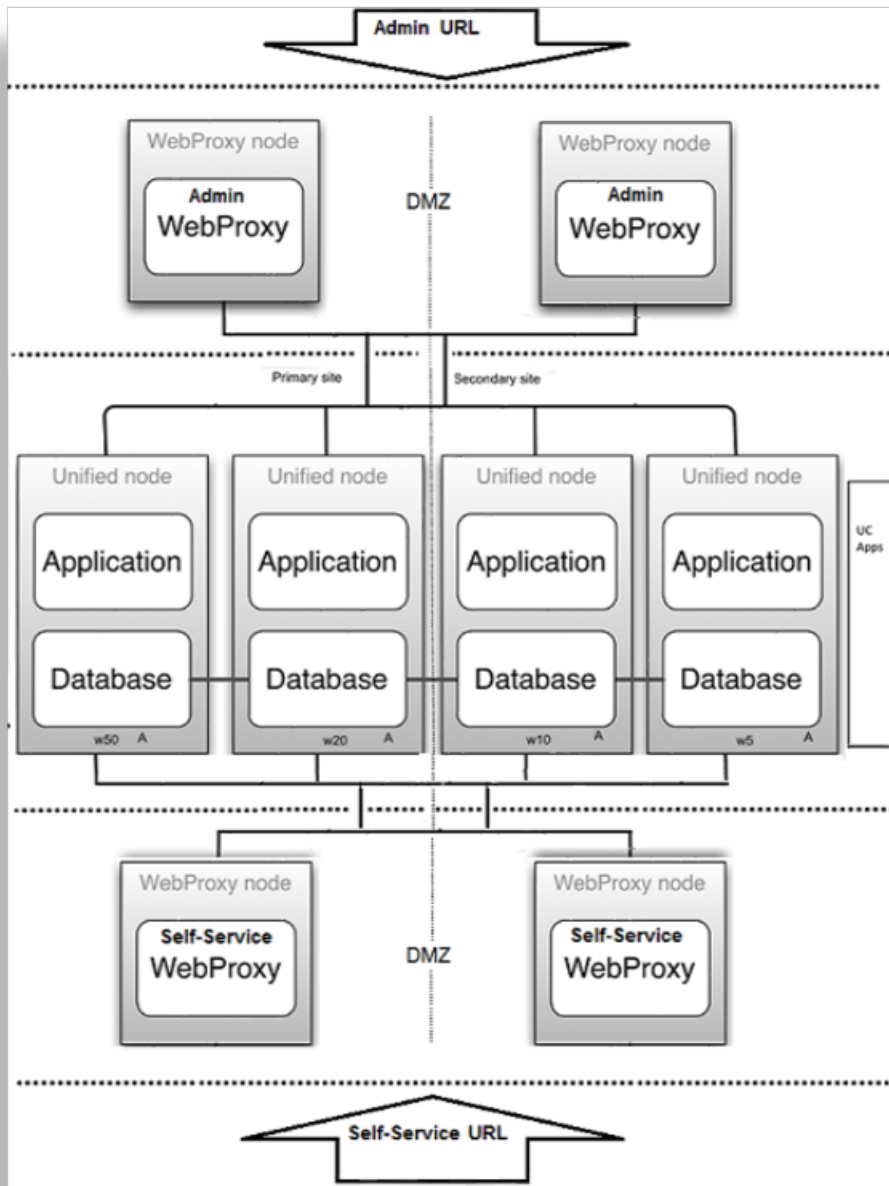
### Example: Two web proxy nodes in a DMZ

The diagram illustrates an example of *two web proxy nodes in a DMZ*:



**Example: Four web proxy nodes in a DMZ (two admin, two Self-service)**

The diagram illustrates an example of *four web proxy nodes (2 admin, and 2 Self-service)* in a DMZ:



### Two node cluster with unified nodes

To achieve geo-redundancy using the unified nodes, consider the following:

- Two unified nodes (each node combining application and database roles) are clustered and optionally split over two geographically disparate locations.
- (Optional) Two web proxy nodes can be used. It may be omitted if an external load balancer is available.
- Web proxy and unified nodes can be contained in separate firewalled networks.
- Database synchronization takes place from primary to secondary unified nodes, thereby offering disaster recovery if the primary node fails.

- If the secondary unified node has *more than 10ms latency* with the primary unified node, it must be configured to be in the *same* geographical location.

---

**Important:**

With only two unified nodes, with or without web proxies, there is no high availability. The database on the primary node is read/write, while the database on the secondary is read-only.

Only redundancy is available in the following instances:

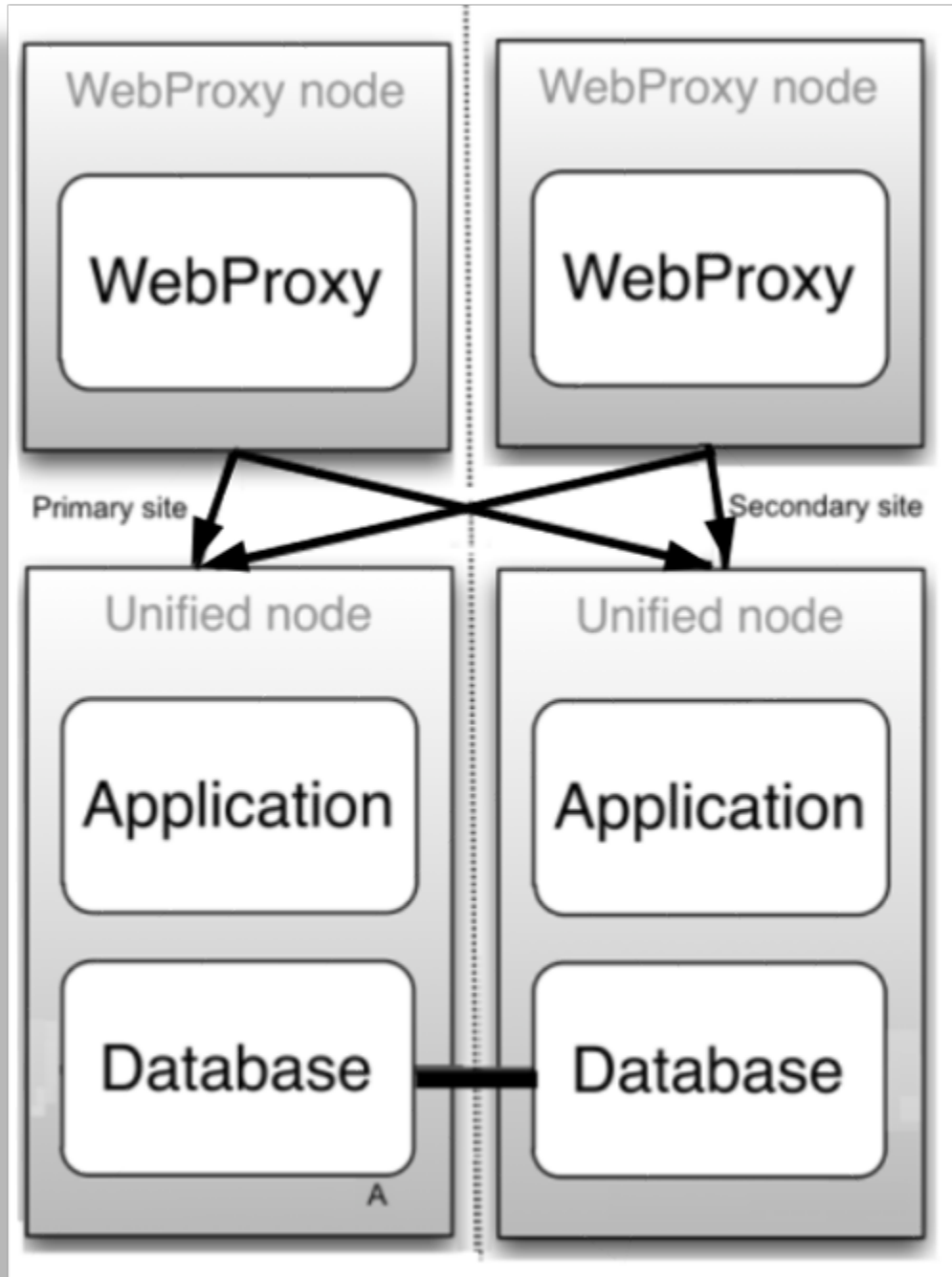
- If the primary node fails, a manual delete of the primary node on the secondary and a cluster provision will be needed.
- If the secondary node fails, it needs to be replaced.

Refer to the topic on *Disaster recovery failover and recovery in a two node cluster* in the Platform Guide.

---

**Example: Two node cluster**

The diagram illustrates a *two node cluster*:



### Four node with web proxies

The table describes the advantages and disadvantages of a *four node with web proxies* deployment topology:

Advantages	Disadvantages
<ul style="list-style-type: none"> <li>• More disaster recovery scenarios supported</li> <li>• More throughput than 3 Node</li> </ul>	<ul style="list-style-type: none"> <li>• More hardware than 3 Node</li> </ul>

### Six node with web proxies

The following are characteristics of a *six node with web proxies* deployment topology:

- Typically deployed for multi-data center deployments
- Supports Active/Standby

## 1.1.4. Modular node cluster deployment topology

### Overview

A *modular node cluster* topology has separate Application and Database nodes:

- Three Database nodes
- One to eight Application nodes
- Web proxies

A *modular node cluster\** topology has the following advantages:

- Increased processing capacity
- Horizontal scaling by adding more Application nodes
- Improved database resilience with dedicated nodes and isolation from application
- Improved database performance by removing application load from the primary database

---

**Important:** Choose between a *Unified Node Cluster* deployment or a *Modular Node Cluster* deployment.

---

Automate is deployed as a *Modular Node Cluster* of multiple nodes, with High Availability (HA) and Disaster Recovery (DR) qualities.

Each node can be assigned one or more of the following functional roles:

Functional role	Description
Web proxy	Load balances incoming HTTP requests across nodes.
Application role node	Clustered with other nodes to provide HA and DR capabilities.
Database role node	Clustered with other nodes to provide HA and DR capabilities.

The nginx web server is installed on the web proxy and application role node, but is configured differently for each role.

### Related topics

Modular Architecture Multi-node Installation in the Install Guide.

Migrate a Unified Node Cluster to a Modular Node Cluster in the Platform Guide.

A load balancing function is required to offer HA (High Availability providing failover between redundant roles).

Automate supports deployment of either the web proxy node or a DNS load balancer. When choosing between a web proxy node and a DNS, consider the following:

- The web proxy takes load off the Application role node to deliver static content (HTML/JAVA scripts). When using DNS or a third-party load balancer, the Application role node has to process this information.
- DNS is unaware of the state of the Application role node.
- The web proxy detects if an Application role node is down or corrupt. In this case, the web proxy will select the next Application role node in a round robin scheme.

---

**Important:** It is recommended that you run no more than one Application role node and one Database role node and one web proxy node on a physical (VMWare) server. When choosing disk infrastructure, high volume data access by database role replica sets must be considered where different disk sub-systems may be required depending on the performance of the disk infrastructure.

---

The following *Modular Node Cluster* topology is recommended (minimum):

---

**Important:** *Single Unified Node* topologies are not available for *Modular Node Cluster* deployments.

---

- Production with nodes (in a clustered system of two data centers):
  - DC1 = Data center 1, a primary data center containing primary database node (highest database weight)
  - DC2 = Data center 2, a data recovery data center

The system comprises of the following nodes:

- Three nodes with application roles (two in DC1; one in DC2)
- Three nodes with database roles (two in DC1; one in DC2)



- Maximum two web proxy nodes if two data centers; offering load balancing. The web proxy nodes can be omitted if an external load balancer is available.

### Multi-node modular node cluster with application and database nodes

To achieve geo-redundancy using Application and Database nodes, consider the following:

- Six Application and Database nodes (three nodes with an application role and three nodes with a database role) are clustered and split over two geographically disparate locations.
- Two web proxy nodes to provide High Availability so that an Application role failure is gracefully handled. More may be added if web proxy nodes are required in a DMZ.

---

**Important:** It is strongly recommended *not* to allow customer end-users the same level of administrator access as the restricted groups of Provider- and Customer administrators. For this reason, Self-service web proxies as well as Administrator web proxies should be used.

Systems with Self-service-only web proxies are *only* recommended where the system is customer facing, but where the customer does not administer the system themselves.

---

- Web proxy, Application and Database nodes can be contained in separate firewalled networks.
- Database synchronization takes places between all database role nodes, thus offering disaster recovery and high availability.
- All nodes in the cluster are active.

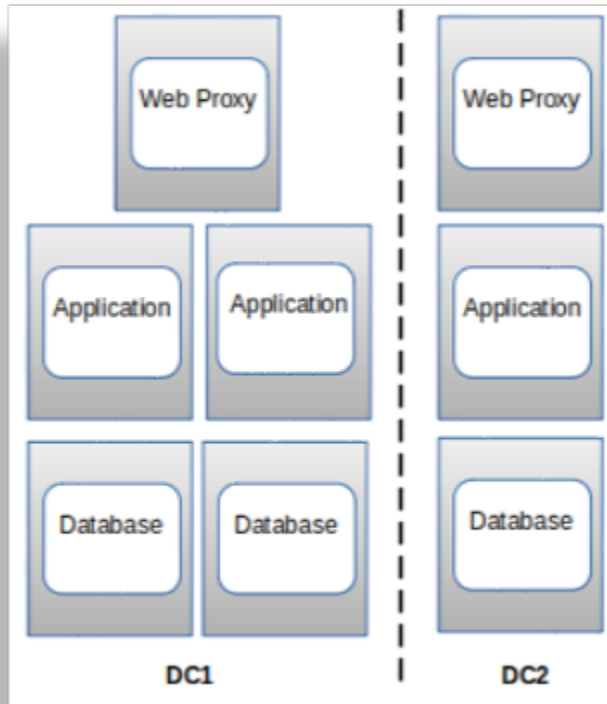
---

**Note:** Primary and fall-back secondary database servers can be configured manually. Refer to the *Automate Platform Guide* for details.

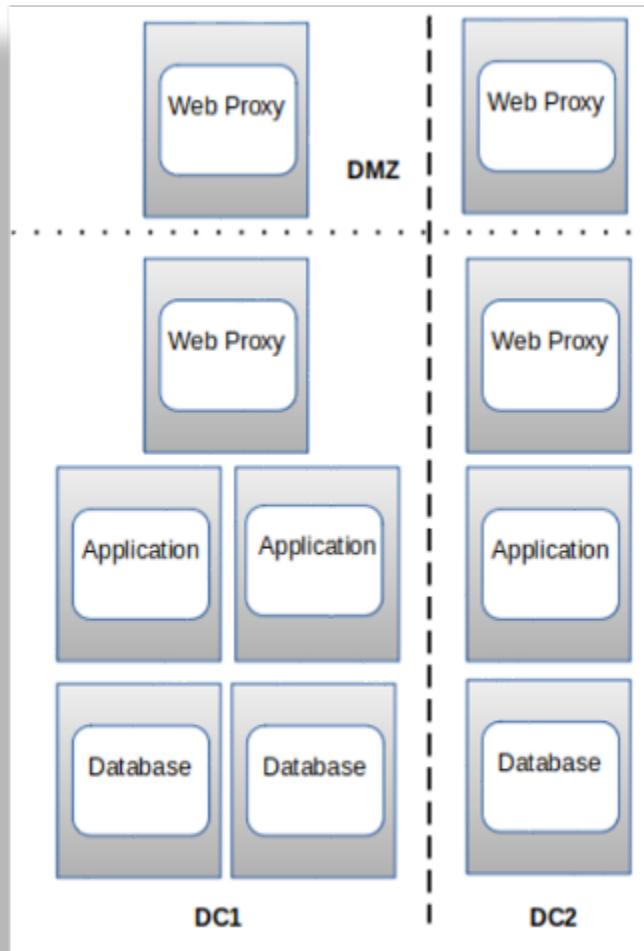
---

### Example: Six node cluster

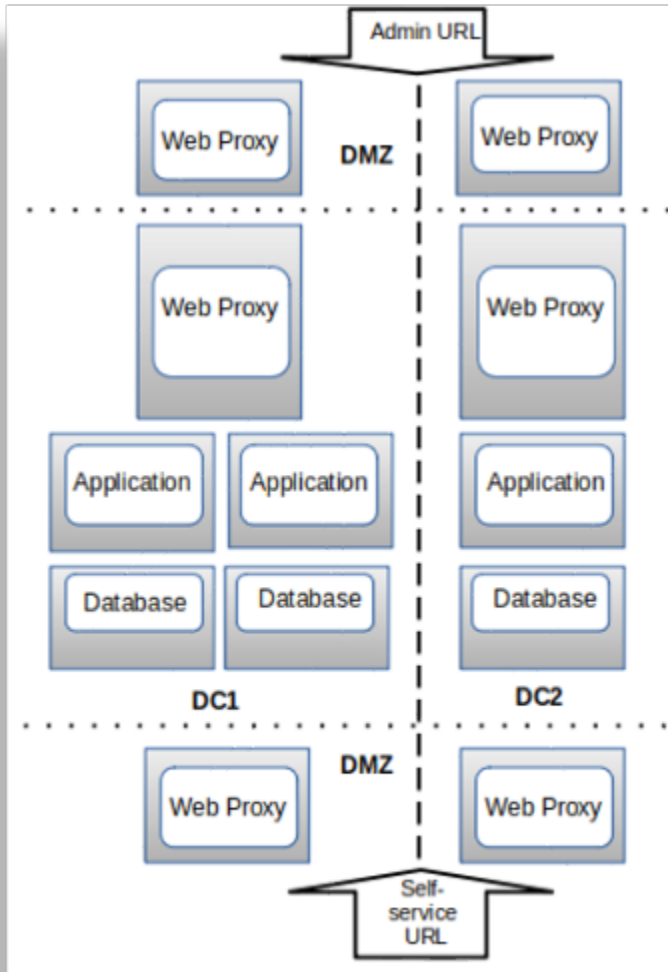
The diagram illustrates an example of a *six node cluster*:

**Example: Two web proxy nodes in a DMZ**

The diagram illustrates an example of *two web proxy nodes in a DMZ*:

**Example: Four web proxy nodes in a DMZ**

The diagram illustrates an example of *four web proxy nodes in a DMZ* (two admin, two Self-service):



### 1.1.5. Cloud deployments

Automate supports the following Cloud deployments:

- Microsoft Azure
- Amazon Web Services (AWS)

Although Google Cloud Platform (GCP) is not officially supported, contact us to discuss your requirements.

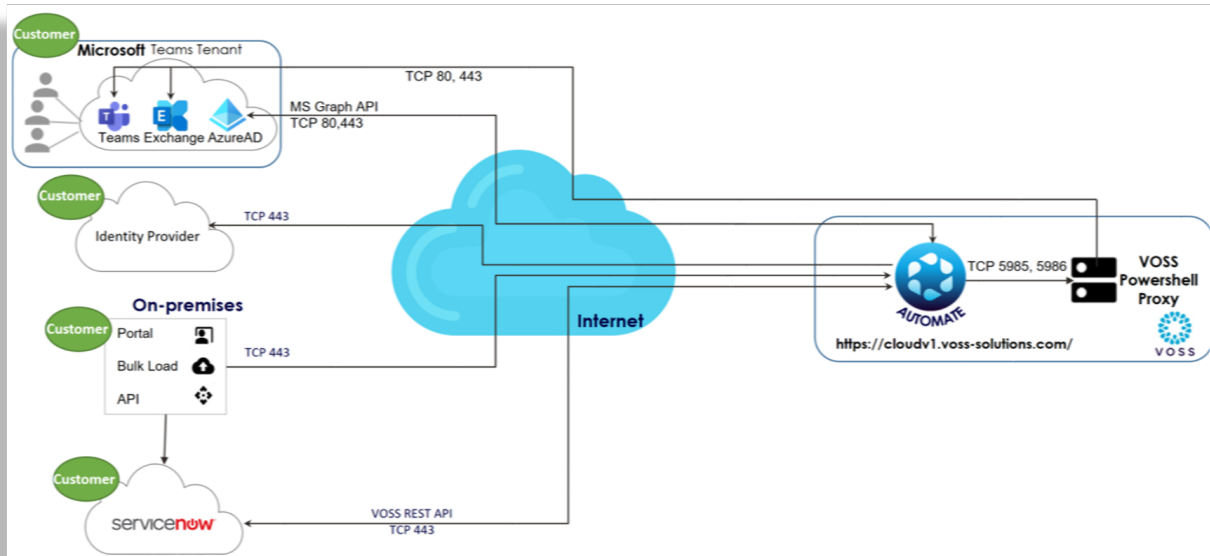
Advantages of a Cloud deployment topology:

- Leverage cloud tooling, such as proxies (which can be used instead of VOSS Web Proxy)

### 1.1.6. VOSS Automate Cloudv1 (SaaS)

VOSS Automate Cloudv1 is a Software-as-a-Service (SaaS) offering hosted on a shared VOSS Automate instance within Microsoft Azure.

VOSS manages this instance, which seamlessly integrates with a customer's Unified Communications (UC) platform, Microsoft Exchange, Microsoft Active Directory, and third-party applications, such as ServiceNow and Identity Providers (IdPs) for Single Sign-On (SSO) authentication.



## 1.2. Automate hardware specifications

### 1.2.1. Overview

**Note:** For details around the open source software components used in Automate, see the *Open Source License Usage Guide*.

#### Virtualized hardware and resource oversubscription

It is recommended that no more than two Unified nodes and one Web Proxy node be run on a physical server (VMware server) and that the disk subsystems are unique for each Unified node.

Automate virtual machines should maintain a 1:1 ratio between virtual RAM and Disk hardware and physical hardware, in other words:

- 1 GB of virtual RAM (vRAM) must map to 1 GB of physical RAM
- 1 GB of virtual Disk (vDisk) storage must map to 1 GB of physical storage

For virtual CPU (vCPU), hyper-threading is supported.

### 1.2.2. Unified node hardware specifications

#### Single-node cluster (cluster-of-one) hardware specification

This section provides the virtual machine specification for a single node cluster deployment topology in Automate.

Node type	Quantity	VM	Memory	CPU	Disk	Network
Single node cluster	1	At least VMware 11, compatible with ESXi 6.0 and up	32 GB with 32 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	Total: 440 GB as allocated below. 370 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application:               <ul style="list-style-type: none"> <li>– 10 GB for logs</li> <li>– 40 GB for our apps</li> </ul> </li> <li>• 50 GB for compressed backups</li> <li>• 250 GB for database</li> </ul> 70 GB database disk to be added after upgrade or installation. Refer to the Upgrade Guide with ISO and Template or Installation Guide.	1 Gbit/s minimum

For Memory and CPU, the Resource Allocation Reservation on VMware is indicated in the table. Hyper-threading is supported.

#### Note:

- From release 24.1, allowance should be made for an additional 70GB database disk to be added after upgrade or installation. This disk will be used for Insights sync.  
Refer to the Upgrade Guide with ISO and Template or Installation Guide and [Add disks in the AWS or MS Azure cloud hosted platform](#).
- If memory allocations are customized, ensure that the memory reservation remains equal to the allocated memory in order to prevent possible negative side-effects due to memory reclamation.

For VMware details, refer to the VMware Support topic in the Compatibility Matrix.

The maximum number of users for a single node cluster is 50,000.

### Multinode cluster hardware specification

Virtual machine requirements are specified in the table below.

Node type	Quantity	VM	Memory	CPU	Disk	Network
Unified	4 or 6	At least VMware 11, compatible with ESXi 6.0 and up	32 GB with 32 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	Total: 440 GB as allocated below. 370 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application: <ul style="list-style-type: none"> <li>– 10 GB for logs</li> <li>– 40 GB for our apps</li> </ul> </li> <li>• 50 GB for compressed backups</li> <li>• 250 GB for database</li> </ul> 70 GB database disk to be added after upgrade or installation. Refer to the Upgrade Guide with ISO and Template or Installation Guide.	1 Gbit/s minimum
WebProxy	2	At least VMware 11, compatible with ESXi 6.0 and up	4 GB with 4 GB reservation	2 vCPU @ 2 GHz with no reservation	70 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application</li> </ul>	1 Gbit/s minimum

For Memory and CPU, the Resource Allocation Reservation on VMware is indicated in the table. Hyper-threading is supported.

#### Note:

- From release 24.1, allowance should be made for an additional 70GB database disk to be added after upgrade or installation. This disk will be used for Insights sync.  
Refer to the Upgrade Guide with ISO and Template or Installation Guide and [Add disks in the AWS or MS Azure cloud hosted platform](#).
- If memory allocations are customized, ensure that the memory reservation remains equal to the allocated memory in order to prevent possible negative side-effects due to memory reclamation.

The OS disk requirement is fixed and logs are rotated to ensure that 10 GB is sufficient. 40 GB for applications is a generous allocation and does not scale with the number of users.

The Database storage partition is sized to support 250 K users. Database backups are compressed and the partition is sized to ensure that sufficient space available to support backup of 250 GB database.

---

**Note:** To change the TRANSACTION\_LOG cap size to greater than 10GB at larger providers for operational reasons, for example for diagnostics and a longer transaction replay window, the `voss db_collection_cap TRANSACTION_LOG <10-50GB>` command can be used from the command line.

Refer to Database Commands in the Platform Guide for more details.

The resize operation will impact the usage on the disk size allocated for the database (typically, 250GB is reserved upon installation). Consider a larger database disk size allocation upon installation if a larger cap size is set.

---

The backup disk should be Thick Provisioned and Eager Zeroed for better performance immediately after installation.

Web Proxies are optional, but if Web Proxies are used, then they form part of the cluster to allow sharing of static data and other content as needed (for example, themes).

To set up the disk requirements, the disk should be set up on the VMware GUI Resources tab where a disk can be created. This task should be done after the OVA import but prior to the boot of the system.

For VMware details, refer to the VMware Support topic in the Compatibility Matrix.

## 2 Node cluster hardware specification

Virtual machine requirements are specified in the table below.



Node type	Quantity	VM	Memory	CPU	Disk	Network
Unified	= 2	At least VMware 11, compatible with ESXi 6.0 and up	32 GB with 32 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	Total: 440 GB as allocated below. 370 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application: <ul style="list-style-type: none"> <li>– 10 GB for logs</li> <li>– 40 GB for our apps</li> </ul> </li> <li>• 50 GB for compressed backups</li> <li>• 250 GB for database</li> </ul> 70 GB database disk to be added after upgrade or installation. Refer to the Upgrade Guide with ISO and Template or Installation Guide.	1 Gbit/s minimum
WebProxy	>= 0	At least VMware 11, compatible with ESXi 6.0 and up	4 GB with 4 GB reservation	2 vCPU @ 2 GHz with no reservation	70 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application</li> </ul>	1 Gbit/s minimum

For Memory and CPU, the Resource Allocation Reservation on VMware should correspond with these requirements.

**Note:**

- From release 24.1, allowance should be made for an additional 70GB database disk to be added after upgrade or installation. This disk will be used for Insights sync.

Refer to the Upgrade Guide with ISO and Template or Installation Guide and [Add disks in the AWS or MS Azure cloud hosted platform](#).

- If memory allocations are customized, ensure that the memory reservation remains equal to the allocated memory in order to prevent possible negative side-effects due to memory reclamation.

For VMware details, refer to the VMware Support topic in the Compatibility Matrix.

### 1.2.3. Modular cluster hardware specifications

#### Multinode modular cluster hardware specification

Virtual machine requirements are specified in the table below.

Node type	Quantity	VM	Memory	CPU	Disk	Network
Application	3	At least VMware 11, compatible with ESXi 6.0 and up	32 GB with 32 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	70 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application: <ul style="list-style-type: none"> <li>– 10 GB for logs</li> <li>– 40 GB for our apps</li> </ul> </li> </ul>	1 Gbit/s minimum
Database	3	At least VMware 11, compatible with ESXi 6.0 and up	32 GB with 32 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	Total: 440 GB as allocated below. 370 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for compressed backups</li> <li>• 50 GB for application: <ul style="list-style-type: none"> <li>– 10 GB for logs</li> <li>– 40 GB for our apps</li> </ul> </li> <li>• 250 GB for database</li> </ul> 70 GB database disk to be added after upgrade or installation. Refer to the Upgrade Guide with ISO and Template or Installation Guide.	1 Gbit/s minimum
WebProxy	2	At least VMware 11, compatible with ESXi 6.0 and up	4 GB with 4 GB reservation	2 vCPU @ 2 GHz with no reservation	70 GB partitioned: <ul style="list-style-type: none"> <li>• 20 GB for OS</li> <li>• 50 GB for application</li> </ul>	1 Gbit/s minimum

For Memory and CPU, the Resource Allocation Reservation on VMware is indicated in the table. Hyper-threading is supported.

#### Note:

- From release 24.1, allowance should be made for an additional 250GB database disk to be added after

upgrade or installation. This disk will be used for Insights sync.

Refer to the Upgrade Guide with ISO and Template or Installation Guide and [Add disks in the AWS or MS Azure cloud hosted platform](#).

- If memory allocations are customized, ensure that the memory reservation remains equal to the allocated memory in order to prevent possible negative side-effects due to memory reclamation.

The OS disk requirement is fixed and logs are rotated to ensure that 10 GB is sufficient. 40 GB for an applications role node is a generous allocation and the size will not have to be increased with the number of users.

The Database storage partition is sized to support 250 K users. Database backups are compressed and the partition is sized to ensure that sufficient space is available to support backup of 250 GB database.

**Note:** To change the TRANSACTION\_LOG cap size to greater than 10GB at larger providers for operational reasons, for example for diagnostics, the **voss db\_collection\_cap TRANSACTION\_LOG <10-50GB>** command can be used from the command line.

Refer to Database Commands in the Platform Guide for more details.

The resize operation will impact the usage on the size of the disk allocated for the database (typically, 250GB is reserved upon installation). Consider a larger database disk size allocation upon installation if a larger cap size is set.

The backup disk should be Thick Provisioned and Eager Zeroed for better performance immediately after installation.

Web Proxies are optional, but if Web Proxies are used, then they form part of the cluster to allow sharing of static data and other content as needed (for example, themes).

To set up the disk requirements, the disk should be set up on the VMware GUI Resources tab where a disk can be created. This task should be done after the OVA import but prior to the boot of the system.

For VMware details, refer to the VMware Support topic in the Compatibility Matrix.

#### 1.2.4. Add disks in the AWS or MS Azure cloud hosted platform

The steps below are required to add a disk that provides for the Insights database in release 24.1 - that should then be assigned to the `insights-voss-sync:database` mount point (refer to the final step in the *Upgrade Guide with ISO and Template* for your topology).

##### AWS

1. Create the EBS Volumes for each DB node in the Amazon EC2 console.

Go to **EC2 > Volumes > Create volume**

For **Volume settings**, enter:

- Volume type: Provisioned IOPS SSD (io2)
- Size (GiB): 70GB
- IOPS: 750

For **Availability Zone**:

- Create 3 volumes in each of the zones (for example: us-east-1a, us-east-1b, us-east-1c)
2. Attach the newly created volumes to each of the database nodes.

Go to **EC2 > Volumes > volume\_id > Attach volume**

- **Instance:** Select the database instance within the same corresponding az
- **Device Name:** /dev/sde (This will display as xvde in drives list)

## Microsoft Azure

1. In the Microsoft Azure portal, search for Virtual Machines
  - Select each of the database nodes
  - Select **Disk** under **Properties**
2. Click **Create and attach a new disk**.
  - **LUN:** Next available
  - **Disk Name:** Label according to your recommended naming convention
  - **Storage Type:** Premium SSD LRS
  - **Size:** 70GB
  - **Encryption:** Set according to your requirements
  - **Host Caching:** Read/Write

## 1.3. Scale and performance

### 1.3.1. Overview

This section describes the supported configurations and corresponding scale for unified deployment topologies and modular deployment topologies, including their Geo-Redundancy and Round Trip Time (RTT) requirements.

The supported configurations and scale are based on a standard profile for each user, which includes the following:

- Two endpoints (physical phone and soft client) per user.
- Voice Mail
- Extension Mobility (EM)
- Single Number Reach (SNR)
- IM and Presence

Scale limits are not enforced by the system. However, exceeding the limits can result in significant performance degradation.

With regards Geo-Redundancy options, the table describes the difference between Active and Standby nodes:

Active Node	A Node that can process transactions
Standby Node	A Node that is powered on and running the Automate software, but cannot process transactions.

For RTT across different UC components, this should not exceed 400ms. RTT will impact the duration of a transaction, but only for the remote leg, in other words, performing an action on the remote device.

### 1.3.2. Unified node deployments

Configuration	Number of Unified Nodes	Number of Web Proxy Nodes	Supported Scale (# Users)	Geo-Redundancy (Y/N)
Single node cluster Automate	1	0	50,000	NA
Two-Node Automate Across Data Centers	2	1 <sup>1</sup>	50,000	Yes <sup>2</sup> (Active-Standby) If RTT <= 10ms
Multi-Node Automate Across Data Centers	4	2 <sup>1</sup>	500,000	Yes (Active-Active) If RTT <= 10ms
	6	2 <sup>1</sup>	500,000	Yes (Active-Standby)
Multi-Node Automate One Data Center	4	2 <sup>1</sup>	500,000	No

<sup>1</sup> Two web proxy nodes can be used. It may be omitted if an external load balancer is available.

<sup>2</sup> With only two Unified nodes, with or without Web proxies, there is no High Availability. The database on the primary node is read/write, while the database on the secondary is read only.

Only redundancy is available.

- If the primary node fails, a manual delete of the primary node on the secondary and a cluster provision will be needed.
- If the secondary node fails, it needs to be replaced.

Refer to the topic on DR Failover and Recovery in a 2 Node Cluster in the Platform Guide.

### 1.3.3. Modular node deployments

Configuration	Number of Modular Nodes	Number of Web Proxy Nodes	Supported Scale (# Users)	Geo-Redundancy (Y/N)
Application Automate Across Data Centers	1-8 <sup>3</sup>	2 <sup>1</sup>	500,000	Yes (Active-Active) If RTT <= 10ms
	1	2 <sup>Page 26, 1</sup>	500,000	Yes (Active-Standby)
Modular Database Automate Across Data Centers	2	2 <sup>Page 26, 1</sup>	500,000	Yes If RTT <= 10ms
	1	2 <sup>Page 26, 1</sup>	500,000	Yes
Multi-Node Automate One Data Center	4	2 <sup>Page 26, 1</sup>	500,000	No
Modular Application Automate One Data Center	1-8 <sup>3</sup>	2 <sup>Page 26, 1</sup>	500,000	No
Modular Database Automate One Data Center	3	2 <sup>Page 26, 1</sup>	500,000	No

<sup>3</sup> Horizontal scaling of Application nodes allows for increased API performance and transaction throughput.

# Index

## V

voss

voss db\_collection\_cap, 20

voss workers, 5, 14

## W

web

web service, 5, 14