VOSS-4-UC
Architecture and Hardware Specification Guide

Release 19.3.2

Apr 08, 2020

Legal Information

Please take careful note of the following legal notices:

- Copyright © 2020 VisionOSS Limited.
  All rights reserved.

- VOSS, VisionOSS and VOSS-4-UC are trademarks of VisionOSS Limited.

- No part of this document may be reproduced or transmitted in any form without the prior written permission of VOSS.

- VOSS does not guarantee that this document is technically correct, complete, or that the product is free from minor flaws. VOSS endeavors to ensure that the information contained in this document is correct, whilst every effort is made to ensure the accuracy of such information, VOSS accepts no liability for any loss (however caused) sustained as a result of any error or omission in the same.

- This document is used entirely at the users own risk. VOSS cannot be held responsible or liable for any damage to property, loss of income, and or business disruption arising from the use of this document.

- The product capabilities described in this document and the actual capabilities of the product provided by VOSS are subject to change without notice.

- VOSS reserves the right to publish corrections to this document whenever VOSS deems it necessary.

- The terms Cisco, Movius, MeetingPlace, Netwise and all other vendor/product names mentioned in this document are registered trademarks and belong to their respective owners. VOSS does not own, nor is related to, these products and vendors. These terms have been included to showcase the potential of the VOSS solution and to simplify the deployment of these products with VOSS should you select to utilize them.

Security Information

This product may contain cryptographic features that may be subject to state and local country laws that govern the import, export, transfer and use of such features. The provision of this software does not imply that third-party authorization to import, export, distribute or use encryption in your particular region has been obtained. By using this product, you agree to comply with all applicable laws and regulations within your region of operation. If you require further assistance, please contact your dedicated VOSS support person.

Contents

1   What's New

2 VOSS-4-UC

2.1. Deployment Topologies

VOSS-4-UC is deployed either as a single node, 2 unified nodes, or 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:

- WebProxy - load balances incoming HTTP requests across unified nodes.
- Standalone - combines the Application and Database roles for use in a nonclustered environment.
- Unified - similar to the Standalone role Application and Database roles, but clustered with other nodes to provide HA and DR capabilities

The nginx web server is installed on the WebProxy, Standalone, and Unified node, but is configured differently for each role.

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

VOSS-4-UC supports deployment of either the WebProxy node or a DNS load balancer. Here are some considerations in choosing a WebProxy node vs. DNS:

- The Proxy takes load off the Unified node to deliver static content (HTML/JAVA scripts). When using DNS or a third-party load balancer, the Unified node has to process this information.
- DNS does not know the state of the Unified node.
- The WebProxy detects if a Unified node is down or corrupt. In this case, the WebProxy will select the next Unified node in a round robin scheme.

We recommend that you run no more than two Unified nodes and one WebProxy node on a physical server (VMware server). Also recommended is that the disk subsystems be unique for each Unified node.

The following deployment topologies are defined:

- Test or Small Production: a single, Standalone node with Application and Database roles combined. No High Availability/Disaster Recovery (HA/DR) is available.
- Production with Unified Nodes: in a clustered system, comprising:
  - 2, or 4 to 6 Unified nodes (each with combined Application and Database roles)
  - 0 to 4 (maximum 2 if 2 Unified nodes) WebProxy nodes offering load balancing. The WebProxy nodes can be omitted if an external load balancer is available.

2.2.   Virtualized Hardware and Resource Oversubscription

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

VOSS-4-UC 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), hyperthreading is supported.

2.3.   Standalone Deployment

   • No High Availability or Disaster Recovery capability is offered in this topology.
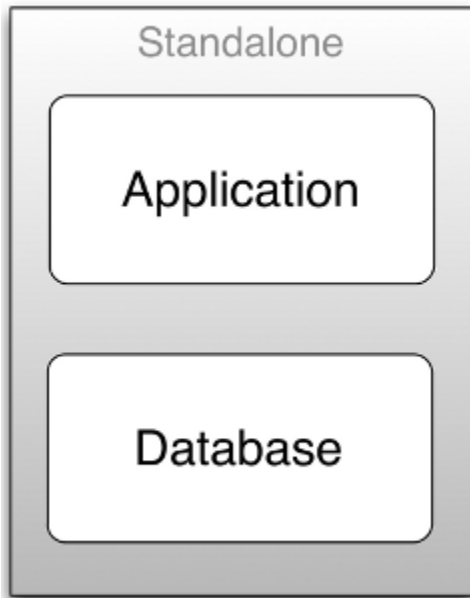
2.4.   Standalone System Hardware Specification

Virtual machine requirements are specified in the table below.

Node type	Quantity	VM	Memory	CPU	Disk	Network
Standalone	1	>= VMware 5.1	16 GB with 16 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	370 GB partitioned: • 20 GB for OS • 50 GB for application: 10 GB for logs, 40GB for our apps • 50 GB for compressed backups • 250 GB for database	1 Gbit/s minimum

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

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

2.5. Two-node Cluster with Unified Nodes

In order to achieve Geo-Redundancy using the Unified nodes, you need the 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.
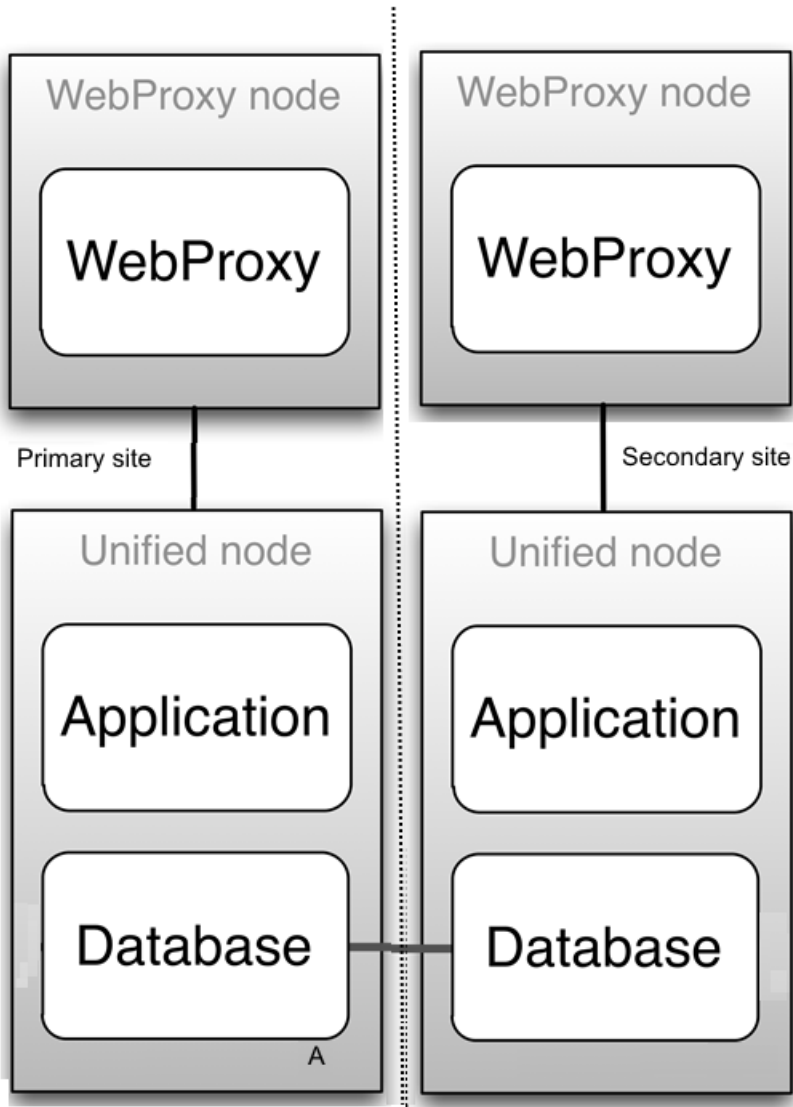
Important:

> 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.

• 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.

The diagram below illustrates the two node cluster:

2.6.   Two-node Cluster Hardware Specification

Virtual machine requirements are specified in the table below.

Node type	Quantity	VM	Memory	CPU	Disk	Network
Unified	= 2	>= VMware 5.1	16 GB with 16 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	370 GB partitioned: • 20 GB for OS • 50 GB for application: 10 GB for logs, 40GB for our apps • 50 GB for compressed backups • 250 GB for database	1 Gbit/s minimum
WebProxy	>= 0	>= VMware 5.1	4 GB with 4 GB reservation	2 vCPU @ 2 GHz with no reservation	70 GB partitioned: • 20 GB for OS • 50 GB for application	1 Gbit/s minimum

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

2.7. Multinode Cluster with Unified Nodes

In order to achieve Geo-Redundancy using the Unified nodes, you need the 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 that ensure an Application role failure is gracefully handled. More may be added if Web Proxy nodes are required in a DMZ.

  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. This is why 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.
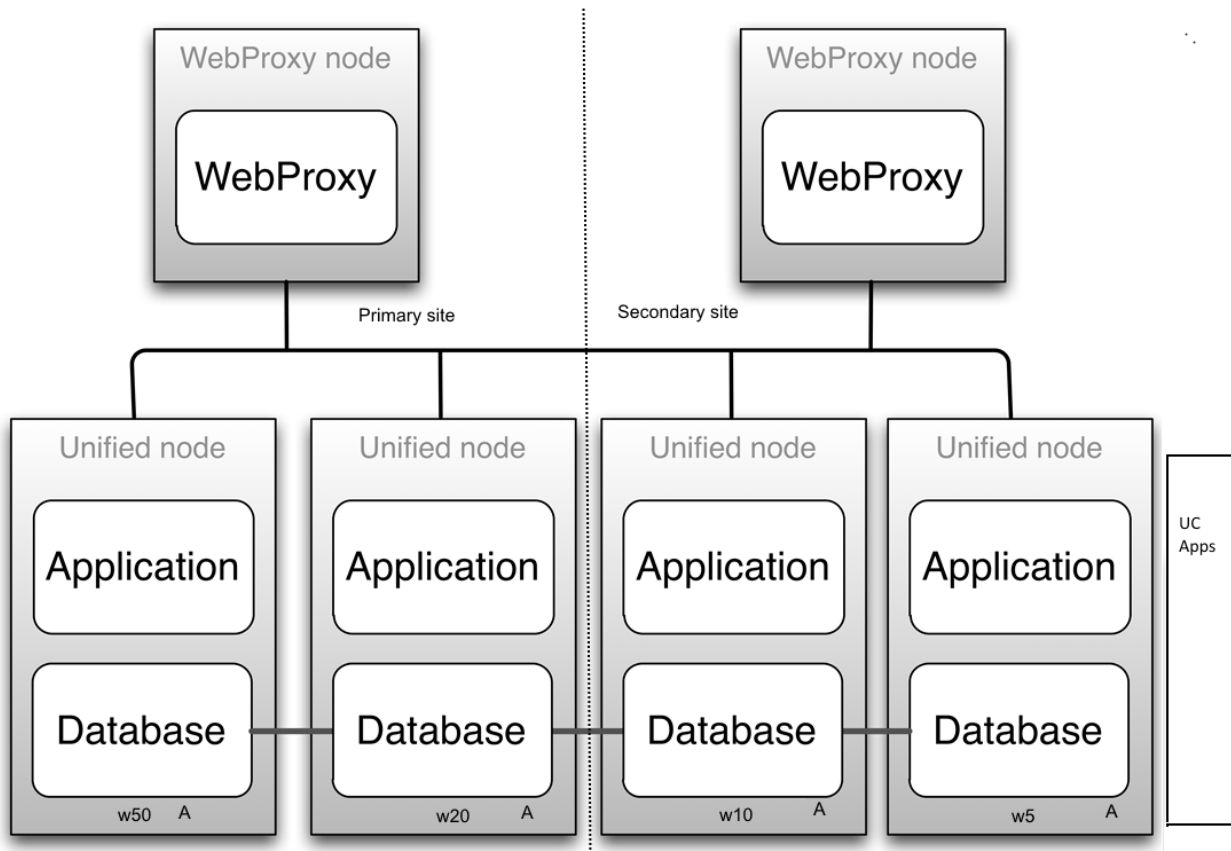
- Web Proxy and Unified nodes can be contained in separate firewalled networks.

- Database synchronization takes places between all Database roles, thereby offering Disaster Recovery and High Availability.

- For 6 unified nodes, all nodes in the cluster are active. For an 8 node cluster (with latency between data centers greater than 10ms) , the 2 nodes in the DR node are passive, in other words, the **voss workers 0** command has been run on the DR nodes.
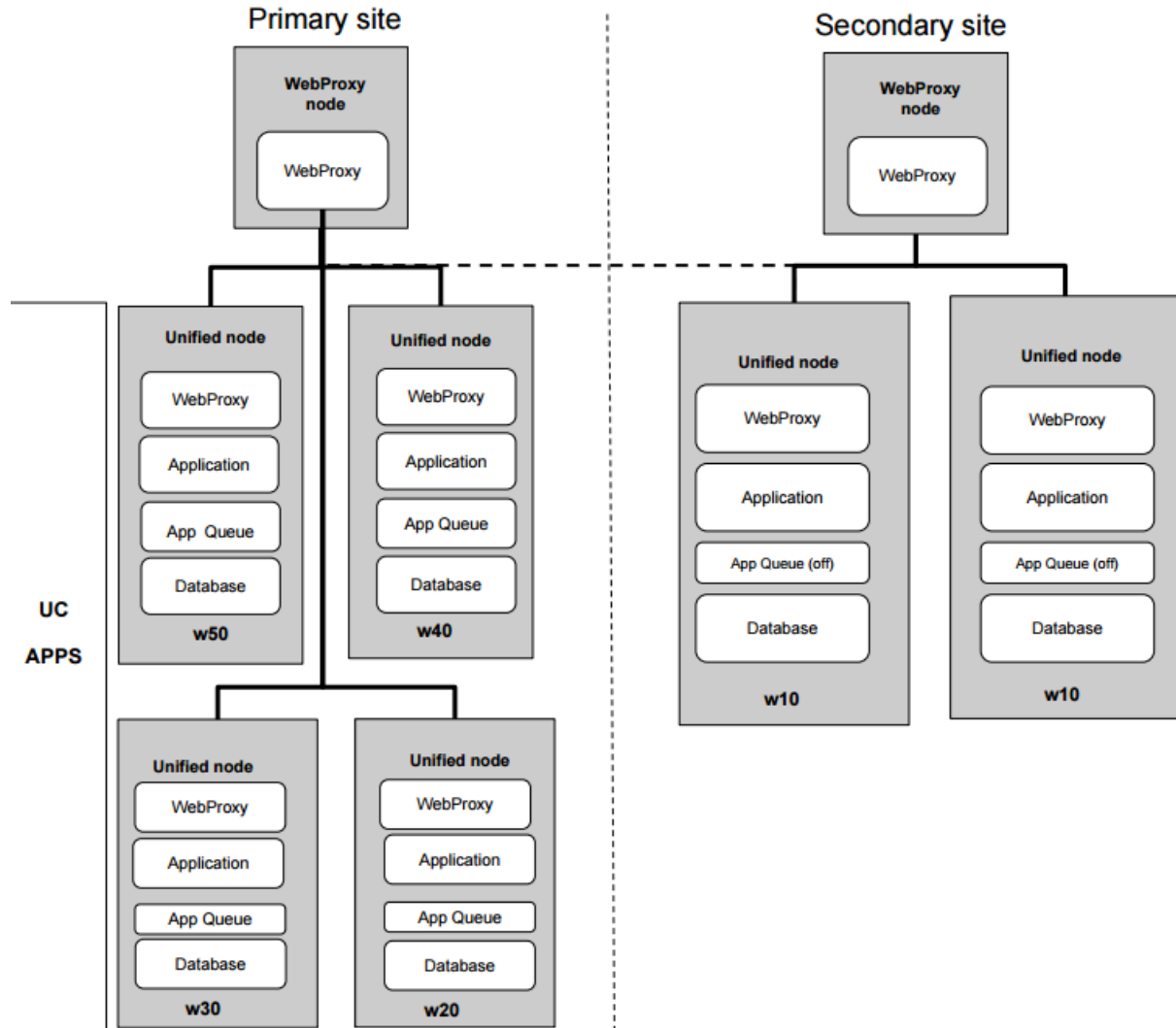
Primary and fall-back Secondary Database servers can be configured manually. Refer to the Platform Guide for further details.
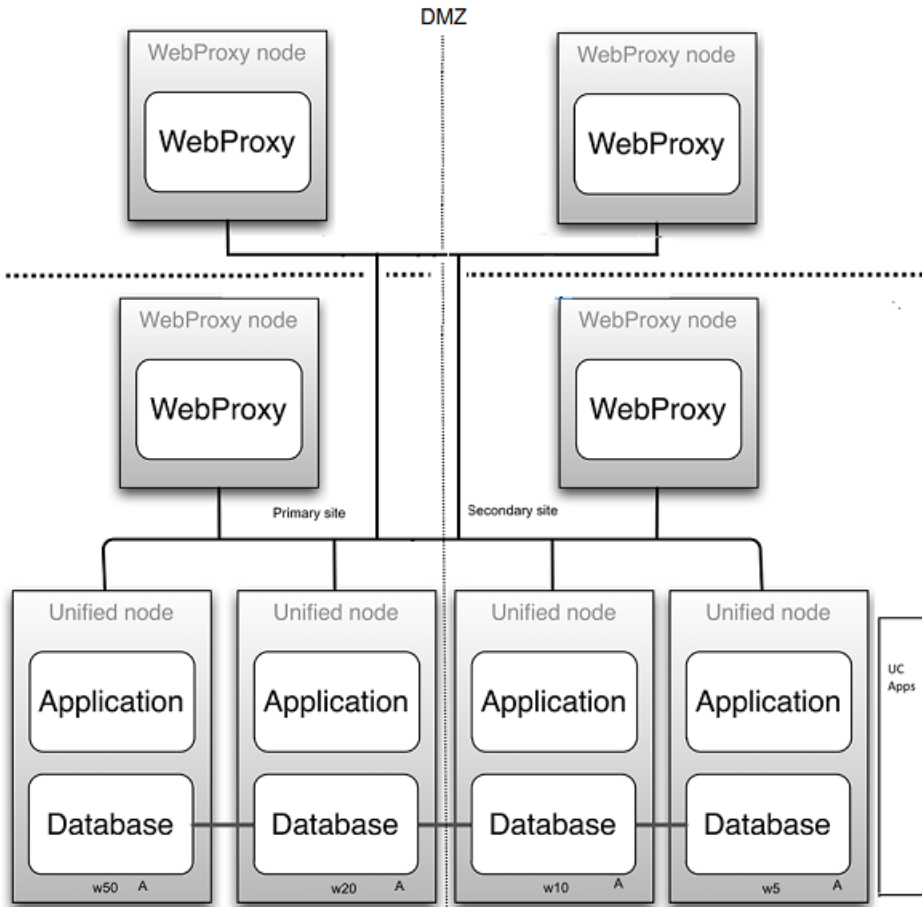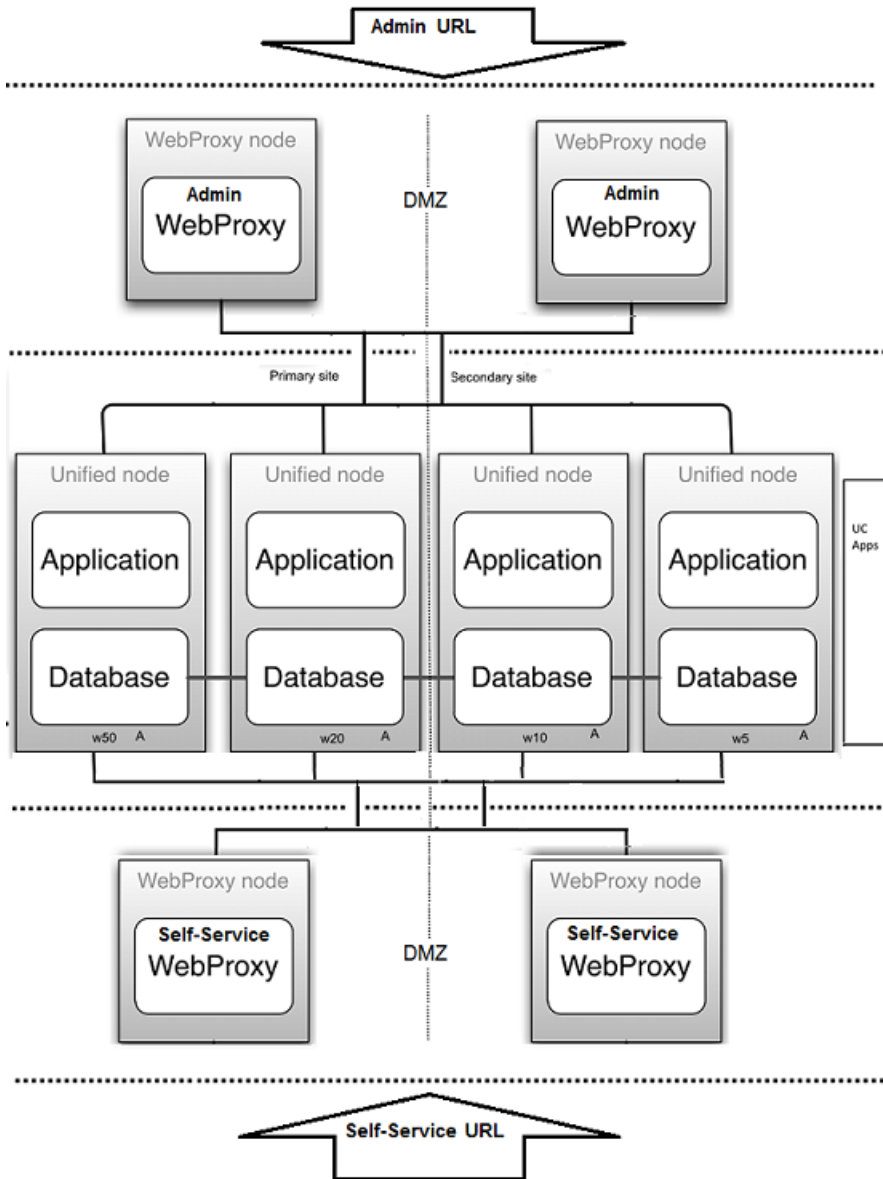
The diagrams in this section illustrate:

- the six node cluster

- the eight node cluster
- 2 Web Proxy nodes in a DMZ
- 4 (2 admin, 2 Self-service) Web Proxy nodes in a DMZ

Primary site

WebProxy node

WebProxy

Unified node

WebProxy

Application

App Queue

Database

w50

Unified node

WebProxy

Application

App Queue

Database

w40

UC
APPS

Unified node

WebProxy

Application

App Queue

Database

w30

Unified node

WebProxy

Application

App Queue

Database

w20

Secondary site

WebProxy node

WebProxy

Unified node

WebProxy

Application

App Queue (off)

Database

w10

Unified node

WebProxy

Application

App Queue (off)

Database

w10

2.8. 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	>= VMware 5.1	16 GB with 16 GB reservation	4 vCPU @ 2 GHz with 4000 MHz reservation	370 GB partitioned: • 20 GB for OS • 50 GB for application: 10 GB for logs, 40GB for our apps • 50 GB for compressed backups • 250 GB for database	1 Gbit/s minimum
WebProxy	2	>= VMware 5.1	4 GB with 4 GB reservation	2 vCPU @ 2 GHz with no reservation	70 GB partitioned: • 20 GB for OS • 50 GB for application	1 Gbit/s minimum

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

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.

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.

2.9. Scale and Performance

This section details the supported configurations and corresponding scale, as well as Geo-Redundancy and Round Trip Time (RTT) requirements.

Configuration	Number of Unified Nodes	Number of Web Proxy Nodes	Supported Scale (# Subscribers)	Geo-Redundancy (Y/N)
Standalone VOSS-4-UC	1	0	50,000	NA
Multi-Node VOSS-4-UC Across Data Centers	4	2	500,000	Yes (Active-Active) If RTT <= 10ms
	6	2	500,000	Yes (Active-Standby)
Multi-Node VOSS-4-UC One Data Center	4	2	500,000	No

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

- Two endpoints (physical phone and soft client) per subscriber.

- Voice Mail

- Extension Mobility (EM)

- Single Number Reach (SNR)

- IM and Presence

The scale limits are not enforced by the system, but exceeding the limits can result in significant performance degradation.

Below follows notes on the difference between Active and Standby nodes, to clarify the Geo-Redundancy options.

- A Unified Node that can process transactions is termed as an Active Unified Node.

- The Standby Unified Node is powered on and running the VOSS-4-UC 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.

Index

V

voss
    voss workers, 6

W

web
    web service, 6