AVAYA
The Power of We™

Avaya Aura® Session Initiation Protocol (SIP) Scalability and Reliability Technical Guide

Table of Contents

A design complement to the Avaya Aura® Scalability and Reliability Overview

Introduction

Avaya Aura® is a service rich, highly scalable solution with multiple levels of redundancy built in to support a variety of communication protocols (traditional analog lines and trunks, digital stations – DCP, ISDN lines and trunks, H.323 stations and trunks, etc.). However solution flexibility and scalability are maximized when the solution components leverage the native protocol of the Avaya Aura core: Session Initiation Protocol or “SIP”.

Avaya Aura® Session Manager is the foundation of Avaya Aura. Based on the IP Multimedia Subsystem (IMS) principles of SIP which establish a practical standards framework, Avaya Aura® Session Manager (SM) is geo-redundant, deployed as always active, and can be implemented with multiple redundancy schemes. This guide describes some of the criteria and operational limits to consider when designing large, highly available solutions.

Executive Summary

The Avaya Aura® Session Manager and Avaya Aura® System Manager solution can currently support 100,000 SIP users (10K users per Session Manager, maximum of 10 Session Managers per Enterprise) and is planned to support 250,000 SIP users in the future. Avaya Aura® Communication Manager (CM) currently scales to 36,000 SIP users and 41,000 non-SIP users or a combination not to exceed 41,000 total users.

Note: The per-instance scale of Communication Manager is not limiting in the Avaya Aura architecture – large enterprise networks are best implemented with multiple Communication Managers, where the Communication Managers serve users closest to them. The current Avaya Aura System Manager solution supports up to 500 Communication Managers in a single Enterprise.

Avaya Aura availability is also well positioned for enterprise customers with requirements similar to those of Service Providers. Since Session Manager and Communication Manager support configurations resilient to more than one failure, more than 6 "9's" of Communication Manager and Session Manager availability can be achieved.

Terminology

For the remainder of the discussion, several terms will be used and are defined as follows:

Term	Description
N	The number of elements (Session Managers) needed for the system to operate. If fewer than "N" Session Managers are operational in the system, the system is considered to be unavailable.
M	The number of redundant Session Manager elements in the system. Up to "M" Session Managers can fail and the system is still considered to be operational however if M+1 Session Managers fail, the system (enterprise core) is considered to be unavailable.
Active-Active	The operational model for the Session Manager. Each Session Manager in the N+M model is actively providing service, and each Session Manager may be replaced by any other Session Manager since all Session Managers share the database necessary for providing service. Active-Active operation is necessary for true N+M redundancy so that failovers do not require initialization and are considered to be "hot".
Geo-Redundant	Geographic redundancy. The implementation flexibility to place elements (Session Managers) randomly throughout a routed network. This is important because it allows Session Managers to handle traffic and provide service on behalf of or in place of failed or unavailable Session Managers wherever they are. Session Managers need not have the same IP address, be on the same subnet, and be geographically nor temporally close to another Session Manager in order to provide redundancy.
SIP Entity	The devices or network elements for which the Avaya Aura Session Manager core provides network routing services. Examples of SIP Entities are Avaya Aura Communication Manager (CM), Avaya Communications Server 1000 (CS 1000), third party PBXs, Gateways, Session Border Controllers (SBC), etc.

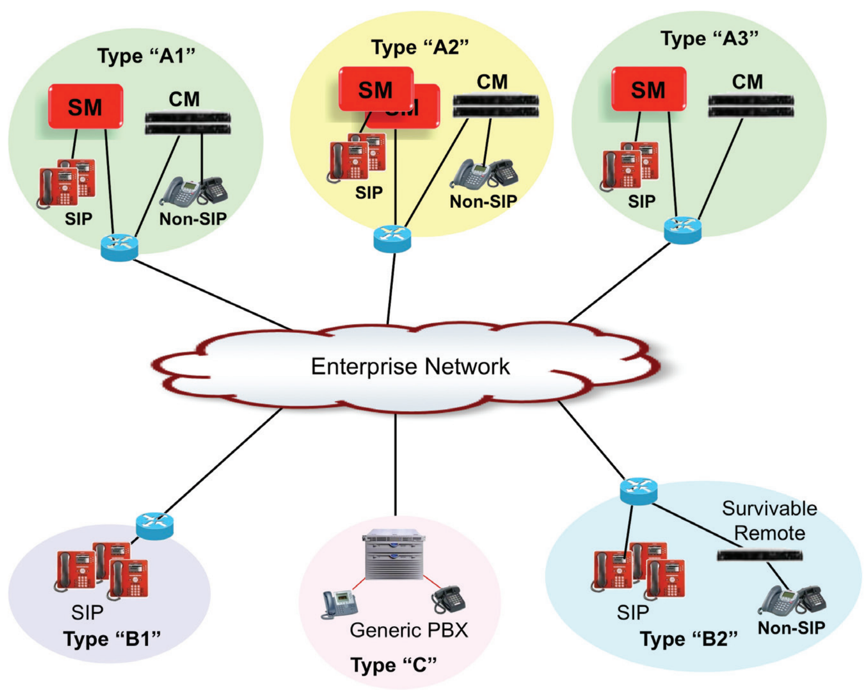

Figure 1: General Enterprise Configuration

Scope

The recommendations in this guide are focused on large SIP deployments employing Avaya Aura components including Session Manager, System Manager, Communication Manager, and endpoints, but also consider non-SIP and non-Avaya existing equipment as possible additions to large multinational worldwide configurations. The general enterprise configuration can be represented as follows:

For the purpose of this design guide, the Enterprise is described as a set of physical locations connected by the enterprise network. We consider three basic types of locations:

- A1–A3 are large locations with or without a “data center” equipment configuration
- B1 and B2 are remote locations that rely on the A1-A3 type locations for normal service
- C is either dominated by non-SIP or non-Avaya type equipment but is still part of the Enterprise and the core SIP design.

Note that any one location type need not contain all of the elements shown in the diagram; some common examples are:

- Location type A1 may not contain a Session Manager
- Location A1 might not have any phones - SIP or non-SIP
- A2 may not actually have any non-SIP phones
- Location B2 may not have any SIP phones.

The design principles offered in the following sections can be applied and are applicable toward real installations that are a combination of two or more locations in the diagram above.

Design Parameters

In order to apply the rules in this guide, the following information is necessary to complete a design:

1. Total Number of Users $\mathbf{U}_T$
2. Number of SIP Users $\mathbf{U}_S$
3. Number of non-SIP Avaya Users $\mathbf{U}_N$
4. Number of non- Avaya PBX Users $\mathbf{U}_P$

5. Average Call Volume per User (in calls per hour) C
6. Number of Large “A” Type Locations (data centers) LA
7. Number of Smaller “B2” Type Survivable Locations with SIP Endpoints LB
8. Desired Redundancy Level

Design Boundaries

The published limitations for SIP designs based on the Avaya Aura 6.2 and set of products including Communication Manager 6.2 and Session Manager 6.2 are:

1. Total Number of Users

   $\mathbf{U_T}$ - there is no design limit, but could be limited by call volume. Please note: $\mathbf{U_T = U_S + U_N + U_P}$

2. Total Number of SIP Users

   $\mathbf{U_S}$ <= 100,000

3. Number of non-SIP Avaya Users

   $\mathbf{U_N}$ <= 100,000

4. Number of non- Avaya PBX Users

   $\mathbf{U_P}$ - there is no design limit, but could be limited by call volume.

5. Number of SIP Users per Session Manager

   The number of SIP users that can be supported on a single Session Manager with no redundancy is 12,000. This includes any overhead associated with dual or triple registrations and there is no “de-rating” required for multiple simultaneous registrations.

6. Communication Manager Non-SIP Users

   41,000 total users can be supported by a single Communication Manager server (or a duplicated pair) based on the Dell or HP Avaya Common Servers. A subset of the 41,000 users can be SIP users as described next.

7. Number of SIP Users per Communication Manager

   Communication Manager supports 36,000 SIP users.

8. Session Manager Performance

   Session Manager can process 350,000 Busy Hour Call Completions (BHCC) per Session Manager in release 6.2. Not all calls are created equally and a detailed Session Manager performance is found in Avaya Aura® SIP Performance Planner (posted on the Avaya Enterprise portal https://enterpriseportal.avaya.com/ptlWeb/bp/products/P0533/WhitePapers).

9. Communication Manager Performance

   Communication Manager SIP call performance is typically not a problem for general business configurations as shown in this guide (call center configurations are not discussed here) and is nearly 120,000 BHCC for each 6.2 Communication Manager. Any Contact Center system design with more than 20,000 BHCC must be carefully reviewed by the ATAC Technical center. (Contact Center performance can vary widely) -

   https://enterpriseportal.avaya.com/ptlWeb/bp/so/CS200662532555952058/C20088137216271038/SN20088138394540 7018/SN20088138394540 7018

10. Administered Communication Managers

System Manager can accommodate 500 main Communication Managers. The System Manager can also accommodate all the Survivable Remotes (SR) and Survivable Core Communication Managers that are associated with the 500 main Communication Managers. Up to 25,000 main Communication Managers may be connected to Session Manager, but only 500 of these Communication Managers can have their users administered via System Manager.

Avaya Aura® Session Manager Scalability

The first exercise in designing a robust scalable solution is a determination of the number of Session Managers required to perform the call processing necessary to meet the customer's requirements. The number of Session Managers is often referred to as "N" in redundancy schemes and will be referred to as such in this document. For this description, the following design parameters are relevant:

- Total Number of Users $\mathbf{U_T}$
- Number of SIP Users $\mathbf{U_S}$
- Average Call Volume per User (in calls per hour) **C**

The total number of Session Managers required (N) for the solution is a function of two principles – call processing performance and user registration. An analysis of both is described below.

SIP User-Based Calculation

$\mathbf{U_S}$	$\mathbf{U_S}$ w Presence	N
1-12,000	1-9,000	1
12,001-24,000	9,001-18,000	2
24,001-36,000	18,001-27,000	3
36,001-48,000	27,001-36,000	4
48,001-60,000	36,001-45,000	5
60,001-72,000	45,001-54,000	6
72,001-84,000	54,001-63,000	7
84,001-96,000	63,001-72,000	8
96,001-100,000	72,001-81,000	9

Session Manager User Scalability

The total number of Session Managers required for SIP users is a simple formula and can be obtained from $\mathbf{U_S}$. The number of Session Managers required is a simple division of the maximum number of users by the maximum number of users supported for a single Session Manager server – 12,000 in R6.2. The Session Manager configuration can support 12,000 SIP users with one, two, or three simultaneous registrations. There is a slight performance impact for SIP users that use Presence and 9000 users with Presence are supported per Session Manager in R6.2. For example, if a customer has 50,000 SIP users with Presence a minimum number of 6 Session Managers would be required. The following table shows the simple calculation for “N” based on the number of SIP users $\mathbf{U_S}$.

Session Manager Performance Scalability

The traffic load for SIP Entities also has an impact on the number of Session Managers required for call processing, although this rarely requires extra Session Managers above what is already planned for the number of data centers in a customer’s network.

The simplified calculation provided here uses the total number of users $\mathbf{U_T}$ and the call rate per user to determine the number of Session Managers (N) and is based on the 350K BHCC capacity of the Session Manager. The formula assumes a call mix shown in the table below.

Please note that a more detailed calculation can be made by referring to the instructions in the Avaya Aura® SIP Performance Planner posted on the Avaya Enterprise Portal at **https://enterpriseportal.avaya.com/ptlWeb/bp/products/P0533/WhitePapers**

Session Manager Performance Call Mix		
Call Mix (% of Total)	Call Topology	
20%	Non-SIP Station	SIP Trunk
10%	Non-SIP Station	Non-SIP Station (Different SIP Entity)
45%	SIP Trunk	SIP Station
10%	SIP Station	SIP Station
10%	SIP Station	Non-SIP Station
5%	Non-SIP Station	Non-SIP Station (Same SIP Entity)

N = (1.725) $\mathbf{U_T}$ **C** / 350,000

A table illustrating examples of the number of Session Managers required for a few popular call rates follows:

Calculation for Minimum Session Manager Quantity Based on Performance			
U_T	N		
	C = 3 Calls/ Hr/ User	**C** = 4 Calls/Hr/ User	**C** = 6 Calls/Hr/ User
50,000	1	1	2
100,000	2	2	3
200,000	3	4	6
300,000	5	6	9
400,000	6	8	Not Supported
500,000	8	Not Supported	Not Supported

Communication Manager Scalability

Communication Managers can be placed anywhere in the customer network, but are normally placed close to the users they serve to conserve roundtrip delay, resources and bandwidth. The number of Communication Managers required for a SIP solution is a simple division of the number of SIP users U_S divided by the total number of users supported on a single Communication Manager – 36,000. For example, a 100,000 SIP user enterprise could be supported by 3 main Communication Managers.

Communication Manager redundancy was delivered in Release 6.0.1 for SIP using Survivable Core (SC) servers. SC installations can be either paired servers or simplex configurations. A high level of availability (greater than 6 "9s") can be obtained with multiple SCs; up to 63 SCs are supported per main Communication Manager.

Branch Scalability

Branch scalability is dependent on the type of branch. 25,000 branches of type B1 and 250 branches of type B2 are supported. The B5800, Advanced Gateway 2330 and Secure Router 2330/4134 are also supported up to a maximum of 25,000.

When branches of type B2 are implemented using the Survivable Remote servers, the number of SIP users should also be designed to the following specifications which have been thoroughly tested to the maximums shown:

Survivable Remote SIP Users	
Server	Max SIP Users
S8300D	700
Dell R610	2000
HP DL360G7	2000
S8510	2000
S8800	2000

Avaya Aura® Session Manager High Availability

After the number of Session Managers minimally required for the design (N) is determined, the placement of these servers and the redundancy scheme can be addressed. Because Session Manager is designed with an active-active shared database architecture, it can flexibly meet a given customer's varied high availability requirements. Three basic Session Manager high availability configurations are considered – **N+M Routing, N+1 Endpoint Support,** and **Redundant Pair**.

The basic formula for calculating availability for any configuration is:

A = **MTBO**/(**MTBO** + **MTTR**), where,

MTBO is Mean Time Between Outage, and indicates the length of time between outages.

MTTR is the Mean Time To Repair (or Recovery), which is a measure of the time to recover.

The key to calculating the availability **A** of any Session Manager system is the underlying availability of the Session Manager single server, the building block of all Session Manager availability models. Session Manager **MTBO** can be approximated using the combination of the availability of the Session Manager hardware server and the Session Manager software availability.

Session Manager Server Hardware Availability

Session Manager runs on several Avaya Aura® Common Server hardware platforms with the latest being the Dell R610 and HP DL360 G7 dual processor quad-core based hardware servers. Both servers are specified because they have excellent redundancy based on:

- Dual redundant hot-swappable power supplies
- Redundant NIC hardware utilizing NIC bonding
- RAID 1 hard drive configuration (2 hard drives – one redundant)
- Assumed externally connected UPS power

The **MTBO** for any of the supported Session Manager server configurations is more than 11 years or approximately 100,000 hours. With a 24 hour replacement strategy, a reasonable worst-case analysis used here is a **MTTR** of 24 hours for the hardware servers.

AHW = 100,000/(100,000 + 24) = 99.97600%

Session Manager Software Availability

The Session Manager software code base is derived directly from the Ubiquity carrier-grade JSR 289 platform designed to provide service under all conditions for the largest service providers in the world. Avaya has added extensive audits, periodic and demand maintenance software processes, along with fault-tolerant code to improve the overall reliability to exceed even the original Ubiquity standards. The **MTBO** for the Session Manager has been verified with thousands of hours of production data on many server hardware platforms and exceeds 1,000 hours.

MTTR for Session Manager software is more complicated as there are numerous recovery times based on the type and extent of the software failure. Software failures range from individual thread recoveries (sub second) to process restarts (one or more seconds), compound software component recoveries (several seconds) to a full Session Manager reset which can take minutes depending on the database size of the customer. For these notes, the worst case of an entire Session Manager reset is assumed and a generous average **MTTR** of 5 minutes will be used. Note that most systems will have far shorter **MTTR** numbers.

ASW = 1,000/(1,000 + .083) = 99.99167%

Combining HW and SW Availability

Assuming independence between HW and SW failures, the model for a combined availability **A** is closely approximated with the Markov serial reliability model:

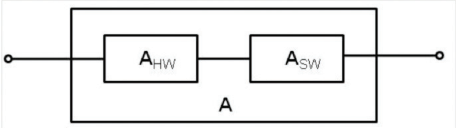

Figure 2: Combined Session Manager Availability

Therefore the combined availability for a single Session Manager server is:

A = ($\mathbf{A_{HW}}$) ($\mathbf{A_{SW}}$) = (99.97600%) (99.99167%) = 99.96767%

Session Manager N+M Routing High Availability

Session Manager can be deployed in an N+M redundant configuration for a supremely reliable core infrastructure that can exceed 6 “9’s” reliability or availability greater than 99.9999%. The description of this industry standard configuration is illustrated below.

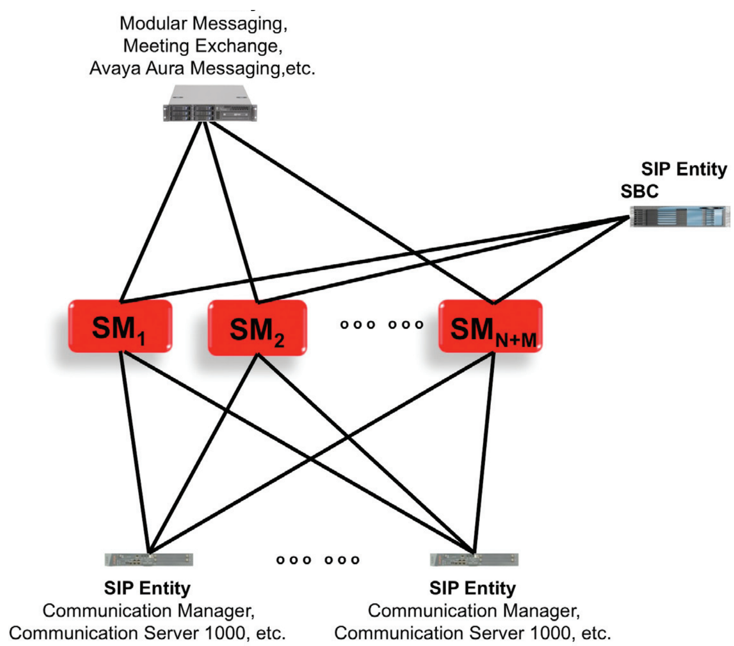

Figure 3: Session Manager N+M Availability

The SIP Entities shown can be implemented with Communication Manager or CS1K systems capable of multiple connections to the Session Manager core as well as the applications and SBCs necessary for proper enterprise connectivity. For the Communication Manager case, the Route Patterns administration page is used to prioritize the multiple connections up to a maximum of 6 links to the Session Manager core. Other advanced administration techniques may also be used to connect to the full maximum of 10 Session Managers, allowing up to 9 of 10 Session Managers to fail while the remaining Session Manager can still provide service.

In an N+M strategy, the system availability $\mathbf{A}_{\mathbf{SYS}}$ for the Session Manager core is approximated by the Markov model:

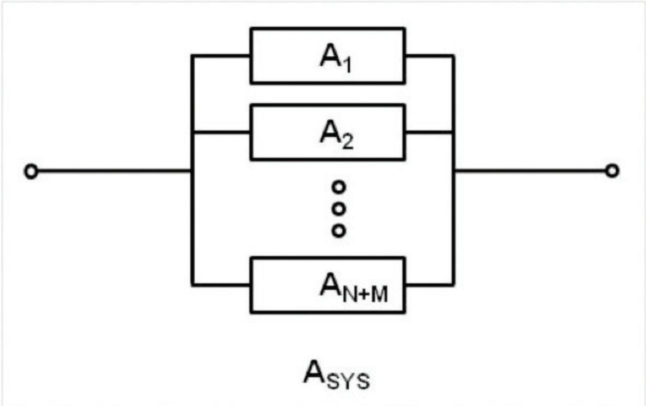

Figure 4: Session Manager N+M Reliability Markov Model

Where the maximum number of Session Manager servers is "N+M", each Session Manager has the same availability **A**, and the total availability for N=1 can be written:

$\mathbf{A}_{\mathbf{SYS}} = [1 - (1 - \mathbf{A})^{M+1}]$

The following table shows the most common availabilities for N=1:

N=1 Session Manager System Availability			
SMs	M	Availability	"9"s
2	1	99.9999%	6+
3	2	99.9999%	6+

A more complicated calculation is necessary for N>1, shown in the following table.

Session Manager Advanced System Availability				
SMs	N	M	Availability	"9"s
3	2	1	99.9999%	6
4		2	99.9999%	6+
5		3	99.9999%	6+
4	3	1	99.9999%	6
5		2	99.9999%	6+
6		3	99.9999%	6+
5	4	1	99.9999%	6
6		2	99.9999%	6+
7		3	99.9999%	6+
6	5	1	99.9999%	5+
7		2	99.9999%	6+
8		3	99.9999%	6+
7	6	1	99.9998%	5+
8		2	99.9999%	6+
9		3	99.9999%	6+
8	7	1	99.9997%	5+
9		2	99.9999%	6+
10		3	99.9999%	6+

Session Manager N+1 Endpoint Support

SIP endpoints in the Avaya Aura R6.2 environment can register simultaneously with three sources: two main Session Managers in the core, and a Survivable Remote. Therefore, an N+1 Session Manager arrangement at the core ensures that a failure of any random Session Manager will not cause any loss of service. Note that a failure of a second Session Manager will only cause a service outage for the endpoints where their primary and secondary Session Manager servers have failed and will only be a subset of the total endpoints in the Enterprise. However to enable a worst-case design, this guide will consider this case a failure of the system even though only a subset of the endpoints are affected (and no inter-SIP Entity routing is affected). This configuration is illustrated as follows:

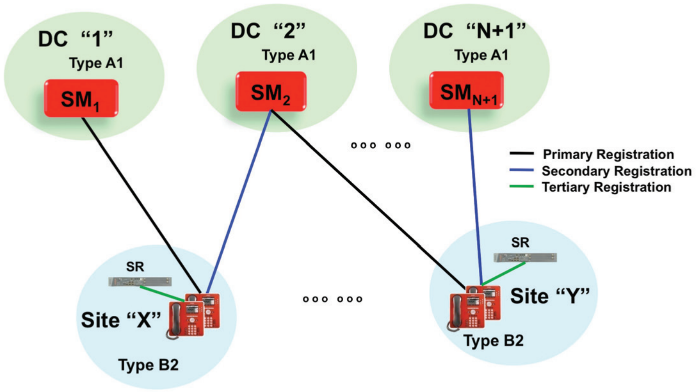

Figure 5: Session Manager N+1 Configuration

The N+1 configuration allows for users to connect to the Session Manager core in user groups or "communities" defined by an ordered triple: (Primary Session Manager, Secondary Session Manager, and Survivable Remote). The Survivable Remote (SR) allows operation when either both Primary and Secondary Session Managers are down or unreachable, or the entire location is isolated from the rest of the Enterprise by a network fault. Since the SR availability is dominated by the WAN availability, the SR availability is ignored for this calculation as it is really a function of WAN availability. However, if the WAN is fully reliable, then the addition of the SR into the calculations below only makes the reliability even greater than is shown below.

The availability of this configuration is determined by the binomial equation for N+M availability where M is (1) and the resulting calculations for the system availability for the values of N are:

Session Manager N+1 System Availability		
N	Availability	"9"s
1	99.9999%	6+
2	99.9999%	6+
3	99.9999%	6
4	99.9998%	5+
5	99.9998%	5+
6	99.9997%	5+
7	99.9997%	5+
8	99.9996%	5+
9	99.9995%	5

The availability of the simplex SR (Survivable Remote) has been empirically determined to be 99.95% (worst case combination of hardware and software availability) for either the S8300/G450 smaller branch configurations or for the larger Aura Common Server Dell 610 or HP DL360G7 servers.

To show the additional availability the SR will provide, the following are the availabilities if the WAN is assumed to be 100% reliable:

Session Manager N+1 System Availability for Survivable Remote w Reliable WAN		
N	Availability	"9"s
1	99.9999%	6+
2	99.9999%	6+
3	99.9999%	6+
4	99.9998%	6+
5	99.9998%	6+
6	99.9997%	6+
7	99.9997%	6+
8	99.9996%	6+
9	99.9995%	6+

Session Manager Data Center Pairs

In this configuration, a pair of redundant data centers is used to provide redundancy for endpoints. Two data centers are employed, with each endpoint simultaneously registering with a primary Session Manager in one data center, and a secondary in another data center. Each region is served by two data centers as shown in the following figure:

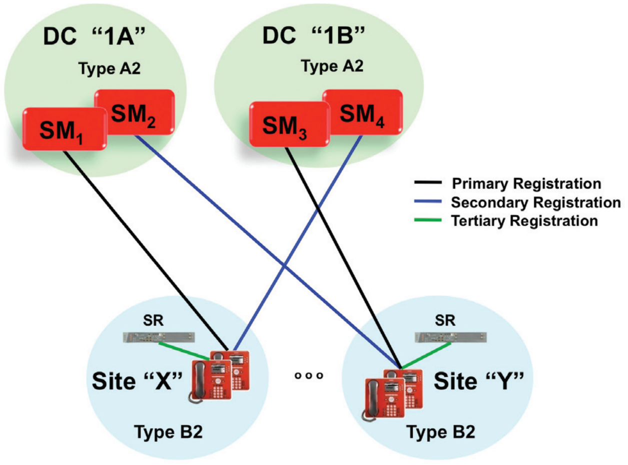

Figure 6: Session Manager Dual Data Center Configuration

The previous figure only shows a single region; there can be another region with duplicated Session Managers, and a third region with unduplicated Session Managers (of location type A1) which reaches the total of 10 Session Managers. Since each region is served by a pair of Session Managers and ignoring the availability of the WAN, the overall system availability due to the Session Manager and SR configuration is closely approximated by:

$$\mathbf{A}_{SYS} = [1 - (1 - \mathbf{A}_{SR})(1 - \mathbf{A})^2]$$

This gives the following calculated availability for a configuration that either type A1 or type A2 Session Manager configurations.

Session Manager Dual Data Center (1+1) System Availability		
SMs	Availability	"9"s
2	99.9991%	5
4	99.99999%	6
6	99.99999%	6+
8	99.99999%	6+
10	99.99999%	6+

Avaya Aura® Communication Manager Availability

When a Session Manager design is sufficiently available and scaled to the number of users and performance required, the availability of the Communication Managers in the design should also be considered. The scalability of main Communication Managers was discussed earlier in Communication Manager Scalability and the correct application of the analysis should result in the number and location of the main server complexes for a customer's design.

Communication Manager Release 6.0, 6.0.1, and 6.2 operate on servers supporting System Platform (SP). System Platform provides the host operating system (CentOS), the Xen Hypervisor, and a virtual machine management console. Communication Manager software provides server redundancy with call and feature preserving fail-over for duplicated S8800, Dell and HP servers. The Avaya S8300D (small branches and main locations) and Dell and HP Servers can further enhance redundancy by serving as Survivable Remotes (SR) and Survivable Cores (SC) within networks. SC and SR servers provide service to users for equipment or network outages as shown below.

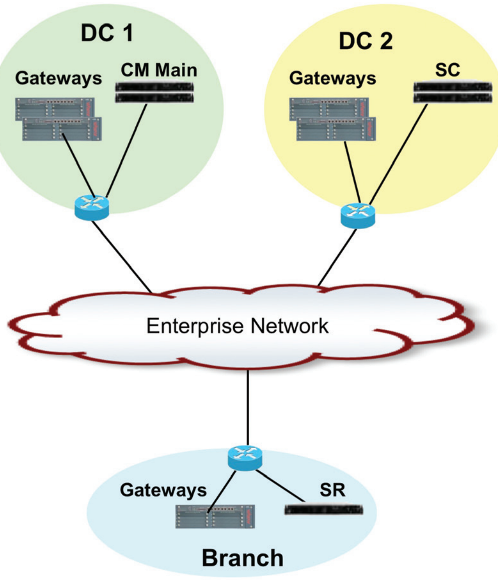

Figure 7: Communication Manager Redundant Configuration

Communication Manager software is delivered on three basic platforms: duplicated Communication Manager server pair, simplex Communication Manager server, and S8300D servers. The duplicated and simplex Communication Manager servers have the following redundancy features:

- RAID 5 Disk arrays
- Redundant hot-swappable power supplies
- ECC Memory
- N+1 Fan Redundancy
- Processor Ethernet NIC Bonding. Processor Ethernet is the NIC interface on the Dell and HP servers and supports high availability with NIC bonding without requiring a port network. Processor Ethernet NICS on the HP and Dell servers can be bonded allowing service even when a NIC or network failure occurs.
- Communication Manager High Availability Platform. The HAP part of Avaya Communication Manager Software incorporates watchdog type monitoring of health and sanity of the applications, and the base operation system, as well as critical environmental conditions and includes extensive automatic software fault recovery.

The G450 Media Gateway is used for high reliability solutions and has the following characteristics:

- Field replaceable supervisor card, DSP VoIP cards, and fan units
- Hot-swappable redundant power supplies
- Media module replacement without power-down

Figure 8 shows an example of a popular networked configuration with duplicated Communication Manager main servers, processor Ethernet with NIC bonding, duplicated SC servers and a dual data center Session Manager design.

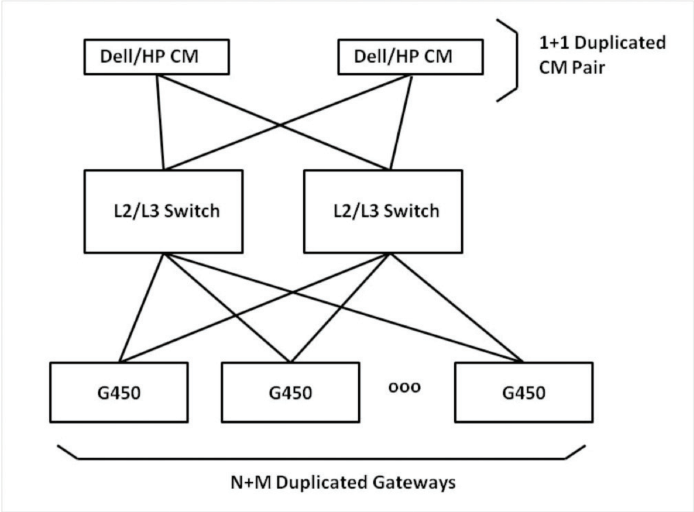

Figure 8: Example Communication Manager Availability

The 1+1 Communication Manager pair availability shown above uses NIC bonding and is a combination of hardware and software Communication Manager availability. The availability of the Communication Manager duplicated pair has been calculated and verified with thousands of hours of operation (with an expected **MTTR** of 24 hours) to be 99.9993% or greater. The individual availability of the G450 gateway has also been determined to be 99.98% or greater. Therefore, the combined reliability of a Communication Manager system is the product of the availability of the servers and the gateways:

$\mathbf{A}_{CM} = (\mathbf{A}_{Server})\ (\mathbf{A}_{GW})$

Since $\mathbf{A}_{\mathbf{Server}}$ is a function of how many SC's are used in multiple data centers, and the $\mathbf{A}_{\mathbf{GW}}$ is a function of the N+M redundant gateway configuration chosen, the following tables outline the total Communication Manager availability and assume a perfectly reliable WAN.

Communication Manager Availability with No SCs and Reliable WAN			
Main GWs "N"	Spare GWs "M"	Availability	"9"s
1	0	99.9870%	3+
	1	99.9994%	5
	2	99.9994%	5
2	0	99.9746%	3+
	1	99.9994%	5
	2	99.9994%	5
3	0	99.9621%	3+
	1	99.9994%	5
	2	99.9994%	5
4	0	99.9497%	3
	1	99.9994%	5
	2	99.9994%	5

Communication Manager Availability with 1 (or more) Simplex SC and Reliable WAN			
Main GWs "N"	Spare GWs "M"	Availability	"9"s
1	0	99.9799%	3+
	1	99.9999%	6+
	2	99.9999%	6+
2	0	99.9600%	3
	1	99.9999%	6+
	2	99.9999%	6+
3	0	99.9400%	3
	1	99.9999%	6+
	2	99.9999%	6+
4	0	99.9200%	3
	1	99.9999%	6
	2	99.9999%	6+

If WAN downtime is considered, the N+M reliability model becomes a function of the WAN reliability. We assume the following:

- The gateway and SC are remoted via WAN
- 24 service time from failure to recovery
- Assume a WAN outage (lasting longer than 45 seconds) of 12 times per year that causes the SR or SC to become active
- Outages less than 45 seconds do not cause a failover
- SC or SR takes over and take approximately 3 minutes to become active
- During the 3 minute switchover time, only 5% (3min/60min) of the users are affected because stable calls are in service

Communication Manager Availability with 1 (or more) Simplex SC with WAN Outages			
Main GWs "N"	Spare GWs "M"	Availability	"9"s
1	0	99.9870	3+
	1	99.9994	5
	2	99.9995	5+
2	0	99.9746	3+
	1	99.9994	5
	2	99.9995	5
3	0	99.9621	3+
	1	99.9994	5
	2	99.9995	5
4	0	99.9497	3
	1	99.9995	5
	2	99.9995	5

Avaya Aura® System Manager Availability

System Manager provides the management interface for provisioning and administering the Avaya Aura solution including Session Manager and Communication Manager. Although not necessary for the processing of communications of any type (voice, video, IM, etc.) System Manager is an important element in the Avaya Aura architecture and its availability is important for the ability to manage changes in the Enterprise such as adding and removing users, changing the dialplan, modifying user profiles, etc.

System Manager is offered in two high availability options which are based on an interim active-standby architecture. Future releases of System Manager will include true geo-redundancy and an active-active architecture; the presently available active-standby capabilities are described following.

System Manager Cold Standby

A System Manager server can be prepared in the Cold standby mode to act as a failover when the main System Manager server fails. The process for restoring System Manager functionality for the cold standby scheme is manual, but can be performed in a few hours should the main System Manager fail, allowing "next day" service for System Manager operations. Note that while the main System Manager is inoperable and before the cold standby is activated, all call processing, routing, and operations continue; Communication Manager, Session Manager, PS etc. do not need a functioning System Manager to perform their normal services.

A cold standby System Manager server is prepared with the same IP address, System Manager software release, and similar hardware as the main System Manager server, but the cold standby is usually powered off. Regular backups of the SGMR database are taken (usually nightly) and stored on a remote node or external storage device. The backups are necessary so that the latest snapshot of the System Manager Database is available to implement the Cold standby procedure.

When the main System Manager fails, the cold standby System Manager is powered on, and the last saved System Manager database is restored to the cold standby System Manager. Administrators continue to access the System Manager as before the failure. Detailed instructions for the failover procedure are described in the System Manager release notes at **http://www.avaya.com/support/**.

System Manager High Availability

System Manager also employs internet protocols to implement a more advanced form of high availability that does not require as much manual intervention as the cold standby described earlier. In this scheme, the active and standby System Manager servers are both powered up and are both administered on the same IP subnet. The active server continuously mirrors all disk data from the active System Manager to the standby System Manager, replicating all changes that occur on the active node to the secondary node. All block changes on the active are written to the standby as part of the same transaction on the active. In case of failure of the active node, System Manager automatically restarts on the standby server with the same IP address it had on the active. To maintain database integrity, as it is restarting on the standby server System Manager aborts any uncommitted transactions.

When a failure to the active System Manager occurs, the standby becomes the active System Manager and resumes operation from the last transaction committed by the active before the failure. System Manager users see a System Manager reboot, but access System Manager as they did before the failure occurred with the same URL and same IP address.

System Manager Manual Geo-Redundancy

A System Manager server can also be prepared in the Cold standby mode to act as a failover when the main System Manager server fails but located in a different location (remote data center for example) from the active System Manager server.

Like the cold standby procedure, the standby System Manager is prepared with the same System Manager version, similar hardware and the same IP address as the active System Manager. However, the standby System Manager is placed in another data center as the only device on an IP subnet that is unroutable from any other location in the Enterprise. The active System Manager is also configured as the only device on its active IP subnet. Backups are taken regularly as with the cold standby procedure.

When a failure to the primary System Manager hardware or network connectivity occurs, a routing change is implemented in the Enterprise for the System Manager to change the routing for the System Manager IP subnet from the active data center to the standby System Manager data center. A script can be used to make the network router change executed in a short amount of time when the failure occurs. The last good System Manager database is then restored to the standby System Manager and administrators can access the System Manager as before the failure.

SIP Scalability Example

As an example of the principles discussed in this guide, a configuration example with the following requirements is shown below.

- New quote for a customer to replace all the existing equipment
- 57,000 total users, of which 50,000 are to be SIP
- Four main locations, with the following configuration:
  - Location A has 3,000 non-SIP users
  - Location B has 3,000 non-SIP and 4,000 SIP users
  - Location C has 12,000 SIP users
  - Location D has 1,000 non-SIP and 4,000 SIP users
- Three main data centers co-located with the main locations
- 500 branches requiring only thin survivability with 40 users each
- 200 branches requiring full survivability with 100 SIP users each
- 5 “9’s” of reliability
- Average busy hour of 4 calls per hour per user

The design will use the dual data center configuration with 8 Session Managers (2 Session Managers in each data center), and 4 Communication Managers with 4 SCs. The 500 thin survivable branches will employ the B5800 server for local thin survivability, while the 200 larger branches will be served with G430 gateways and the S8300D SR. The configuration is shown:

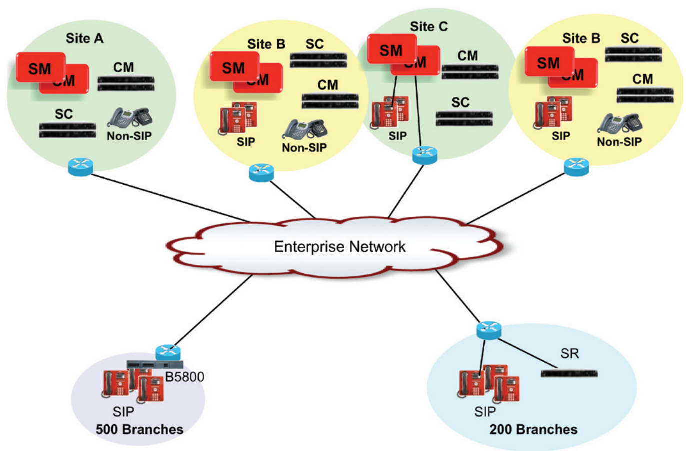

Figure 9: Example System Configuration

Location A and B form a region where each location adds redundancy to the other, and Location C and D form a second region. A simple performance analysis using the general formula yields the minimum number of Session Managers required:

N = (1.725) $\mathbf{U_T}$ **C** / 350,000 = 1.725*57000*4/350,000 = 1

Since less than 12,000 users are required to register to each Session Manager and the call processing traffic requirements are met with just 2 Session Managers, the 8 Session Managers in the design are more than sufficient.

Each Communication Manager has less than 36,000 SIP users, and the redundant SCs in the neighboring regional data centers provide the required availability.

About Avaya

Avaya is a global provider of business collaboration and communications solutions, providing unified communications, contact centers, data solutions and related services to companies of all sizes around the world. For more information please visit **www.avaya.com**.

12/11 • UC4750-02