Cisco Preferred Architecture for Webex Calling

Design Overview

Aug 2022

Contents

What's New in This Guide

Table 1 provides a historical list of updated and new topics added to this guide.

Table 1 Cisco Preferred Architecture for Webex Calling publication history.

Date	Updated or New Topic	Update Details and Location
July 2020	Initial document publication	Initial release
April 2021	Throughout document	Rebranding "Webex app". Minor edits to text and illustrations to correct spelling, grammar, etc. based on feedback.
	Premises-based PSTN	New term used throughout document to refer to PSTN access via on-premises Local Gateway
	Webex Calling Regional Datacenters	New Webex Calling datacenters in Canada (Architectural Overview, Architectural Overview > Webex Calling Regions).
	Cisco Calling Plans	Cisco Calling plans added as additional PSTN access option (Architectural Overview > PSTN Access and On-Premises Interconnect)
	Trunks and Route Groups	New section (Architectural Overview > PSTN Access and On-Premises Interconnect > Trunks and Route Groups). Local Gateway deployment design guidance (Case Study > Local Gateway deployment)
	Dial Plans	New section (Architectural Overview > Dial Plans). Updated description of routing behavior for interconnect with on-premises call control (Architectural Overview > Dial Plans > Interconnect with On-Premises Call Control). Combining premises trunks and Cloud PSTN as new deployment option (Architectural Overview > Dial Plans > Combining Premises Trunk and Cloud PSTN for PSTN Access). Use Webex Calling dial plans to establish routing to on-premises call control (Deployment Aspects > Dial Plan > Abbreviated On-net Dialing) Call routing between Webex Calling and Unified CM based on Webex Calling dial plans (Case Study > Call Routing Considerations > Calls from Webex Calling to Unified CM). Best practices for Webex Calling dial plan configuration (Case Study > Webex Calling dial plan configuration)
	Local Gateway concurrent calls	Design Guidance on Local Gateway concurrent calls scalability and network requirements (Deployment Aspects > Local Gateway Deployment Options).
	Call Detail Records in Analytics	CDR availability in Analytics (Deployment Aspects > Service Assurance > Analytics)
	Combining Cloud Connected PSTN	Combining Cloud Connected PSTN and Premises Based PSTN is now an option (Case Study: Unified CM with Centralized Call Processing and multiple Webex Calling Locations)

Date	Updated or New Topic	Update Details and Location
	and Premises Based PSTN	
	Called party number format for calls from Unified CM to Webex Calling	Clarification of called party number format options for calls from Unified CM to Webex Calling (Case Study > Call Routing Considerations > Calls from Unified CM to Webex Calling)
Oct 2021	ICE media path optimization	Mention optimized media paths (Architectural Overview > Media Flows, Architectural Overview > Webex Calling Regions)
	Tokyo and Osaka location	Updated datacenter map (Architectural Overview, Architectural Overview > Webex Calling Regions)
	Webex rebranding	Throughout the document
	Opus codec bitrate	Changed per call bandwidth from 70 to 40 kbps
	Call routing flow	Description of Webex Calling call routing flow (Architectural Overview > Dial Plans > Call Routing Overview)
	ESN caller ID without location code	Clarification of ESN caller ID format when no location prefix is configured (Architectural Overview > Dial Plans > Allowed transit calls and caller ID selection, Table 5)
	Clarification about +E.164 caller ID from premises to Webex Calling	Caller ID must be in +E.164 format for screening services to work. (Architectural Overview > Dial Plans > Trunk calls to Webex Calling)
	Location's main number required for trunk to work	Location main number must be configured for trunk to work (Architectural Overview > Dial Plans > Interconnect with On-Premises Call Control)
	Configurable dual identity support	Caller ID handling with configurable dual identity support (Architectural Overview > Dial Plans > Trunk calls to Webex Calling, Architectural Overview > Dial Plans > Allowed transit calls and caller ID selection, Table 5)
Aug 2022	Dedicated Instance	Network connectivity, PSTN options, directory integration, Case study.
	Trunk caller id format requirements	Clarification of caller id format (+E.164, ESN, or extension) requirements for calls sent to Webex Calling on trunks (Identity sent to Webex Calling on trunks)
	Calls from trunks to ESN/extension	Clarification of prerequisites for successful routing of calls to extensions or ESNs (Allowed transit calls and caller ID selection)
	Emergency Calling	The Emergency Calling section has been updated and now includes both CCP emergency calling and Enhanced Emergency Calling.

Date	Updated or New Topic	Update Details and Location
	Detailed Call history, Troubleshooting Webex Calling Calls	Detailed Call History, Troubleshooting (Deployment Aspects > Service Assurance > Analytics)
	Singapore location	Updated datacenter map (Architectural Overview > Webex Calling Regions)
	Video Considerations	Video Call devices, codecs, and bandwidth considerations have been included.

Preface

Cisco Preferred Architectures provide recommended deployment models for specific market segments based on common use cases. They incorporate a subset of products from the Cisco Collaboration portfolio that is best suited for the targeted market segment and defined use cases. These deployment models are prescriptive, out-of-the-box, and built to scale with an organization as its business needs change. This prescriptive approach simplifies the integration of multiple system-level components and enables an organization to select the deployment model that best addresses its business needs.

Documentation for Cisco Preferred Architectures

- Cisco Preferred Architecture (PA) design overview guides help customers and sales teams select the appropriate architecture based on an organization's business requirements; understand the products that are used within the architecture; and obtain general design best practices. These guides support sales processes.
- Cisco Validated Design (CVD) guides provide details for deploying components within the Cisco Preferred Architectures. These guides support planning, deployment, and implementation (PDI).
- Cisco Collaboration Solution Reference Network Design (SRND) guide provides detailed design options for Cisco Collaboration. This guide should be referenced when design requirements are outside the scope of Cisco Preferred Architectures.

Use Case

This Cisco Preferred Architecture (PA) for Webex Calling document describes how organizations can embrace growth without necessarily replacing their entire calling infrastructure. Additionally, the use case details call routing considerations for:

- **Calls between Webex Calling and Dedicated Instance**
- **Calls between Webex Calling and Unified CM on-premises**
- **Class of Service (CoS)**
- **Dial Plan Integration**

Information about Cisco Collaboration Technologies and additional use cases is available on Cisco.com.

About This Guide

This *Cisco Preferred Architecture for Webex Calling* provides architectural guidance on Webex Calling. Specifically covered, are deployments using Webex Calling as cloud-based call control in an isolated deployment or combined with an on-premises collaboration deployment as described in the enterprise PA documents.

Unless explicitly mentioned otherwise, the term "Webex Calling" in the context of this document always refers to Webex Calling multi-tenant. Design guidance for integrations of Webex Calling multi-tenant with Unified CM also applies to integrations between multi-tenant and Dedicated Instance.

For specificities that apply only to Dedicated Instance and not to Unified CM on-premises, a new section has been added.

Readers of this guide should have a general knowledge of Cisco Voice, Video, and Collaboration products and a basic understanding of how to deploy these products.

This guide simplifies the design and sales process by:

- Recommending products in the Cisco Collaboration portfolio that are built for the enterprise and that provide appropriate feature sets for this market
- Detailing a collaboration architecture and identifying general best practices for deploying in enterprise organizations

For detailed information about configuring, deploying, and implementing this architecture, consult the related CVD documents on the Design Zone for Collaboration.

Introduction

Webex Calling Solution Overview

Enterprise-level decision makers recognize that installing, securing, and maintaining an on-premises private branch exchange (PBX) is both complex and expensive. As vendors develop additional features and functionality, privately deployed PBX systems tend to become outdated and less secure.

Prior to the introduction of Webex Calling, cloud-based PBX solutions lacked features, functionality, performance, and security, and were not able to replace an enterprise PBX or PBX network.

Webex Calling is part of an integrated, intelligent, and modular team collaboration suite. It provides enterprise-grade PBX features, functionality, and performance previously only possible with an on-premises PBX network. Licensing is subscription-based and managed with the Cisco Collaboration Flex Plan. The solution integrates with Webex, specifically Webex devices, and optionally with Webex Meetings and Webex Contact Center.

Webex Calling is deployed as a cloud-only solution, or if you require a mixed network of both cloud and on-premises PBXs, it is deployed as part of a hybrid cloud.

Webex Calling data centers are globally distributed and geo-redundant. Cisco Value-Added Reseller (VAR) channel partners distribute Webex Calling.

Dedicated Instance: The Dedicated Instance option is an entitlement within Webex Calling that provides a Cisco Unified Communications Manager based stack of applications, in a private cloud, dedicated to a single customer.

In addition to enterprise-grade PBX features, Webex Calling also provides:

- Webex Calling group features, including unlimited subscriptions of auto-attendants, hunt groups, and call queues
- Integrated calling from within the Webex App for soft client use with mid-call features, or alternatively with control of a user's desk phone
- Webex App with messaging, screen sharing, and audio and video conferencing
- The option to add Webex Meetings for advanced meeting experiences including meeting room recording, meeting room locking, remote dial-in access over PSTN, and supporting up to 1000 meeting participants
- The ability to deploy all models of Cisco Multiplatform Phones (MPP)
  The list of supported devices is available online:
  https://help.webex.com/en-US/article/qkwt4j/Supported-Devices-for-Webex-Calling
- Access to public switched telephone networks (PSTN)

Supported Devices

Webex Calling provides a variety of user interfaces that can be selected depending on the requirements of the organization or end users. Webex Calling supports a wide range of Cisco MPP Series IP Phones (6800, 7800, and 8800 Series), Webex Devices and a variety of third-party devices. The administrator selects the default calling application for the organization, however individual users may be configured by the administrator to use other applications as needed. Software options include Webex App with messaging, screen and file sharing, with complete meetings capabilities. Further information is available at https://help.webex.com/en-us/article/qkwt4j

The Webex App supports commonly used mid-call features, rich presence with MPP and a single line for each user. The Webex App also allows control of the user's Cisco MPP phone. Cisco's line of headsets for end users is supported both when connected to the MPP devices or to the user's workstation or mobile phone. In addition to these Cisco software options, Webex Calling users can be set up to use Webex Calling services from within Microsoft Teams. More information on this integration is available at https://help.webex.com/en-us/article/ngmx08cb.

Video Support

Webex Calling supports video calls between the following endpoints and app:

- Video capable MPP phones (8845, 8865, 8875)
- Webex Devices
- Webex App (desktop and mobile) when configured with Webex Calling for users

Users can take advantage of a speed dial with a SIP address to call into Webex meetings with a properly configured soft key on the video-capable phone. (e.g., user's Personal Meeting Room). For Webex devices in shared mode, you can register them to the cloud and then add Webex Calling PSTN service to the Workspace.

Architectural Overview

Webex Calling is a cloud-based enterprise calling solution hosted in Webex datacenters in multiple geographically distributed locations.

Figure 1 Globally Distributed Datacenter

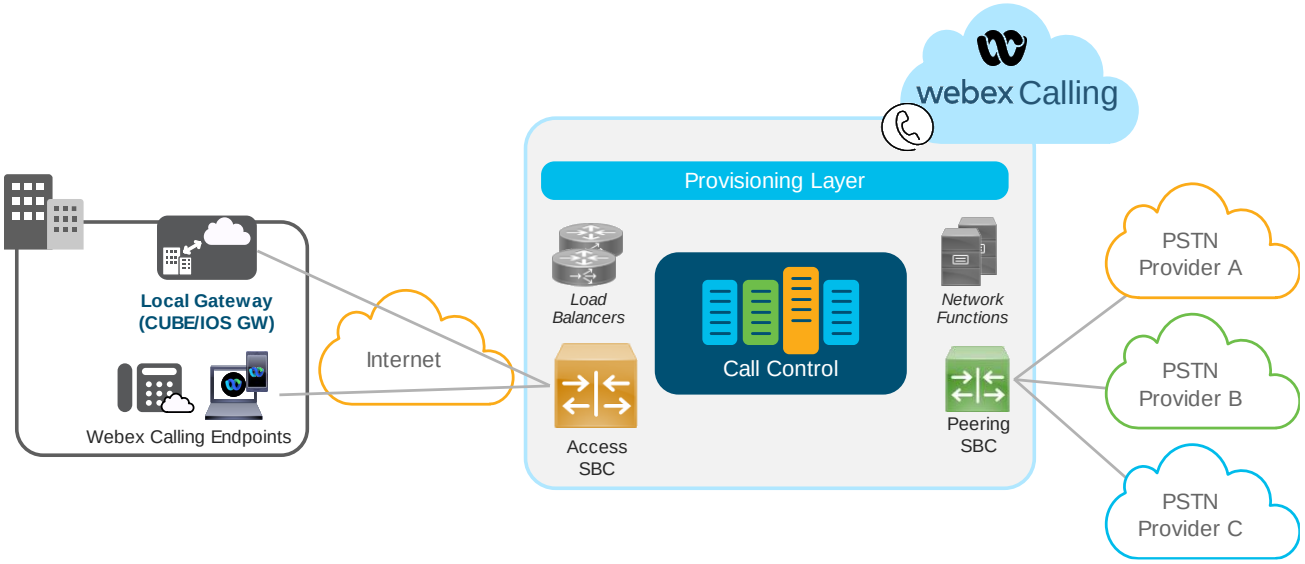
Webex Calling is available globally and is delivered from redundant data centers in six regions: US (Dallas, Chicago), Canada (Vancouver, Toronto), Europe (Frankfurt, Amsterdam), UK (redundant data centers in London), Australia (Melbourne, Sydney), and Japan (Tokyo, Osaka). The data centers in New York and Singapore provide media services to optimize media round trip times. The Singapore data center specifically is used to optimize media round trip times for Webex Calling customers in Asian countries where the round trip times to either the Australia or Japan region might be suboptimal. The datacenters are interconnected by a multi-gigabit and fully redundant backbone.

Webex Calling uses SIP for signaling and SRTP for media. Dynamic NAT can be used for IP Phone and Local Gateway IP addresses. Phones and Local Gateways initiate a TLS connection to Webex Calling and are authenticated by Webex Calling. Webex Calling continues to use the same connection to send traffic back to the phone or Local Gateway, thus providing firewall traversal.

Webex Calling Datacenters

Each Webex Calling datacenter hosts call routing functions and provides provisioning interface access to Webex Calling.

Figure 2 Webex Calling's Functional Elements



The datacenters also host the access and peering Session Border Controllers (SBCs). The access SBCs terminate all customer-facing SIP connections from Local Gateways, endpoints and soft clients and the peering SBCs terminate the SIP peering connections to SIP service providers.
Load balancers and other network functions are required to build a scalable, redundant datacenter architecture are also part of each Webex Calling datacenter.

Access connectivity options

Reliable Network connectivity with sufficient bandwidth is a base requirement to ensure the best possible user experience end-to-end for all voice and video capable endpoints, clients, and applications using Webex Calling..

Customers and partners have access connectivity options beyond Over-the-top (OTT) Internet that can optimize the connection to Webex Calling and these include Webex Edge Connect or Private Network Connect.

Over-the-top TLS Connections

Webex Calling endpoints on the customer's network use the public Internet as the access network to connect to the Webex Calling datacenters and establish over-the-top TLS connections.

Webex Edge Connect

Webex Edge Connect is a solution that peers your Webex meetings and Webex Calling traffic with an Equinix Cloud Exchange (ECX) location. This peering improves the calling and in-meeting user experience by providing guaranteed bandwidth and quality of service (QoS), which minimizes network latency, packet loss, and jitter. For design details please refer to the Webex Edge Connect Preferred architecture
https://www.cisco.com/c/dam/en/us/td/docs/solutions/PA/EdgeConnect/PA_Edge_Connect_Design.pdf

Private Network Connect

The Private Network Connect (PNC) solution allows Webex Calling customers to extend their private network to the cloud. This is done through either a partner cloud or direct connects to ensure high quality of service and low latency for voice calls. For Private Network Connect design guidelines, please refer to the Private Network Connect for Webex Calling Preferred Architecture https://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Collaboration/cloud/PA-PNC.pdf

Figure 3 Over-the-top Connections between Webex Calling Endpoints and Datacenters

PSTN Access and On-Premises Interconnect

Webex Calling can access the PSTN in three ways: Cisco Calling Plans, Cloud Connected PSTN, and Premises-based PSTN:

- Cisco Calling Plans (Cisco PSTN)

  Cisco PSTN enables partners to sell Cisco provided PSTN options to their customers, simplifying the overall purchasing experience of a complete collaboration solution. With Cisco PSTN, number ordering and initiating a port ordering is available from within Control Hub.

- Cloud Connected PSTN (CCP)

  Cisco has set up shared SIP integration with a few Cloud Connected PSTN providers. Webex Calling customers contract directly with a Cloud Connected PSTN provider of their choice and then, in Webex Control Hub, select their Cloud Connected PSTN provider to route calls to the PSTN.

- Premises-based PSTN

  The Local Gateway function running on a Cisco voice gateway or Cisco Unified Border Element (CUBE) leverages existing enterprise PSTN connections to route calls between Webex Calling and the PSTN, including on-premises resources. The Local Gateway function is commonly deployed on the customer's premises but can also be hosted by a partner. The Local Gateway registers with Webex Calling and handles all calls between the PSTN and Webex Calling. Premises-based PSTN requires that a trunk or a route group with multiple trunks is selected as the PSTN choice in Webex Control Hub. Each trunk represents a connection to a Local Gateway.

Figure 4 CCP and Local Gateway PSTN Access

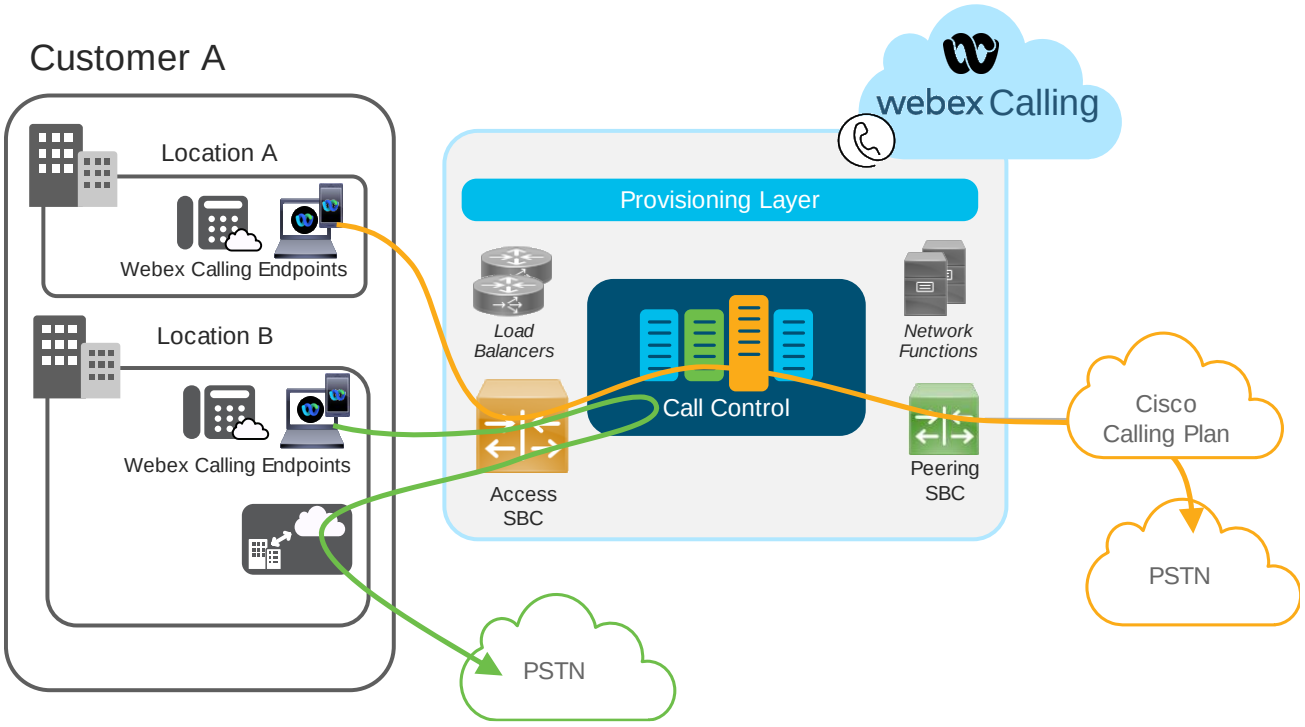


Figure 4 shows that the selection of PSTN type (Cisco Calling Plan, Cloud Connected PSTN, or Premises-based PSTN) can be configured per location. While devices in location A use Cisco Calling Plan for PSTN access, the devices in location B use a trunk to a Local Gateway. Cloud Connected PSTN is not shown in the picture.

In addition to providing PSTN access, a Local Gateway can also connect Webex Calling with an on-premises call control instance. To accomplish this, the Local Gateway must be connected to the on-premises call control instance via SIP.

Calls between different Webex Calling customers are always routed through the PSTN via the configured PSTN choice (Cisco Calling Plan, Cloud Connected PSTN, or Premises-based PSTN) to meet legal requirements.

Multiple PSTN Providers

A Webex Calling customer is not limited to using only a single Cloud Connected PSTN provider or Cisco Calling Plan for PSTN access. Unique Cloud Connected PSTN providers can be selected for different Webex Calling locations to include geographical presence of the provider in case of multi-national deployments and tariff structure, for example.

Figure 5 Multiple PSTN Providers

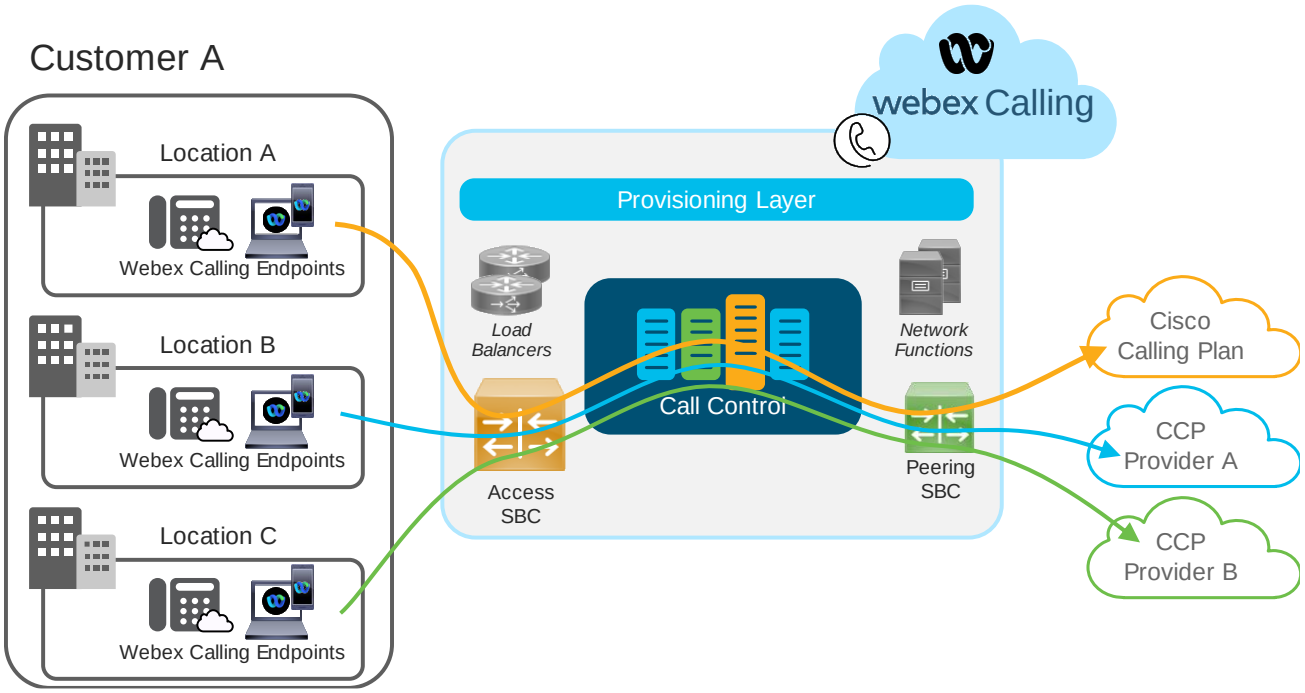


Figure 5 shows an example where each location uses a different PSTN choice. While PSTN services for location A are provided using Cisco Calling Plan, two different Cloud Connected PSTN providers are used for locations B and C. The PSTN choice used for a given call depends on the source (location) of the Webex Calling call. No other attributes related to the call affect the PSTN selection, not even the called address.

Trunks and Route Groups

Trunks connect Webex Calling with Local Gateways or to Dedicated Instance, if included in the solution. Each trunk in Control Hub represents a trunk to one Local Gateway instance configured on a CUBE (or a Cisco voice gateway). Each CUBE can be located either within the customer's premises or in a partner datacenter. Each trunk must be assigned to a location. Multiple trunks can be grouped together in a route group for redundancy or to provide more capacity. Each trunk can belong to multiple route groups. Route groups and trunks can be used as destination for PSTN calls or premises calls.

Using route groups as a routing choice is preferred to using individual trunks even if a route group at the start only contains a single trunk. Using route groups everywhere allows to add capacity or redundancy later simply by adding additional trunks to the route group without having to update all places where that specific route is used.

The dial plan section provides more details about the Webex Calling routing choices.

Figure 6 Route Group for Scale and Redundancy

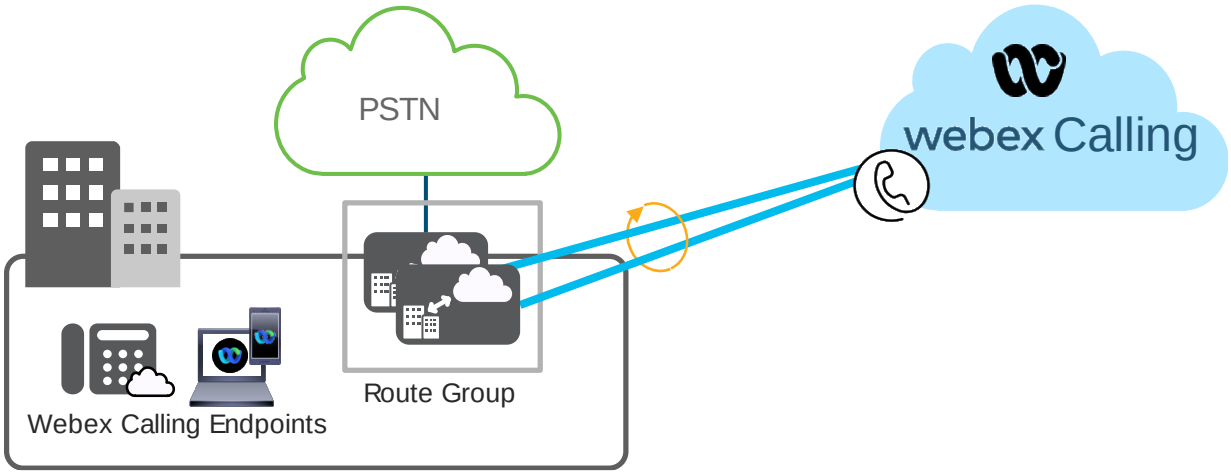


Figure 6 shows an example where trunks to two Local Gateways in the same location are combined into one route group configured as premises-based PSTN choice for that location.

A priority is assigned for each trunk within a route group. When calls are sent to route groups these calls are randomly distributed among all trunks with the highest priority (1 = highest priority). If a call cannot be sent to a trunk, then rerouting occurs and a different trunk is randomly selected from the group of trunks with the highest priority. If rerouting occurs for all trunks with the highest priority, then trunk selection continues with the trunks of the next lower priority level until either the call is routed successfully or all trunks within the route group have been tried without success.

Rerouting is triggered by:

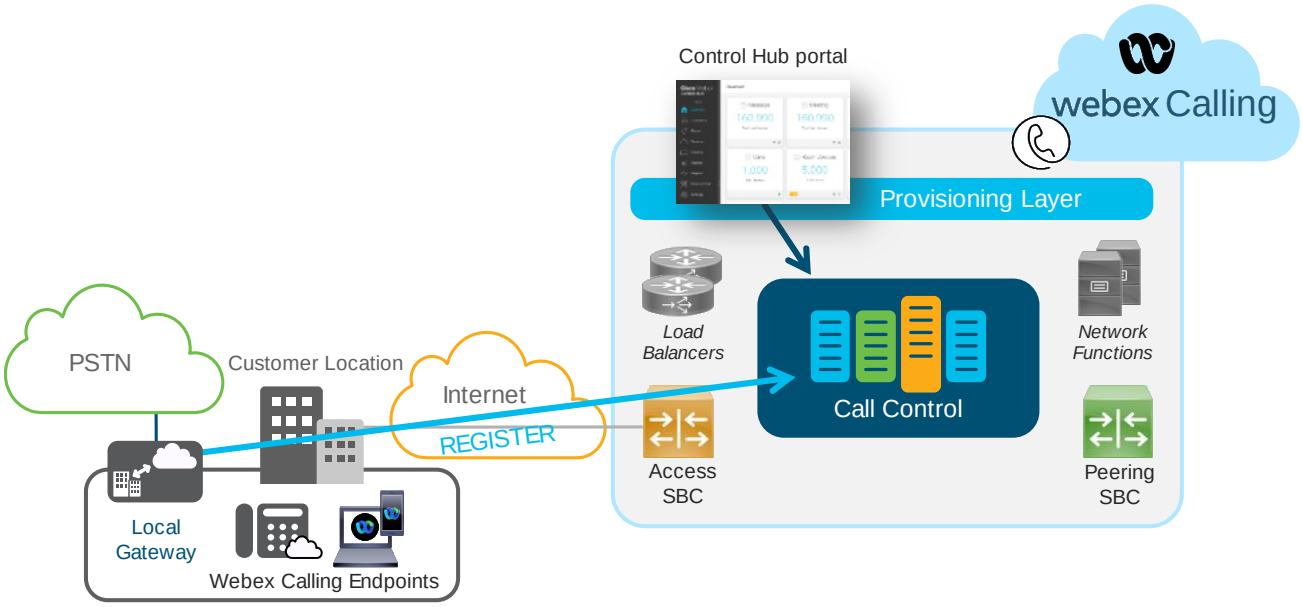
- SIP responses 401, 470, 480, and 606 if there has been no prior 18x response
- All other 4xx SIP responses except for 403, 404, 410, 413, 484 and 486
- All 5xx SIP responses
- A SIP timeout (12 seconds)

A maximum of 100 trunks is allowed per location, a route group can have a maximum of 10 trunks, and a maximum of 10,000 route groups can be configured for a Webex Calling customer.

Local Gateway Registration

Local Gateways use authenticated and registered SIP trunks for the connection to Webex Calling.

Figure 7 Local Gateway Registration



The Webex Control Hub, as part of the provisioning process of a trunk, provides connection parameters and digest credentials for SIP authentication.

These SIP/TLS (SIPS) connections are only initiated from the customer's network (from endpoints and Local Gateways) to Webex Calling. No inbound connections need to be configured on the enterprise firewall because SIPS connectivity between the enterprise and Webex Calling is only outbound.

This deployment is referred to as "Registration-based trunk". Other options, such as "Certificate-based trunk", are not covered in this document.

Dedicated or Co-Resident Local Gateway

A Local Gateway connects to the PSTN either directly by terminating a PSTN trunk (TDM or IP) on the same box or by connecting to an existing PSTN gateway via a SIP trunk.

Figure 8 Local Gateway Deployment Options

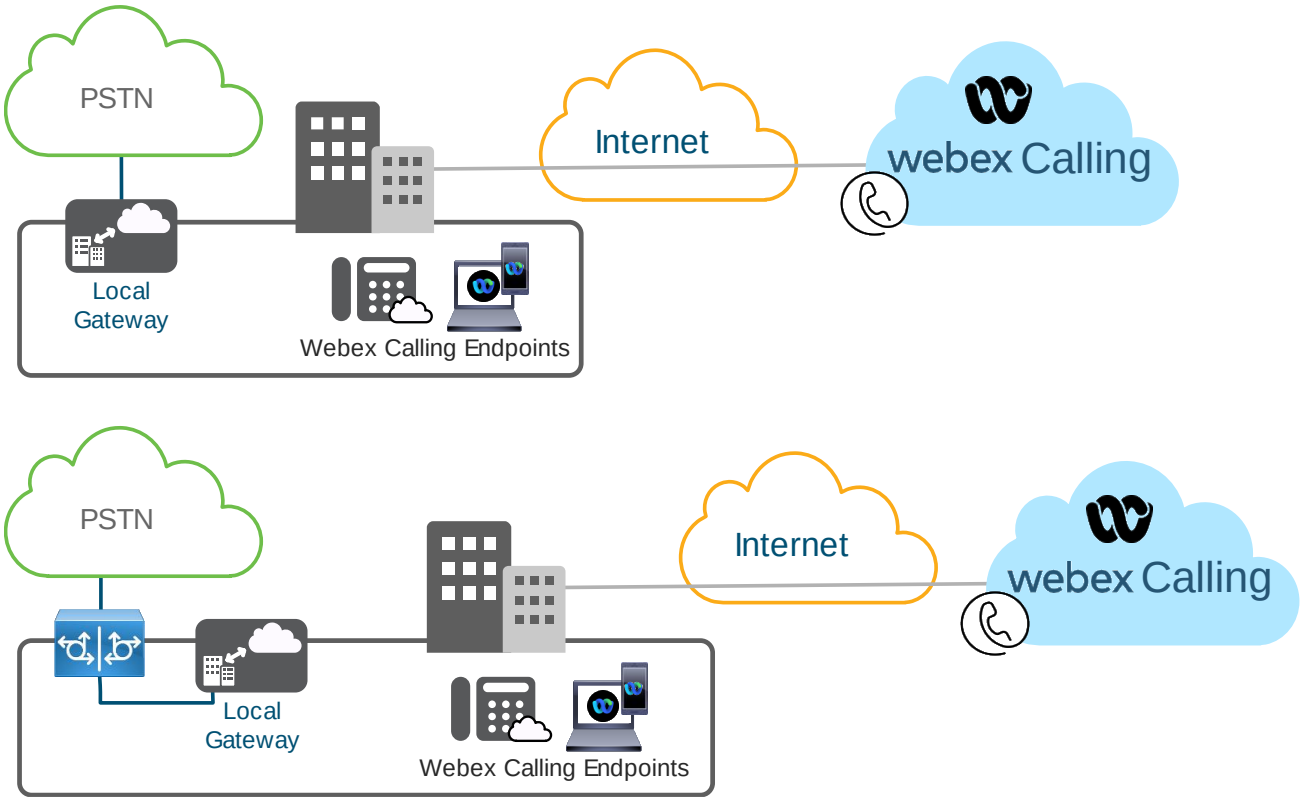


Figure 8 shows both options. While combining PSTN access and the connection to Webex Calling requires less hardware to be installed and maintained on the customer's network, implementing both functions on separate devices can be a preferred option in cases where an existing PSTN gateway continues to be used after migration of a site to Webex Calling.

Partner Hosted Local Gateway

Instead of deploying individual Local Gateways on each customer's network a partner can also host a customer's Local Gateway in their own datacenter.

Figure 9 Hosted Local Gateways

Figure 9 shows that the partner in this example does not need to deploy separate Local Gateways in the partner data center. The dial peer-based call routing configurations of each individual customer can be combined on a single CUBE. Separation of traffic between customers is achieved by proper dial-peer routing and voice class tenant configurations. The Local Gateway configuration guide available at https://help.webex.com/en-us/jr1i3r/ describes how to configure dial-peer matching. This allows you to deterministically on the Local Gateway, map calls received from Webex Calling to customer specific PSTN trunks, and vice versa. There is no need to deploy multiple virtual routing functions (VRFs): there is no need for direct IP connectivity between the partner's data center and customers' networks because both signaling, and media are anchored on the access SBCs of Webex Calling in the cloud. This deployment model enables the partner to more efficiently deploy, maintain, and operate Local Gateways for various customers.

No data connectivity is needed between the partner's datacenter and the customer networks for this deployment model. Individual per-customer registrations with Webex Calling ensure that different customers' calls can easily be identified on the combined Local Gateway to avoid calls leaking between customers.

IP Connectivity between Local Gateway and Customer's Network

Voice media streams need IP connectivity between the Local Gateway and the on-premises call control system and related endpoints.

A Local Gateway connecting Webex Calling with an existing on-premises call control system (see next section) requires IP connectivity between the Local Gateway and the customer network to enable voice media streams between the Local Gateway and the on-premises call control system and/or endpoints controlled by the on-premises call control system.

When a Local Gateway is deployed in a partner's datacenter, the customer's network must be extended to the partner's datacenter. If the Local Gateway functionality is implemented on a shared platform, network separation is achieved with network virtualization mechanisms, for example, VRF configuration. The additional overhead and maintenance of this complex Local Gateway configuration combined with network security reservations may negate any benefit of sharing a single Local Gateway platform between multiple Webex Calling customers. Deploying a dedicated on-premises Local Gateway may be a more efficient deployment option if connectivity to an on-premises call control system is needed.

Local Gateway Call Setup

Figure 10 SIP/TLS Connection Setup Flow

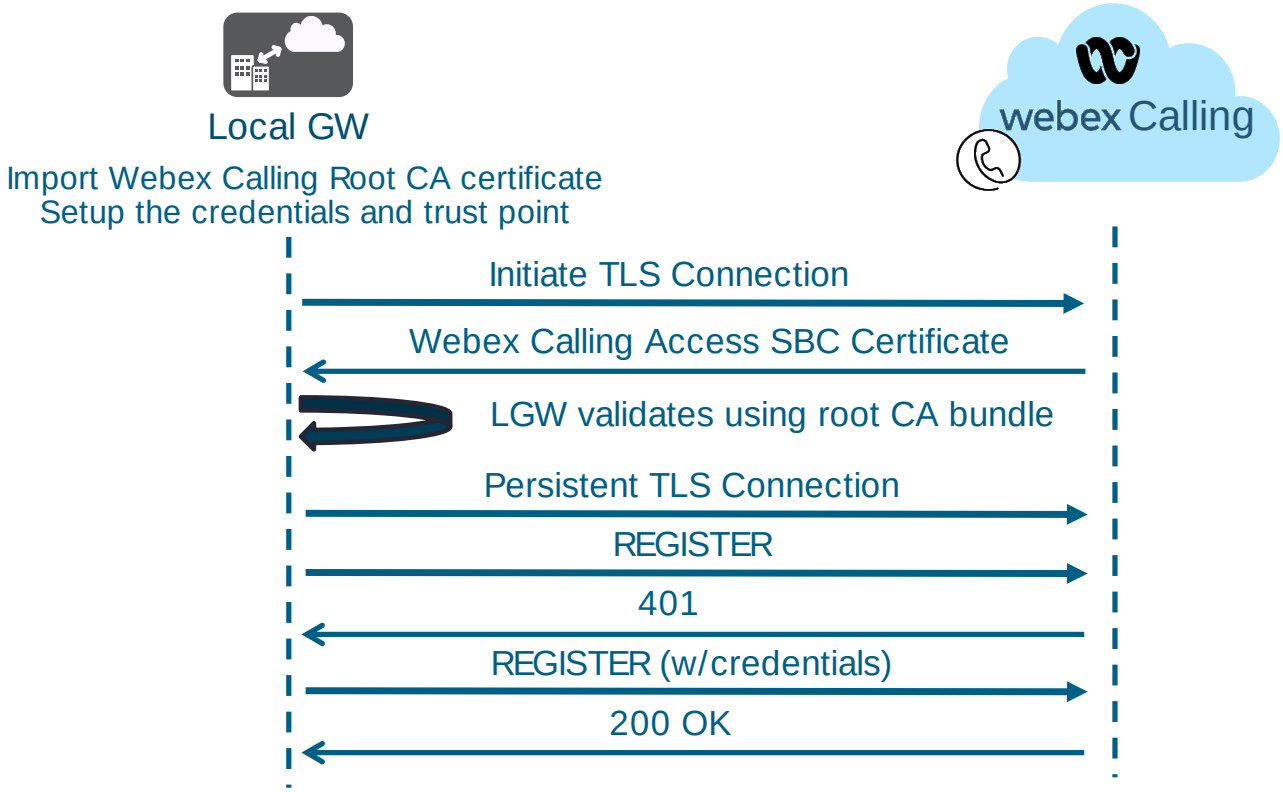


Figure 10 shows the detailed flow of the SIPS connection setup. As part of the provisioning process of the Local Gateway, the Cisco Trusted Core Root Bundle is installed. The bundle contains a set of public CA trust anchors. The digest credentials obtained from Control Hub are also provisioned on the Local Gateway.

The Local Gateway initiates a SRV query to the DNS to get the outbound proxy DNS A-record, TCP port, priority and weight based on the SRV domain obtained in Control Hub during the Local Gateway provisioning process. For redundancy, multiple records are returned. The Local Gateway will select that one with the highest priority (corresponding to the lowest priority number) returned by the DNS and then subsequently sends a DNS A-record query to determine the IP address of the Webex Calling Access SBC. During TLS connection setup, the Local Gateway verifies the authenticity of the Webex Calling Access SBC by validating the presented certificate with the root CA bundle.
On Webex side, the Local Gateway is authenticated and authorized based on the presented digest credentials during SIP registration.

The Local Gateway's connection and secure registration follows RFC3261. Once a persistent TLS connection is established as transport layer for SIP, the Local Gateway sends a SIP REGISTER to the Access SBC. The first REGISTER message is sent without SIP digest credentials, the access SBC then denies registration with a SIP response 401 indicating that authentication is required. On receiving the 401 response, the Local Gateway sends another REGISTER message which contains the required authentication information based on the Local Gateway's SIP digest credentials.

PSTN providers typically do not support video calls. To avoid call setup failures it is recommened to configure the video suppression feature on the Local Gateway as described at https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/voice/cube/configuration/cube-book/voi-audio-forced.html.

Dial Plans

Dial plans add the ability to route calls to premises-based call control instances or between multiple premises-based call control instances based on matches against dial patterns within a dial plan. Each dial plan can have up to 10,000 patterns and a maximum of 10,000 dial plans can be configured for each customer. The routing choice for each dial plan can be a trunk or a route group. Whenever a destination matches a pattern in a dial plan the respective call is sent to the trunk or route group selected as routing choice on the dial plan. As mentioned earlier always using a route group as the routing choice instead of an individual trunk simplifies adding capacity or redundancy later.

Dial plans are defined globally for an enterprise and apply to all users regardless of location. The route selection defined by dial plans is the same for all users.

Patterns and pattern matching

Numeric patterns can either represent E.164 numbers, enterprise numbers, or extensions. Patterns for E.164 numbers start with a leading “+” followed by a sequence of digits “1” to “9” and finally optional wildcard characters. An enterprise number pattern is represented by a sequence of digits “1” to “9” followed by optional wildcard characters. The only valid wildcard characters are “!” and “X” where “!” matches any sequence of digits and “X” matches a single digit “0” to “9”. The “!” wildcard can only occur once at the end and only in an E.164 pattern. For example, “+496100773!” matches all national numbers in Germany (country code 49) starting with “6100773” and 84969XXX matches all dial strings starting with “84969” followed by exactly 3 arbitrary digits.

In addition to numeric patterns, a dial plan can also contain domain patterns to enable routing of alphanumeric SIP URIs. The main use case for this is to enable URI routing between premises-based call control instances interconnected using Webex Calling trunks. Domain patterns are used to match the host portion of SIP URIs. A domain pattern can be a fully qualified domain for an exact match or a domain with a “*.” for a domain suffix match. All domain patterns have at least one dot (“.”).

Dial plan patterns (numeric and URI) are unique within the enterprise; the same pattern cannot exist in two different dial plans. This is enforced to guarantee deterministic routing behavior.

When handling SIP calls if the host portion (right-hand side) of a URI does not refer to Webex Calling, then Webex Calling for route selection only considers the domain patterns by matching the host portion of the URI against the provisioned domain patterns. If the host portion of the URI refers to Webex Calling (any numeric call originating from a Webex Calling phone or received from a trunk falls into this category) then for URIs with a numeric user portion (left-hand side of the URI) Webex Calling will try to find a match on +E.164 patterns first and then on enterprise patterns. No dialing normalization is applied before trying to match +E.164 patterns. If no numeric match is found, then Webex Calling uses the host portion to check for a match on one of the domain dial patterns.

If multiple dial plan matches are found, then best match routing logic is applied, and the most specific pattern is selected. Range patterns using the “X” wildcard are preferred over prefix patterns using “!” to determine the best match.

Interconnect with On-Premises Call Control

Local Gateways provide PSTN access for Webex Calling and can also connect Webex Calling to existing on-premises call control services. This allows customers to keep existing on-premises call control while they transition to Webex Calling. If needed, the on-premises call control can remain in permanent use, co-existing with Webex Calling.

Figure 11 Local Gateway providing connections for both PSTN and On-premises Call Control

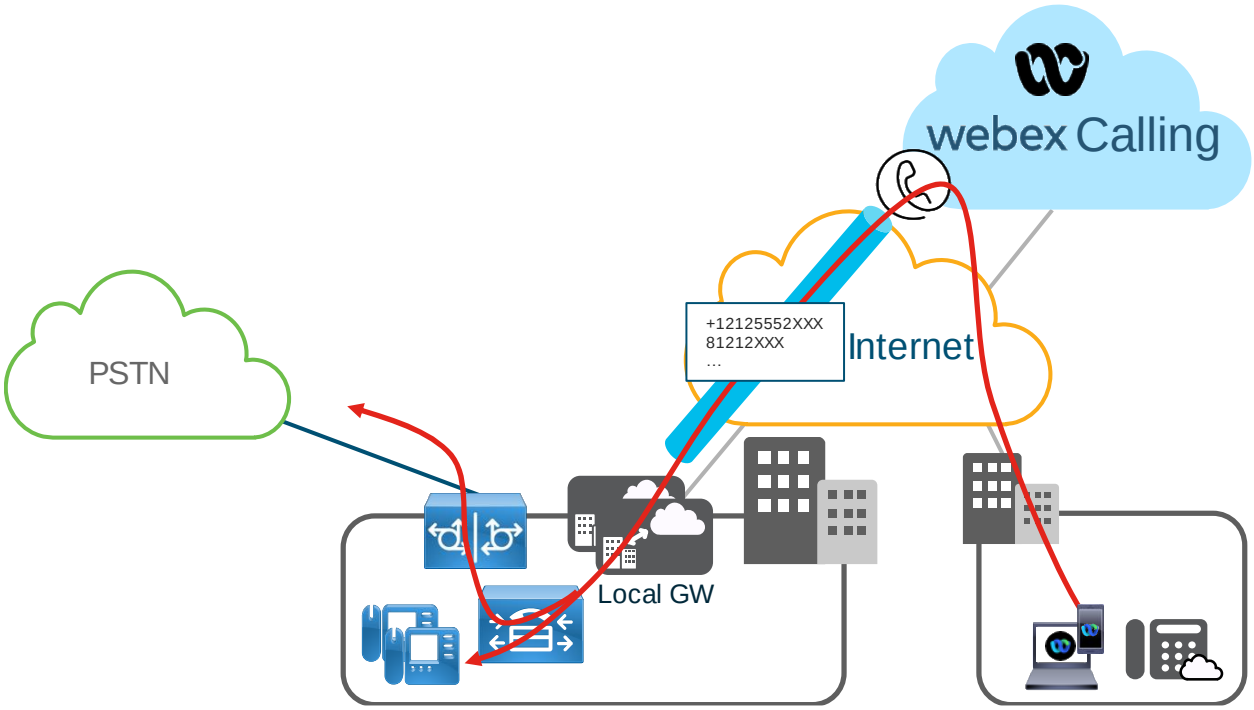
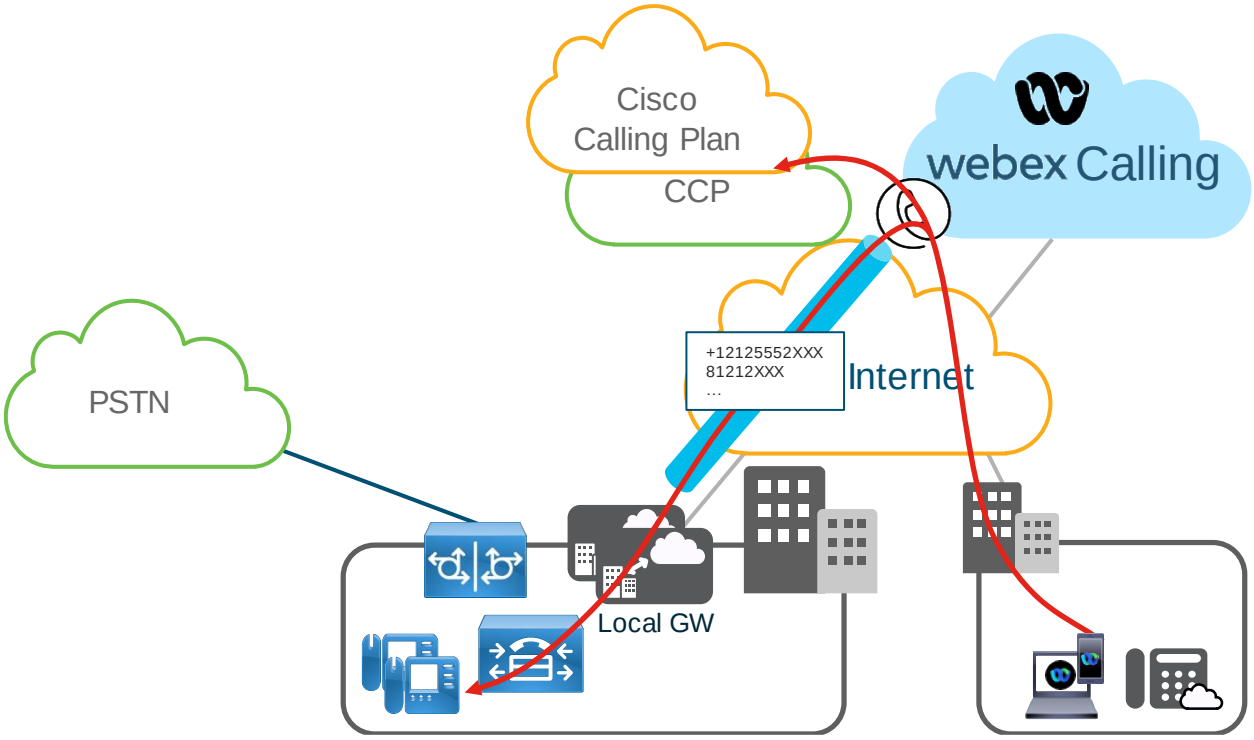


Figure 11 shows a high-level overview of the coexistence of Webex Calling with an on-premises call control instance. A dial plan configured in Webex Calling ensures that enterprise destinations on Unified CM when called from Webex Calling are sent to the trunk or route group terminating to the on-premises call control instance as enterprise calls while all other calls, calls not matching any of the patterns of the dial plan, are sent to the trunk or route group as PSTN calls. In this example, calls sent from Webex Calling first go to the on-premises call control instance where the enterprise dial plan is used to differentiate between on-net and off-net calls. Off-net calls are sent to the PSTN via the regular PSTN gateways of the on-premises call control instance and on-net calls are routed to endpoints registered to the on-premises call control instance or to services provided by the on-premises call control instance.

Inbound and outbound calls via trunks require that the main number of the trunk's location is configured even if that location is only used for trunks and no users are configured within that location.

Combining Premises Trunk and Cloud PSTN for PSTN Access

Figure 12 Combining Premises Trunk and Cloud PSTN for PSTN access



For each Webex Calling location, a PSTN choice (Cisco Calling Plan, Cloud Connected PSTN, or Premises-based PSTN) must be configured. Only one PSTN choice can be selected for each location. Using a Webex Calling dial plan allows to use a choice of cloud-based PSTN (Cisco Calling Plan or Cloud Connected PSTN) for a location while still connecting to an on-premises Unified CM via a trunk or route group. In this case, all destinations matching patterns in the dial plan configured in Webex Calling are sent to the on-premises Unified CM via the trunk or route group as an enterprise call while the configured PSTN choice is used for PSTN calls.

Multiple On-Premises Call Control Instances

Dial plans combined with trunks or route groups can also be used to establish interworking with multiple on-premises call control instances.

Figure 13 Interconnecting multiple On-Premises Call Control Instances

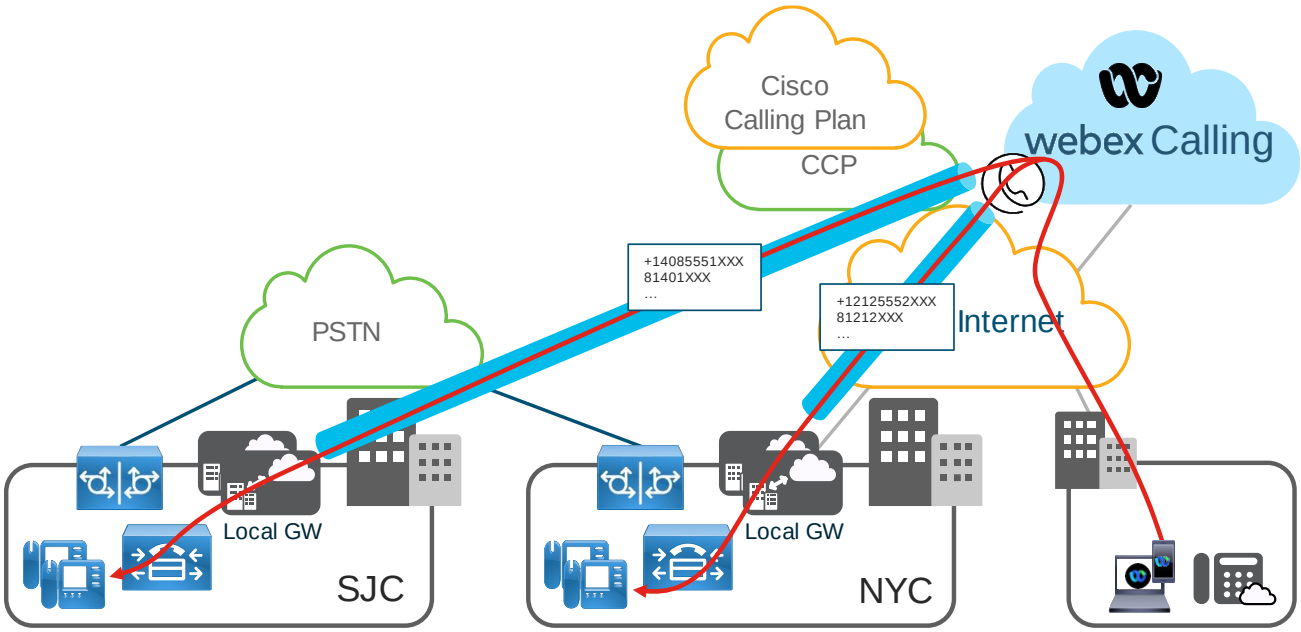
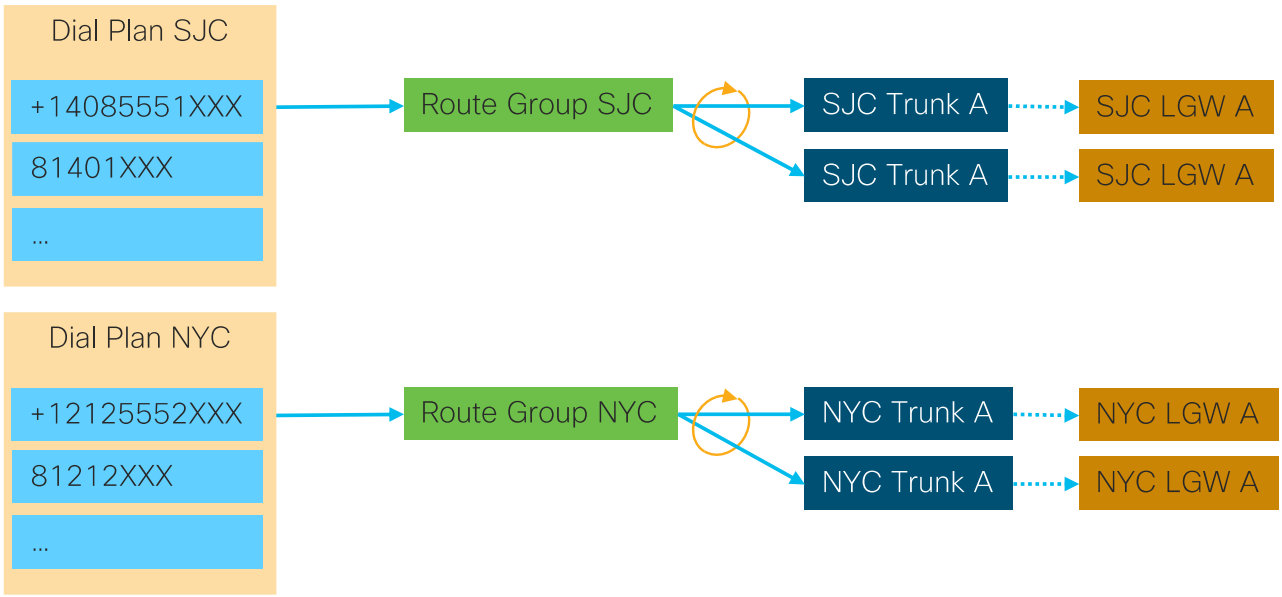


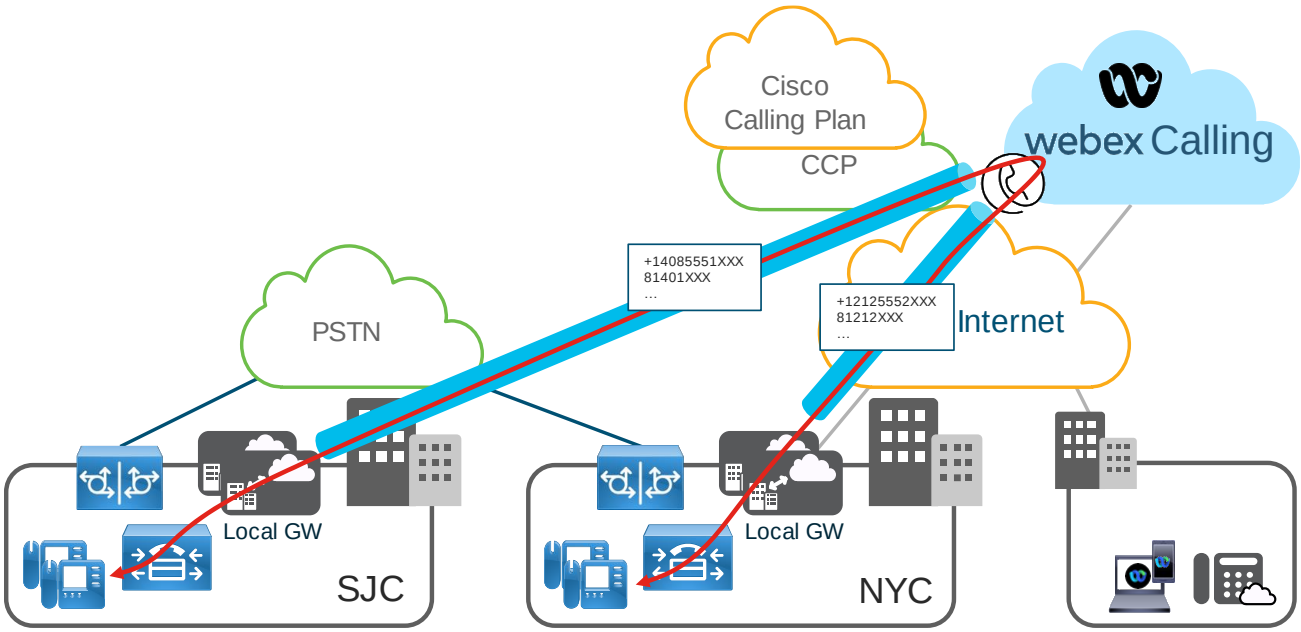
Figure 13 shows an example where two on-premises call control instances are connected to Webex Calling at the same time. Two sets of Local Gateways establish the connection to Webex Calling. Each set of Local Gateways is represented by a route group in Webex Calling where each route group contains the trunks connecting to the individual Local Gateways. This is shown in the following illustration

Figure 14 Dial Plan Setup for two On-Premises Call Control Instances



For calls originating from a Webex Calling user, the dialed destination is matched against all patterns and if a match is found, the call is routed to the route group or trunk configured as the routing choice on the dial plan which contains the pattern. Normalization of the dial string according to the relevant national dial plan is applied prior to the match. Using the example above all, “81212345”, “+12125552300”, and “9125552300” (assuming an outside access code of “9”), are routed to the NYC on-premises location where for the last two dial strings the called destination on the SIP INVITE going out to the Local Gateway is “+12125552300”.

Figure 15 Routing Between On-Premises Call Control Instances



The numeric routing schema established by Webex Calling dial plans not only is used to steer calls originating from Webex Calling users to a specific call control **instance**, but it also applies to calls originating from on-premises users and which are sent to Webex Calling by the on-premises call control instance via one of the trunks. If the called address of a call received by Webex Calling on a trunk does not match any Webex Calling address (user or any service) then like calls originating from Webex Calling users, the dial plan patterns are matched to determine which trunk or route group to send the call to. Webex Calling in this case acts as a tandem between all on-premises call control instances.

Call Routing Overview

For calls originating from Webex Calling users the dial string is first compared with the emergency numbers defined in the national numbering plan for the calling user. In case of a match the call is routed as emergency call.

Then the dial string is checked for a match on a TN, extension, or enterprise number of any Webex Calling user. In this step only very limited (naïve) number normalization is applied. For example, a call dialed as 9-1-972-555-0101 (with PSTN access code and national prefix) would not match TN +1-972-555-0101 right away. In case of a match the call is routed to the respective user.

Next the dial string is matched against virtual on-net extensions and in case of a successful match the call logic for virtual on-net extensions is invoked. Next the dial string is matched against numeric dial plan patterns and in case of a match the call is sent to the dial plan’s destination as premises call.

Finally, the dial string is analyzed based on the national dial plan of the calling user. As a result of the dial plan lookup the dial string is transformed to +E.164 if possible. For toll-free, international, and national calls if the dial string was transformed call routing starts again above with trying to find a match for a Webex Calling user. This secondary lookup makes sure that calls are routed on-net independent of the dialing habit used to dial the call.

If no secondary lookup is needed and the national numbering plan lookup led to a match, then the call is sent to the PSTN. Handling of unknown (not matched by the national numbering plan) dial strings is covered in the next section.

The Remote Office and Office Anywhere services do not follow above process. The Remote Office/Office Anywhere destination is always treated as external and hence there is never an attempt to match the destination to an user, virtual on-net extension or dial plan.

Unknown Number Handling

The routing behavior for unknown numbers is controlled by two settings. The enterprise wide “Unknown Number Handling” setting and the “Calls to On-Premises Extension” setting at the location level. The enterprise level “Unknown Number Handling” setting can be set to “Standard behavior” or “Legacy”. For new deployments “Standard behavior” should be selected and call routing to on-premises should be established by configuring Webex Calling dial plans. The other option exists mainly to provide backward compatibility for deployments that existed prior to the introduction of Webex Calling dial plans.

For calls originating from a Webex Calling user, the dialed number is first checked against emergency numbers, Webex Calling numbers defined for that customer, virtual on-net extensions, dial plan patterns, and finally national numbering plan patterns. This has been described in the previous section in more detail. If no match is found, and the dial string has between two and six digits (the dialed number is considered an unknown extension), and “Calls to On-Premises Extensions” are enabled for the location the call is originating from, then the call is routed as an “unknown extension” call to the trunk or route group configured for calls to on-premises extensions for that location. With “Calls to On-Premises Extensions” disabled any call to any unknown destination is sent to the trunk or route group configured as the location’s PSTN only if enterprise level “Unknown Number Handling” is set to “Legacy”. Calls to unknown numbers are never sent to cloud PSTN (Cisco Calling Plan or Cloud Connected PSTN). In all other cases calls to unknown numbers are rejected. Sending all unknown numbers to the premises based PSTN when “Unknown Number Handling” is set to “Legacy” is backward compatible with the routing behavior of Webex Calling prior to the introduction of Webex Calling dial plans.

Premises calls (calls originiating from a trunk where the calling party matches a pattern in a Webex Calling dial plan) to unknown numbers are only routed if the called destination is an extension (2-6 digits) and “Calls to On-Premises Extensions” is enabled on the location of the originatring trunk. In all other cases the call is rejected.

Table 2 summarizes the routing behavior for an unknown number.

Table 2 Routing Logic for Unknown Numbers

			Global "Unknown Number Handling"	
From	**To**	**"Calls to On-Premises Extensions"**	**"Standard"**	**"Legacy"**
Webex Calling	**Unknown extension (e.g., 4099)**	**Enabled**	Route to premises PSTN choice configured on location level "Calls to On-Premises Extensions" option	
		Disabled	rejected	Route to PSTN (if premises based PSTN is selected)
	Other unknown number	**any**	rejected	Route to PSTN (if premises based PSTN is selected)
Premises	**Unknown extension (e.g., 4099)**	**Enabled**	Route to premises PSTN choice configured on location level "Calls to On-Premises Extensions" option	
		Disabled	rejected	
	Other unknown number	**any**	rejected	
External	**Any unknown number**	**any**	rejected	

Identity sent to Webex Calling on trunks

For PSTN calls over a Webex Calling trunk (in either direction) identity (calling, connected) has to be sent in +E.164 format. Other formats (extension, ESN) are only acceptable for calls to or from from premises users.

Trunk calls to Webex Calling

Inbound to Webex Calling, the originating trunk is identified by the otg tag inserted into the "From" header by the Local Gateway. Then the unscreened calling identity of the call is used to determine the type of the call.

When the inbound call from Local Gateway requests privacy, the asserted identity carried in the P-Asserted-Identity (PAI) header is selected as the unscreened calling line identity if the header is present. If no privacy is requested, or privacy is requested but no PAI header is present, the presentation identity carried in the "From" header is selected as the unscreened calling line identity.

If Dual Identity Support is enabled on the trunk, then the received "From" and PAI headers are treated independently, and both are kept for further processing. With Dual Identity Support turned off only one received identity is kept for further processing with the PAI header taking precedence.

Webex Calling screening services use a set of selective criteria to perform differently depending on the result of the screening. For example, Call Forwarding Selective can be used to forward an incoming call to a specified destination when the calling address matches calling address criteria specified by the user. In the cases of inbound call from premises PBX to a user, the user's screening services require the calling number to be in +E.164 format. The screening services should not specify premises ESN/unknown extension as a calling party criterion.

For an inbound call, the unscreened caller ID is matched against dial plan patterns configured for the customer. The user portion is used for matches against numeric dial plan patterns and the host portion is used for domain pattern matches. If a match is found, then the call is classified as premises call.

If no dial plan match is found and the global “Unknown Number Handling” setting is set to “Standard”, then the “Calls to On-Premises Extension” setting of the inbound trunks is checked. If “Calls to On-Premises Extension” is enabled and the numeric caller id length is between two and six digits, then the call is considered to originate from an enterprise extension and classified as premises call.

Finally, if the global “Unknown Number Handling” setting is set to “Legacy” then the call is also classified as premises call.

If none of these checks succeed, the call is classified as a network (PSTN) call.

Table 3 summarizes this behavior

Table 3 Call classification for inbound calls on trunk

	Global “Unknown Number Handling”		
	“Standard”		**“Legacy”**
	“Calls to On-Premises Extensions”		
Incoming caller id (number)	**Enabled**	**Disabled**	
Dial plan match	Premises		Premises
Unknown extension (e.g., 4200)	Premises	External	Premises
Other unknown number	External		Premises

The call classification impacts how services work which consider whether a call is originating from an on-net source or from the PSTN. Additionally, the call classification determines which destinations can be reached (see next section for details).

Allowed transit calls and caller ID selection

Webex Calling dial plans enable deterministic routing between on-premises call control instances and from Webex Calling users to on-premises call control instances. Webex Calling dial plans do not aim at providing cloud PSTN for on-premises users. Transit calls from on-premises call control instances to Cloud Connected PSTN or Cisco PSTN are not allowed. Table 4 summarizes which transit calls are allowed and what is sent as caller ID.

Table 4 Transit call matrix with caller IDs

		To		
		Webex Calling User	**Trunk (DP match)**	**Cloud Connected PSTN or Cisco PSTN**
From	**Webex Calling User**	Allowed, internal caller identity of the user	See next section, Table 5	Allowed, external caller identity of the user
	Trunk	Allowed, incoming “From”	Allowed [1)], incoming “From”	rejected
	Cloud Connected PSTN or Cisco PTSN	Allowed, incoming “From”	rejected	rejected

[1)]Trunk to trunk calls are only allowed if the incoming calls are classified as premises calls.

Incoming calls from a trunk are only routed to a dialed extension/ESN if either the caller id is an extension (2-6 digits) or the caller id matches a dial plan pattern and thus the call is classified as a premises call. An extension (2-6 digits) as caller id is only accepted if for the trunk's location routing of unknown extensions to the premises as internal calls has been enabled. Calls not meeting above requirements are rejected with status code 604.

Outbound caller ID on trunks for calls from Webex Calling users

When premises based PSTN is used as the PSTN choice for a location, then for PSTN calls sent to a trunk the presentation identity is always determined by the user’s caller ID configuration.

For calls from Webex Calling users to premises destinations, Webex Calling checks the “Caller ID Format for Calls from and to On-premises” parameter to determine the presentation identity to send to the premises-based Local Gateway.

When “Caller ID Format for Calls from and to On-premises” is set to “ESN (Location routing prefix + user extension)”, then the desired presentation identity format can only be satisfied if the calling user has an extension and the caller’s location has a site prefix configured. If either the user has no extension or the site prefix is missing on the calling user’s location, then instead the presentation identity value is determined according to the user’s caller ID configuration (direct line, location number, or assigned number from user’s location) and the presentation identity is sent as +E.164.

When “Caller ID Format for Calls from and to On-premises” is set to “+E.164 phone number”, then Webex Calling ignores the user’s caller ID configuration and tries to use the user’s direct line as presentation identity. If the user doesn’t have a phone number, then the presentation identity value is determined according to the user’s caller ID configuration.

Webex Calling supports separate presentation and asserted identities. On the outgoing side, the presentation identity is included in the “From” header, and the asserted identity is included in the P-Asserted-Identity header. The asserted identity for users with a direct line is always the direct line and only for users without a direct line the location’s main number is used.

The following table summarizes what is sent as presentation identity in the “From” header and as asserted identity in the PAI header assuming that the location prefix is set.

Table 5 Webex Calling caller ID presentation

Does the user have phone number and/or extension configured?			"Caller ID Format for Calls from and to On-premises" setting	
			"+E.164 phone number"	"ESN (Location routing prefix + user extension)"
Phone Number	**Extension**	**PAI**	**From**	
Yes	Yes	Phone number, +E.164	DN, +E.164	Location prefix + extension[1]
Yes	No	Phone number, +E.164	DN, +E.164	According to user's caller ID configuration, +E.164
No	Yes	Location main number, +E.164	According to user's caller ID configuration, +E.164	Location prefix + extension[1]
No	No	No calls allowed without either DN or extension configured		

[1]If no location prefix is configured for the location then only the user's extension is used as caller ID

Note: if Dual Identity Support is disabled on the trunk, then the PAI header is always set to the same value as the "From" header.

For the cases marked in red, no correct presentation ID can be sent, and thus, a callback from the called user on the terminating side, for example, on-premises Unified CM, is not possible. If all Webex Calling users at least have an extension provisioned using ESN presentation ID ensures that a callback is possible. The same can only be achieved with +E.164 presentation ID if all Webex Calling users have a phone number, which is less likely.

If ESN presentation ID is configured, then when a call from Webex Calling to an on-premises user is forwarded to the PSTN, no valid caller ID is available for the PSTN call leg (ESN is not a valid PSTN caller ID). In that case, the caller ID must be masked to some fixed value (main number). Mapping the caller ID from ESN to +E.164 can be achieved if the next-hop call control instance has all ESN to +E.164 mappings of all Webex Calling locations but there is a risk that the originating user doesn't have a phone number and that the naïve ESN to +E.164 mapping results in a number which does not belong to this customer.

Service interactions

Many Webex Calling services use routing lookups involving dial plan matching. For example, for call forward always and auto-attendant (transfer call to operator) services, Webex Calling performs a dial plan lookup to route the call to on-premises PBX in the same way as a regular user origination. The administrator can provision an ESN number as the forward-to/transfer-to numbers in the same way as a user extension.

However, other services, like Office Anywhere, do not invoke dial plan lookups when forking calls to their network locations. In these cases, this feature does not apply.

The screening services use a set of selective criteria to perform differently depending on the result of the screening. The screening service requires the calling number to be in E.164 format. The screening services should not specify premises ESN/unknown number as a calling party criterion.

Video Considerations

Video calls are possible within a single Webex Calling org between video-capable devices or Webex App by dialing a work number or extension. Video calls are also possible by dialing a conference or video address (SIP address) (such as username@example.webex.com). By default, PSTN only carries voice, calls routed via PSTN will be audio calls only.

To video call someone from the Webex App, it can be used the app header by clicking in the Search, meet, and call bar on the top of the app, the system automatically dials the phone number assigned to the contact. If there is not a phone number to dial, the assigned email address is used.

Additionally, the dial pad can be used on the Webex App, phones and Webex devices; from which it can be placed a video call. Just enter a person's number, name, email address, or video address (such as for a room device) in the dial pad and select the video call option.

Media Flows

Media flows initially are always anchored on the Webex Calling Access SBCs. For point-to-point calls involving phones, Local Gateways and the Webex App media path optimization using ICE tries to establish a direct media path between the involved entities. See the Media Path Optimization section for further information

Media Flows for Co-located and PSTN Calls

Figure 16 Webex Calling Media Flows for Co-located and PSTN Calls

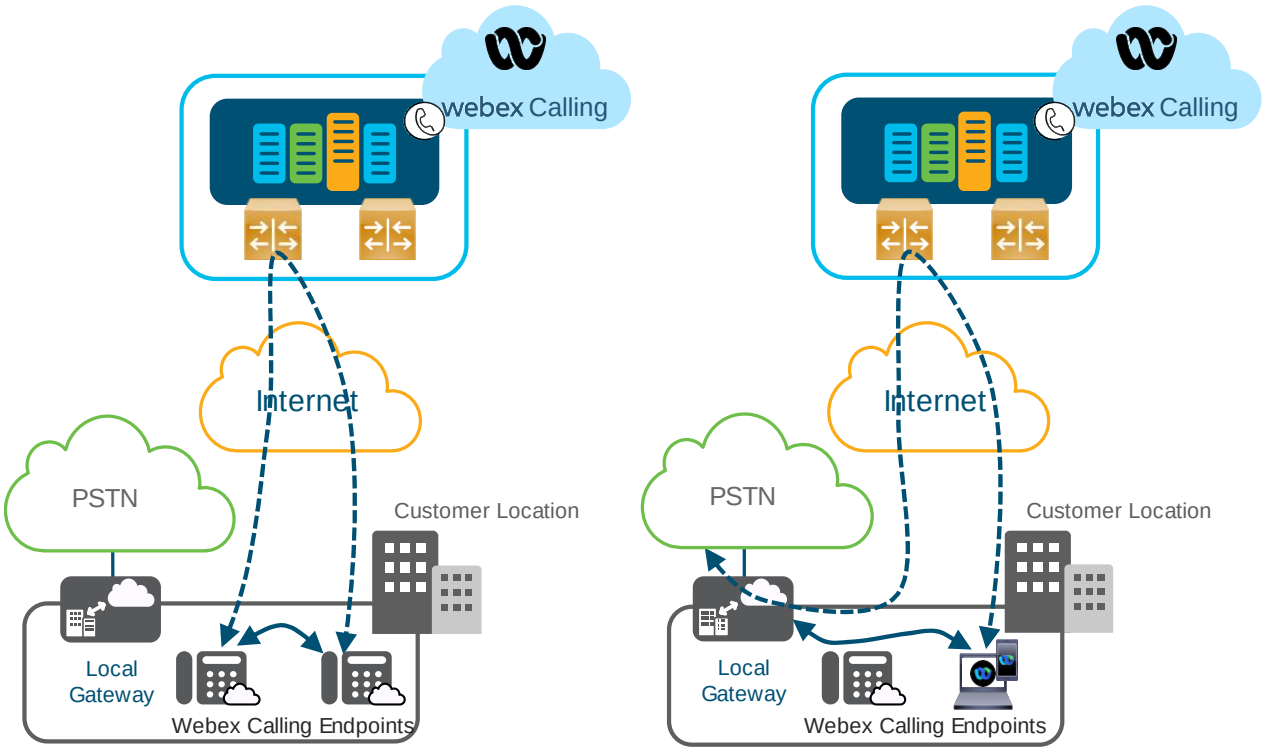


Figure 16 shows two media flow examples. The example on the left shows a call between two co-located Webex Calling endpoints. The media flow is direct between the two endpoints if ICE negotiation is successful (solid line) or anchored on the Webex Calling Access SBC if ICE negotiation fails, causing media to flow from the originating phone via the customer's Internet edge to the Webex Calling Access SBC and then back to the destination endpoint (dotted line). Media in the opposite direction follows the same path.

The example on the right shows that media for a PSTN call is sent from the originating endpoint via the customer's Internet edge to the Webex Calling Access SBC and from there back again via the customer's Internet edge to the Local Gateway.

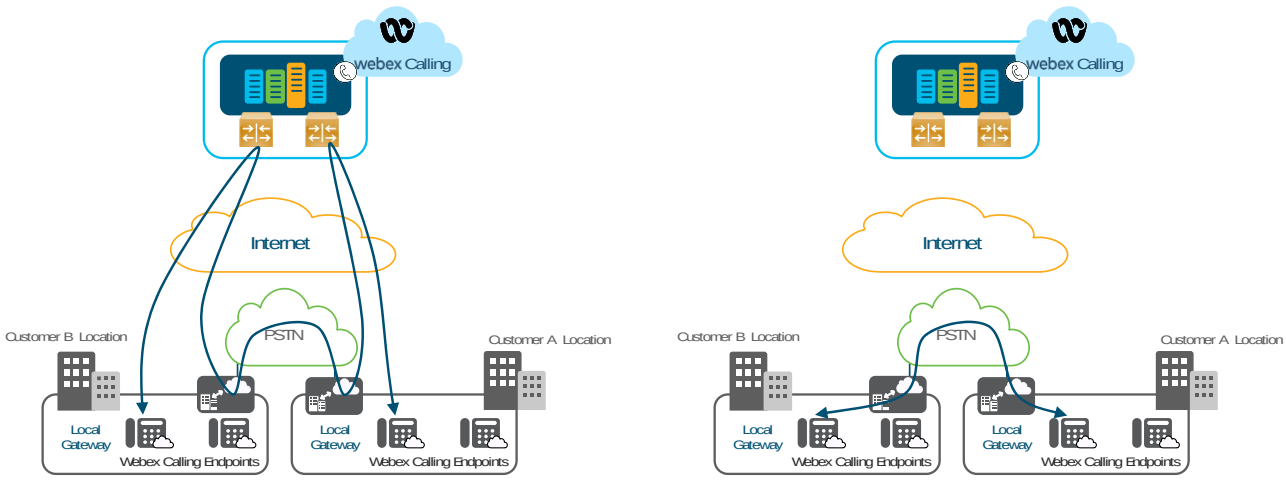
Two full-duplex audio and/or video streams traverse through the customer's Internet edge only if ICE negotiation is not successful; that is, direct UDP media between the two parties is blocked. This means that in most cases the media will stay local. If instead security settings in the corporate network prevent ICE negotiation from being successful, bandwidth for each location needs to be sized accordingly to be able to handle these two full-duplex audio and/or video streams for each intra-location, inter-location, and PSTN call in the busy hour. Bandwidth calculation also needs to account for the fact that all calls initially are anchored on a Webex Calling Access SBC and only switch to the optimized path shortly after call establishment. These calls can be factored into the bandwidth calculation as calls with a short average call holding time representing the time it takes to switch to the optimized media path.

Media Flows for Calls between Different Webex Calling Customers

Calls between two Webex Calling customers must be routed through the PSTN to meet legal requirements, for example lawful intercept.

Figure 17 Webex Calling Media Flows between Two Webex Calling Customers

.



The picture on the left in Figure 17 shows that without media path optimization the media flow is anchored twice in the Webex Calling access layer and traverses through the Local Gateways of both the Webex Calling customers.

With media path optimization (picture on the right) the call still traverses through the Local Gateways of both the Webex Calling customers, but media goes direct between the Local Gateways and the Webex Calling endpoints.

Webex Calling Regions

Webex Calling operates six regional platforms: US (Dallas, Chicago, New York), Canada (Vancouver, Toronto), UK (London, London), Europe (Amsterdam, Frankfurt), APJC Japan (Tokyo, Osaka) and APJC Australia (Melbourne, Sydney). Each Webex Calling instance provides redundant datacenters within that region. The datacenter in Singapore provides media services for users in certain countries in the APAC region to avoid excessive media round-trip-time with the Webex Calling platforms in Australia or Japan.

Figure 18 Webex Calling's Global Backbone

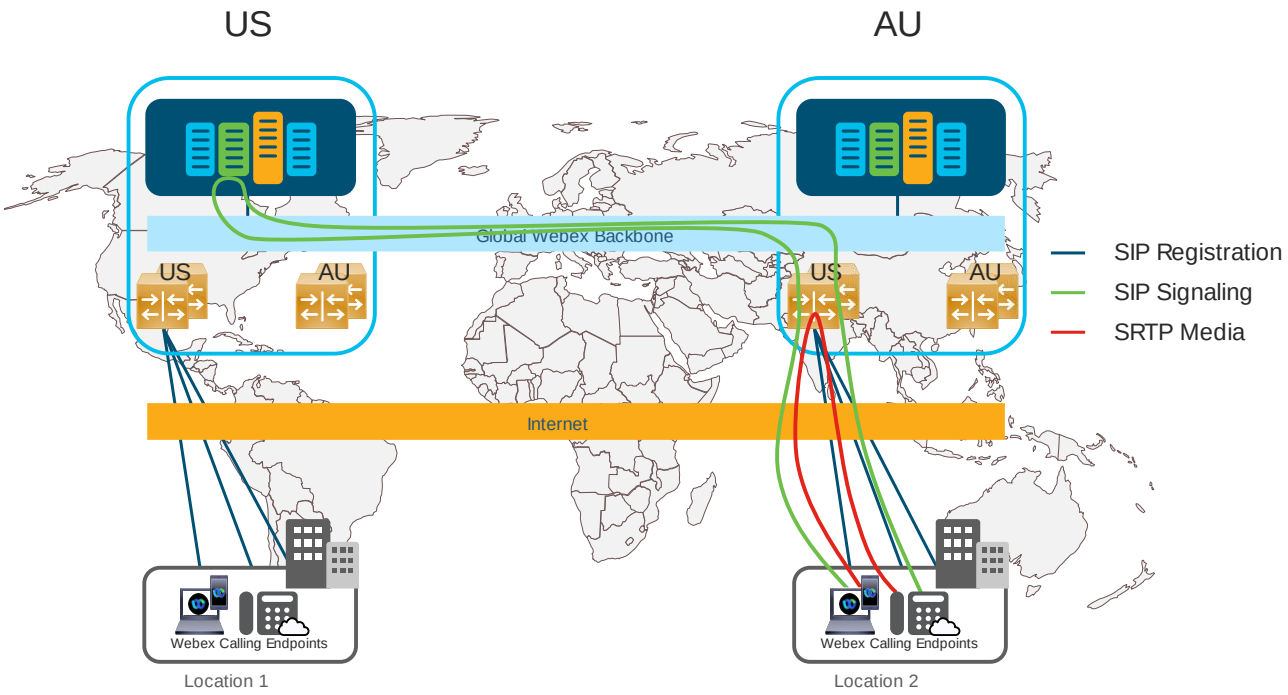
Each Webex Calling customer is provisioned on one of the six Webex Calling instances. All provisioning information of that customer is stored in that Webex Calling instance and the SIP signaling of all endpoints and Local Gateways provisioned for that customer is tied to the Webex Calling instance the customer is provisioned on. Because the initial Webex Calling region selection cannot be changed later it is important to consider all relevant factors as part of the decision process leading to the Webex Calling region selection. To avoid excessive signaling round-trip delay it is important to decide early in the transition process which Webex Calling instance should be used. Cisco recommends selecting the Webex Calling instance which provides the lowest signaling round-trip times for the largest number of users within the deployment.

Another factor to consider in the Webex Calling region selection is the country availability of PSTN services provided by Cloud Connected PSTN (CCP) providers available within that region. While during the transition period PSTN access for Webex Calling devices can be facilitated via a Local Gateway to also enable interworking with Unified CM registered devices, at any point during or after successful completion of the transition, PSTN access for Webex Calling may be switched to Cisco Calling Plan or Cloud Connected PSTN. At that point, the country availability of Cisco Calling Plan or the CCP providers available within the Webex Calling region becomes an important factor.

Refer to the Cloud Connected PSTN providers list available at https://community.cisco.com/t5/collaboration-voice-and-video/cloud-connected-pstn-provider-partners-for-cisco-webex-calling/ta-p/3916211. In addition, for Webex Calling

country availability refer to the *Where is Webex Available* article: https://help.webex.com/en-us/n6fwepj/Where-is-Cisco-Webex-Available#id_98285.

Figure 19 Regional Access SBCs



The region selection for a Webex Calling customer determines the location of the authoritative call control entity for all calls of that customer. To avoid excessive media RTT for calls that are anchored on the Webex Calling access layer, all Webex Calling call control entities can use access resources in each region.

Figure 19 shows the call flow between two phones in Australia. For simplicity, the figure shows only the US and Australia regions, but this concept applies to all regions. Even though the customer is assigned to the US region, the media path is kept on the access SBCs in the Australia region controlled by the US Webex Calling instance. This is only the fallback media path; with ICE media path optimization a direct media path between the endpoint can be achieved. The SIP signaling for this call goes back to the Webex Calling call control entity hosted in the US region because this is the Webex Calling instance on which the customer is provisioned.

Figure 20 Inter-region Media Flow with Regional Access SBC

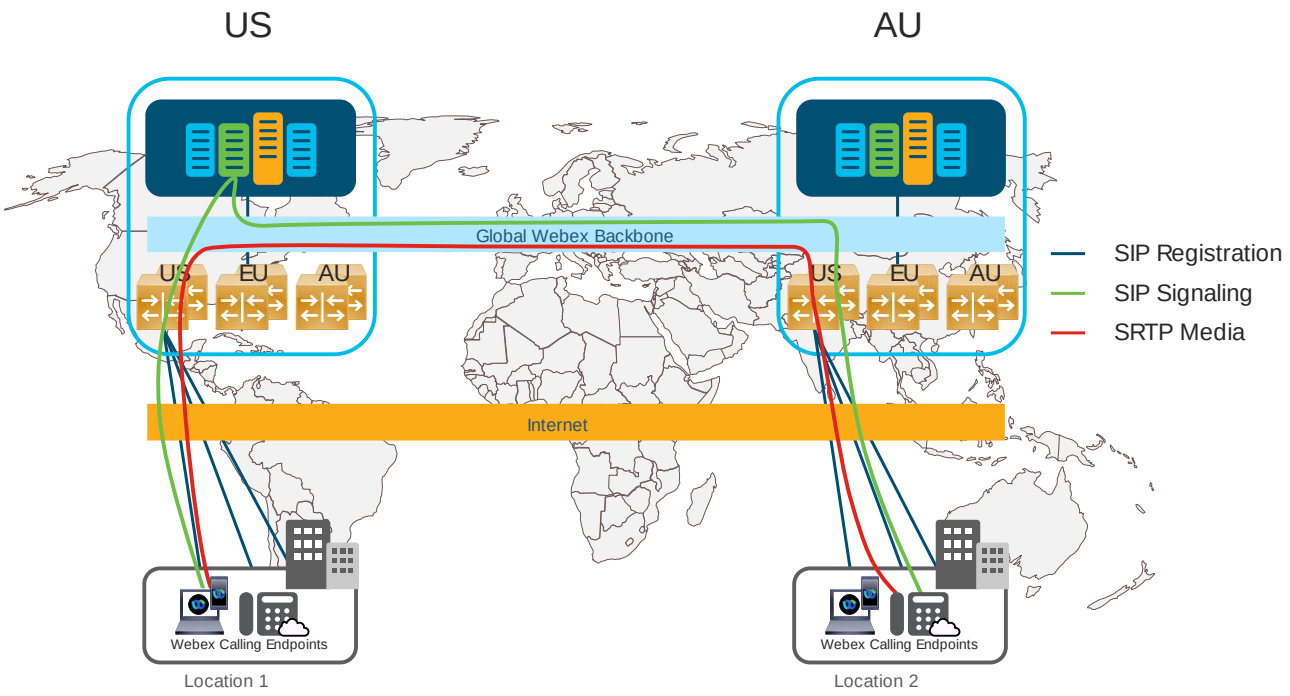


Figure 20 shows how the Access SBCs in both regions anchor the media for inter-region calls. The traffic between the Access SBCs traverses the global Webex backbone and only part of the media flow between the endpoints and the Access SBCs traverses the Internet as the access network.

Media flows involving Local Gateways follow the same schema as described above: media flows involving Local Gateways are also always anchored on the access SBCs of the region in which the Local Gateway is deployed.

Deployment Aspects

This section includes details on various aspects of a Webex Calling deployment including the concept of a Location, Local Gateways, and Directory integration.

Location Definitions

For Webex Calling deployments, the division of the organization into “Locations” will require consideration of many factors. In addition to the PSTN configurations and dial plan considerations, factors like how emergency services are notified and internal telephony services are distributed are significant.

Table 6 User features available across locations

Feature	Permitted for users in other locations	Not permitted for users in other locations
Executive-Assistant	X	
Hunt Group Agent	X	
Call Queue Agent	X	
Auto Attendant user lookup	X	
Line Monitoring (Busy Lamp)	X	
Single Number Reach	X	
Paging Group	X	
Call Park Group		X
Pickup Group		X
Number Movement		X
User Movement		X
Shared Line Appearance		X

Each location must have a PSTN connection defined and may share that connection with other locations. The details of PSTN connectivity will be addressed later in the document, however, there is a default scale limit of 250 concurrent calls that does impact the definition of locations using a Local Gateway. Cloud Connected PSTN is not impacted by this limit.. The use of site dialing prefixes also needs to be considered as will be discussed later.

When adding a location in Control Hub, a physical address and contact name and information are required. For Webex Calling VAR, these are not utilized in emergency contact functionality either to the PSAP or within the organization and are defined separately.

Emergency Calling

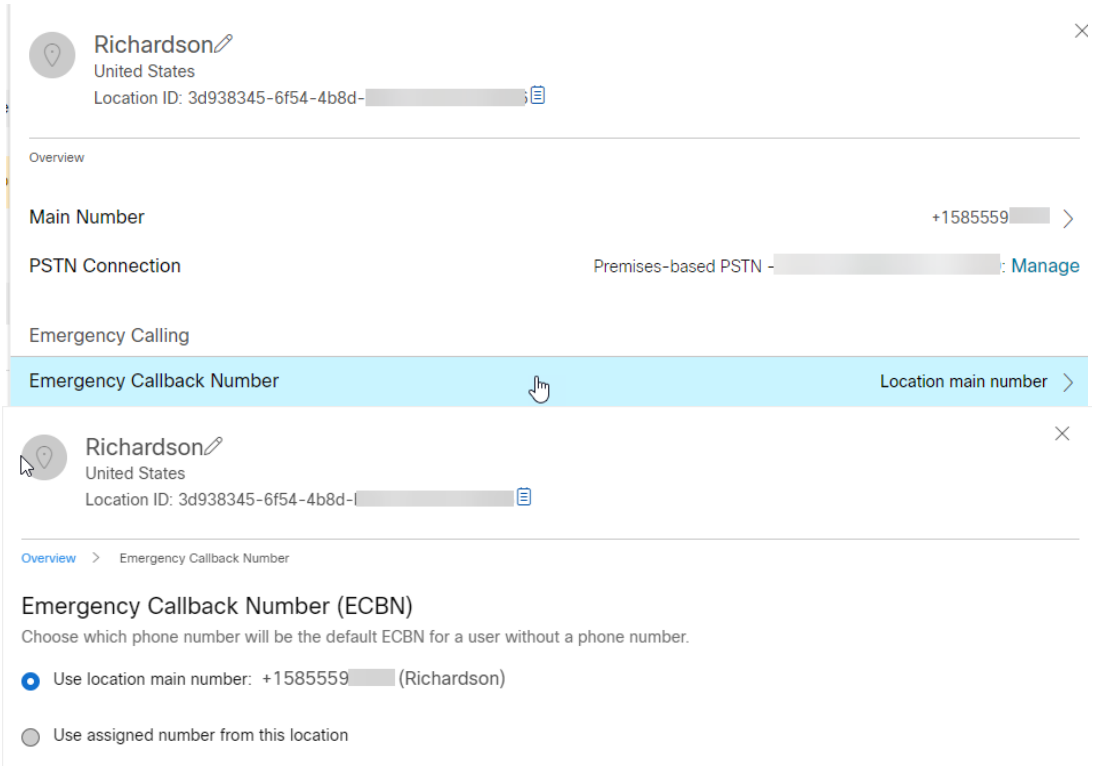
This section describes emergency calling features as required by countries in North America and other regions of the world. In these countries localization's specifications might be so granular to require additional implementation.

The requirement to route an emergency call to the appropriate dispatch center is a requirement for any calling service that offers PSTN service. With Webex Calling, the routing of emergency calls is native to the solution and includes support for all national emergency numbers in the countries that Webex Calling supports. The routing of an emergency call in Webex Calling is based on the location defined within control hub and the PSTN access method of the location (CCP or LGW).

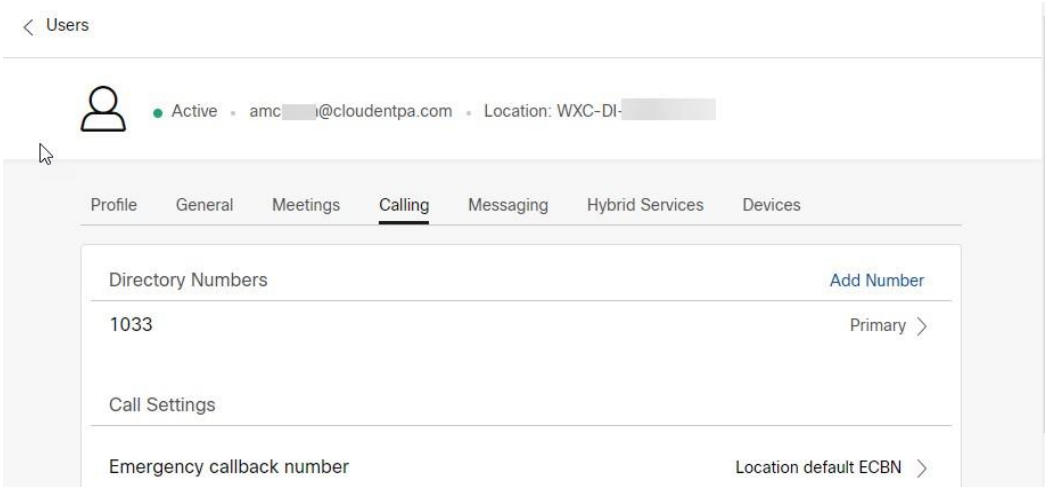
The emergency numbers in Webex Calling are predefined and specific to the country that the Webex Calling users and devices are deployed in.

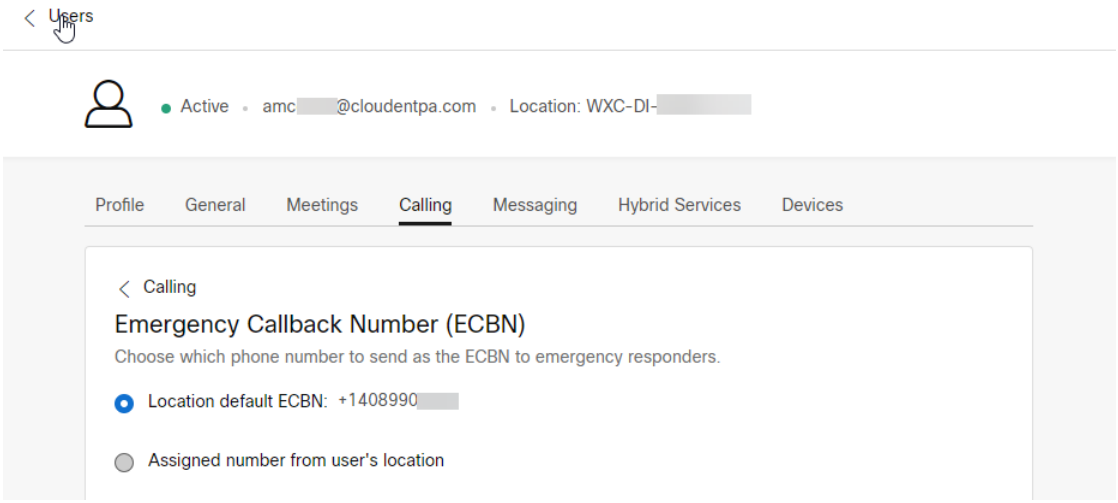
There are two methods to deliver emergency calls in Webex Calling. There is a basic emergency call routing service and an Enhanced Emergency Call routing service. The basic emergency call routing service will use an admin selected number to identify the location and call route to reach emergency services. For basic emergency calling the call path it typically through the customer's CCP provider, but can also use LGW for the emergency call routing. Webex Calling also has Enhanced Emergency call routing designed for US and Canada deployments that have regulatory compliance requirements that require the use of a nationwide provider to deliver emergency calls to the correct dispatch center.

All customers should deploy, at minimum, the basic emergency calling configuration. Basic emergency calling requires that at least one customer owned E164 number be assigned to each location defined in Webex Calling. For basic emergency calling, each location will be defined by a street address that police, fire or ambulance services are dispatched to in case of emergency. In most cases, the main number for the location is the best choice to represent the physical location of the emergency. Typically, the assignment of the address to the E164 number is coordinated with the CCP or LGW provider. The images below show the assignment of the main number to be used as the Emergency Callback Number for the Richardson location.



In most situations, a building address is sufficient for the dispatch address for the location. But if additional location details are needed for specific users or devices, then an administrator can use the same process described above and assign those devices to a specific address or a more precise location inside the address (like a floor or room). In User management in Control Hub, the 'Calling' tab allows for a specific number to be used for a user and their devices to get specific dispatch address. The following images illustrate how a specific number can be assigned to a device. The administrator is responsible to make sure the number used by the device will have the correct dispatch address assigned to it. The address assignment is typically done through the CCP or LGW PSTN service provider.





For US based telephony deployments that must provide enhanced emergency calling solutions, Webex Calling uses RedSky's Horizon Mobility integrated into Webex Calling for emergency call routing. When using RedSky for call routing, an administrator must enroll for an account through Cisco and configure the appropriate information in the Calling->Service Settings to enable this feature. Once the RedSky service has been enabled at the system level, an administrator will enable the RedSky service at each Location level. The enablement of Enhanced Emergency Calling in a Webex Calling Location will activate the service for all devices that are assigned to that Location. Devices that support Enhanced Emergency Calling are Cisco MPP phones and Cisco's Webex App.

There are two setting for enabling Enhanced Emergency Calling at a Location. The "Allow RedSky to receive network connectivity information and test calls" should be used to verify that the RedSky configuration for device and infrastructure mappings are correct. This setting also allows test calls to be place to 933 to perform location verification using RedSky's IVR system to read out the location of the caller. Although this document will not cover the RedSky configuration for location tracking, an administrator should ALWAYS test their location discovery prior to activating emergency calls to route to RedSky. Once testing has completed and verified as accurate, the administrator will route calls to RedSky by toggling the 'Route Emergency Calls to RedSky'. This toggle will direct all emergency calls for the location to RedSky for delivery to the answering center for the location.

The Enhanced Emergency Calling settings also apply to Webex App clients both on-premises and off-premises. When on-premise, the Webex App can be tracked the same way that MPP phones are tracked. When off-premise, the user will be able to set their location dynamically directly within the Webex App. For additional information on emergency calling, refer to the Webex Help Center article Enhanced Emergency Calling for Webex Calling at https://help.webex.com/en-us/article/av6oo3/Enhanced-Emergency-Calling-for-Webex-Calling

Telephony services

Most calling features are available to users across defined locations. These include auto attendants, hunt groups, call queues, paging groups, and executive assistant for users and group members in different defined locations. Some features like call park groups and shared line appearance do require users to be within the same defined location.

PSTN DID numbers are also added by Location, and there is no validation of numbers (For example, is the 302-area code available in San Jose, California?). This will allow the organization to have locations that may cross geographic boundaries but make logical use of user allocation (with the correct number allocation for emergency calling).

Network Connectivity

Consider existing provider data connections (MPLS, SD-WAN, and so on) and generally plan for direct Internet access at each location within the customer deployment. Because cloud-based services will be consumed, reliable Internet connectivity with sufficient bandwidth is a base requirement.

Provided reliable network connectivity is available, Webex Calling offers global reach from all the customer locations thus eliminating the need for endpoint survivability.

Local Gateway Deployment Options

Each trunk in Control Hub represents a connection to a single Local Gateway instance. Multiple trunks can be grouped together in a route group to provide more capacity or redundancy. All calls originating from a location that are destined for premises based PSTN are sent to the trunk or a route group containing this trunk based on call routing configuration. A Webex Calling customer can utilize the same Local Gateway instance for multiple locations.

Currently, Webex Calling allows no more than 250 concurrent sessions from a single Local Gateway instance, which by default becomes the session count limit for Local Gateway based calls, that is, premises based PSTN or Inter-site calls between Unified CM and Webex Calling endpoints. Multiple Local Gateways can be combined using the trunk and route groups for increased capacity. However, if a single Local Gateway deployment requires more than 250 concurrent calls, please contact your Cisco account team to explore other deployment options.

Poor network conditions between the Local Gateway and Webex Calling access SBC can limit the performance of the signaling connection leading to an even lower concurrent calls limit. One-way latency between the Local Gateway and the Webex Calling data center should not exceed 100 ms, jitter should be less than 10 ms, and packet loss should be no more than 0.5%.

Any calls exceeding this capacity limit are rejected with a “403 Forbidden”. The “show call active voice” command can be run on the Local Gateway at any instance to determine the total number of active calls.

```
LocalGateway# show call active total-calls

 Total Number of Active Calls :  153
```

If the output of the above command on the Local Gateway shows more than 250 calls (153 in the above example), and troubleshooting reveals some calls getting rejected by the Webex Calling Access SBC with a “403 Forbidden” SIP response code, Cisco Technical Assistance Center (TAC) may be contacted for further assistance.

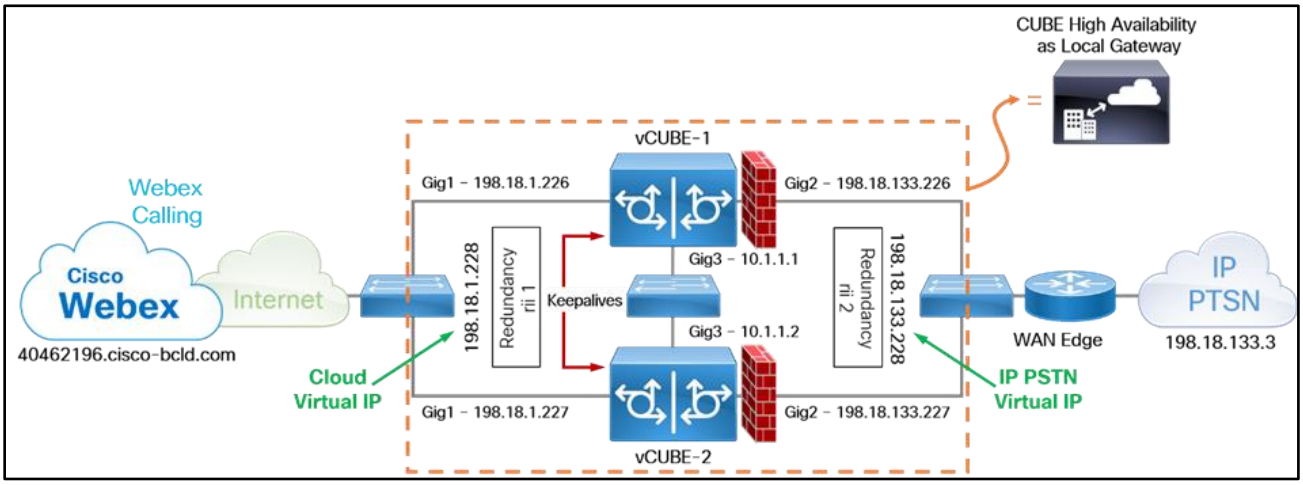
CUBE High Availability as Local Gateway

For all IP-based environments, customers have the option to deploy CUBE high availability (HA) as Local Gateway (LGW) for call preservation. CUBE high availability Layer 2 box-to-box redundancy uses the Redundancy Group (RG) Infrastructure protocol to form an active/standby pair of routers. The active/standby pair share the same virtual IP address (VIP) across the respective interfaces and continually exchange status messages. CUBE session information is check-pointed across the active/standby pair of routers enabling the standby router to immediately take over all CUBE call processing responsibilities if the active router should go out of service, resulting in stateful preservation of signaling and media.

Note**:** Checkpointing is limited to connected calls with media packets. Calls in transit are not checkpointed, for example, Trying or Ringing state.

Refer to Figure 21 below which depicts a typical CUBE high availability as Local Gateway setup.

Figure 21 CUBE High Availability with Local Gateway



The following requirements exist for using CUBE High Availability as Local Gateway for stateful failover of calls:

- CUBE HA as LGW deployment option is available on supported ISR 4000 and CSR1000 series platforms
- CUBE HA cannot have TDM or analog interfaces co-located
- Gig1 and Gig2 are referred to as traffic (SIP/RTP) interfaces and Gig3 is Redundancy Group (RG) Control/data interface
- No more than 2 CUBE HA pairs can be placed in the same layer 2 domain, one with group id 1 and the other with group id 2. If configuring 2 HA pairs with the same group id, RG Control/Data interfaces needs to belong to different layer 2 domains (VLAN, separate switch)
- Port channel is supported for both RG Control/data and traffic interfaces
- All signaling/media is sourced from/to the Virtual IP Address
- Anytime a platform is reloaded in a CUBE-HA relationship, it always boots up as Standby
- Lower address for all the interfaces (Gig1, Gig2, Gig3) should be on the same platform
- Redundancy Interface Identifier (RII) should be unique to a pair/interface combination on the same Layer 2
- Configuration on both the CUBEs must be identical including physical configuration and must be running on the same type of platform and IOS-XE version
- Loopback interfaces cannot be used as bind as they are always up
- Multiple traffic (SIP/RTP) interfaces (Gig1, Gig2) require interface tracking to be configured
- CUBE-HA is not supported over a crossover cable connection for the RG-control/data link (Gig3)

- Both platforms must be identical and be connected via a physical Switch across all likewise interfaces for CUBE HA to work, i.e., GE0/0/0 of CUBE-1 and CUBE-2 must terminate on the same switch and so on.
- Cannot have WAN terminated on CUBEs directly or Data HA on either side
- Both Active/Standby must be in the same Data Center
- It is mandatory to use separate L3 interface for redundancy (RG Control/data, Gig3). i.e., interface used for traffic cannot be used for HA keepalives and checkpointing
- **Upon failover, the previously ACTIVE CUBE goes through a reload by design, preserving signaling/media**

For additional information on CUBE HA as Local Gateway, visit https://help.webex.com/en-us/ndiofhf/Implement-CUBE-High-Availability-as-Local-Gateway.

Firewall Requirements

From an enterprise firewall's perspective, both provisioning and registration are set up using an outbound TLS connection so that no inbound connections need to be allowed on the firewall.

Figure 22 Firewall Traversal Mechanism for TCP Connections

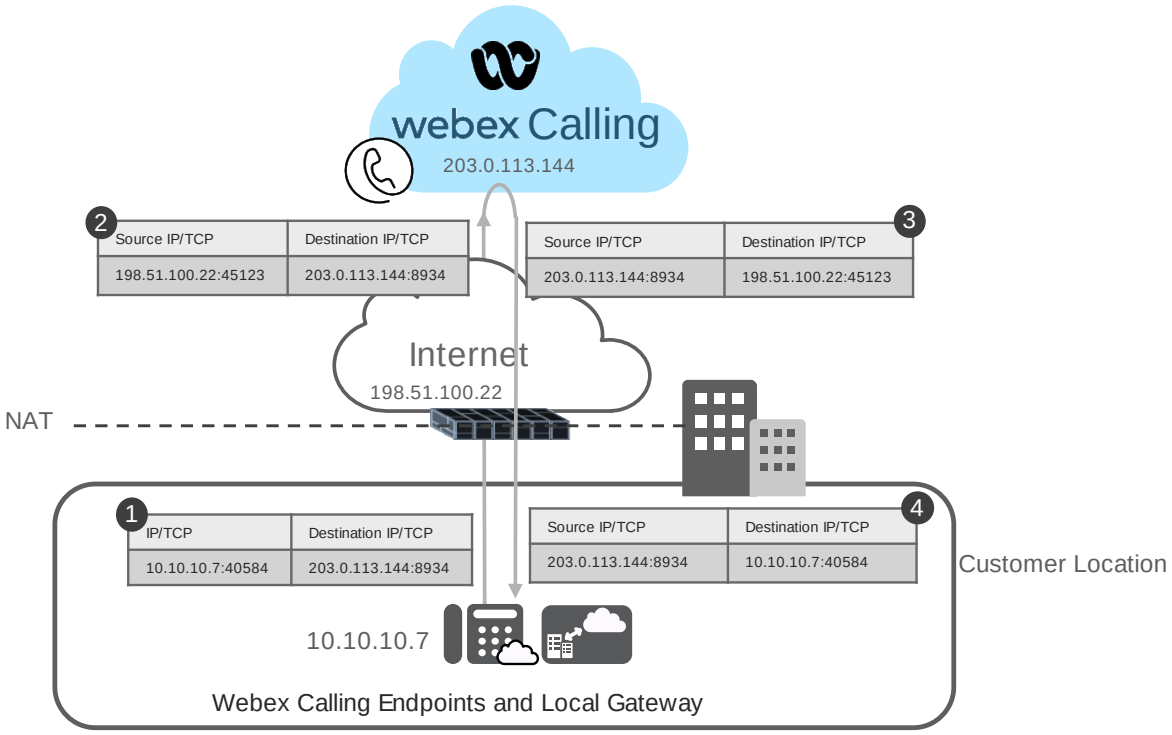


Figure 22 shows this process for an IP phone establishing a TCP/TLS connection for SIP. The firewall translates the private IP address of the phone into a public IP address and changes the source port as shown in step 1 and step 2. The same external transport address is used for the return traffic sent by Webex. The firewall allows the return traffic and translates the public IP address and port back to the internal private IP address and port used by the phone (step 3 and step 4).

Once the TLS connection is established, the phone can send and receive SIP signaling messages to and from Webex Calling. SIP signaling traffic from Webex Calling to the phone reuses the same SIP/TLS connection created by the phone registration. Local Gateways use registering SIP trunks to Webex Calling so that the same mechanism of creating the SIP/TLS transport connection from the inside of the corporate firewall also applies to Local Gateways. The SIP/TLS connections require outbound (egress) TCP port 8934 to be open on the corporate firewalls.

In environments where NAT is applied the transport addresses advertised by Local Gateways and IP phones within the SIP signaling messages are always internal (private) IP addresses. Those private addresses are not reachable by an external entity such as the Webex Calling access. However, the Webex Calling access layer can detect NAT by checking the layer 3 transport addresses of received SIP signaling messages against the addresses contained within the SIP messages. Differing addresses point to the presence of NAT between the Webex Calling access and the endpoint or Local Gateway. When NAT is detected, the SBC always reuses the existing TCP connection and does not send replies to the IP addresses contained in the Via header, or new SIP requests using the addresses contained in the Contact Header.

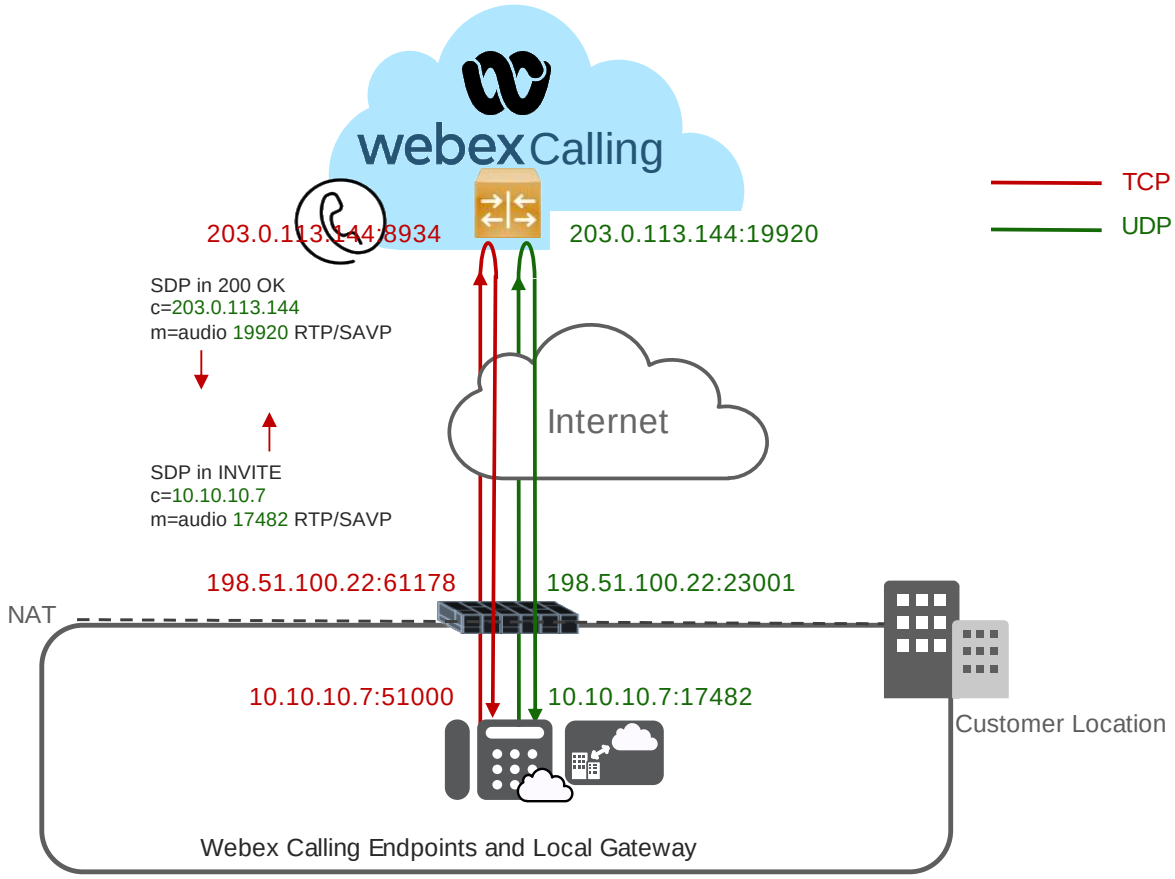
IP addresses to be used to send and receive media are advertised by endpoints and Local Gateways as part of the SDP media negotiation within the SIP signaling. However, in most cases those are private addresses, and the Webex Calling access layer cannot use those addresses to stream media to. There is no way to determine the public IP address and port these private transport addresses get translated to by NAT. To determine the public transport address to be used for return media the Webex access layer waits of for the reception of the first media packet on the listening port negotiated for the access layer as art of the SDP exchange and then uses the source transport address of the received media packet as the destination for media traffic sent back to the phone or Local Gateway.

Using the same logic, the Webex Calling access layer is also able to detect if a NAT change has happened on a Local Gateway or IP phone, for example following a resume after a call hold, and accordingly update IP address and port information for the return traffic which can traverse through the corporate firewall from the outside because they belong to an existing flow.

The mechanism to extrapolate return media address and port from received media packets instead of relying on SDP is called “Media Latching”. Media Latching requires that the endpoints (Local Gateway and IP phones) inside the corporate firewall always send media packets to the Webex Calling access before any media packets can be returned in the opposite direction. The media packets sent from within the corporate network create a NAT binding and open a connection through the firewall.

Figure 23 shows an example for firewall traversal of signaling and media between an on-premises Webex Calling endpoint or Local Gateway and Webex.

Figure 23 Firewall Traversal Mechanism for Webex Calling



This picture shows an example of the firewall traversal for Webex Calling signaling and media.

1. The endpoint or Local Gateway use private IP 10.10.10.7 and port 51000 to initiate a TLS connection. The destination address is the Webex Calling access IP and port, that is 203.0.113.144 and 8934.
2. SIP uses this TLS connection as transport. The endpoint or Local Gateway sends out an INVITE to the SBC, containing the media IP and ports (i.e., 10.10.10.7 and 17482 for audio) it will use to send and receive media.
3. The SBC replies with 200 OK on the same TLS connection. SDP information in that 200 OK message contains the Webex access media IP address and port (i.e., 203.0.113.144 and 19920 for audio).
4. The endpoint or Local Gateway starts sending media to the IP address and port obtained from the SDP received with the 200 OK. The source port of these media packets is the UDP port advertised in the SDP sent with the initial INVITE, in this case, UDP port 17482. The Webex Calling access for now cannot send return packets because IP address and port obtained from the SDP in the initial INVITE by the access will both be translated to still unknown values by the firewall. Once the first media packet hits the Webex Calling access on the port advertised for this call by the access SBC, the SBC learns the public transport address from the source IP address and port of the received media traffic (198.51.100.22:23001 in the example). The connection state in the Webex Calling access is now updated accordingly with this transport address and media from Webex Calling back to the phone or Local Gateway now is streamed using this transport address.

In some scenarios with Local Gateways, there can be situations where both call legs are inbound from a firewall's perspective, such as in the case where an IP phone connected to Webex Calling calls a PSTN number through the Local Gateway, and the called PSTN device transfers the call to another IP Phone belonging to the same network. These two call legs are inbound from the firewall's perspective, and the firewall will block them because no RTP media packets are sent from the inside which would otherwise open a connection on the firewall. To solve this deadlock STUN needs to be

configured on the Local Gateway. With STUN configured the Local Gateway will send STUN packets to the media IP address and port negotiated via SDP. Although no actual media packets are sent the UDP STUN packets from the firewall perspective still constitute outbound UDP packets so that the firewall creates a connection and allows inbound media to flow on the same connection. Failure to configure STUN can prevent bidirectional media in some scenarios.

In summary, Webex Calling does not require to open inbound ports on the firewall. Only outbound UDP and TCP traffic to specific IP addresses and ports must be allowed. The required destinations, including a full list of ports, IP addresses, and DNS domains are included in the Port Reference for Webex Calling document available at. https://help.webex.com/en-us/b2exve/Port-Reference-for-Cisco-Webex-Calling-Value-Added-Resellers#id_112963.

ICE Media Path Optimization

To move from media flows that are anchored on the Webex Calling access to direct peer-to-peer media, where the Webex Calling access is out of the media path, it is fundamental that firewall traversal operations are performed by the clients. This is achieved by using the Interactive Connectivity Establishment (ICE) protocol framework.

ICE introduces a specific terminology that will be used in this present document. The most used terms are explained below.

Transport address: combination of IP address, transport protocol and transport port

Host address: the transport address associated to the endpoint's local interface

Server reflexive address: the transport address learned through the STUN server, corresponding to the firewall NAT translated transport address of the endpoint

Peer reflexive address: the firewall NAT translated transport address learned through connectivity checks

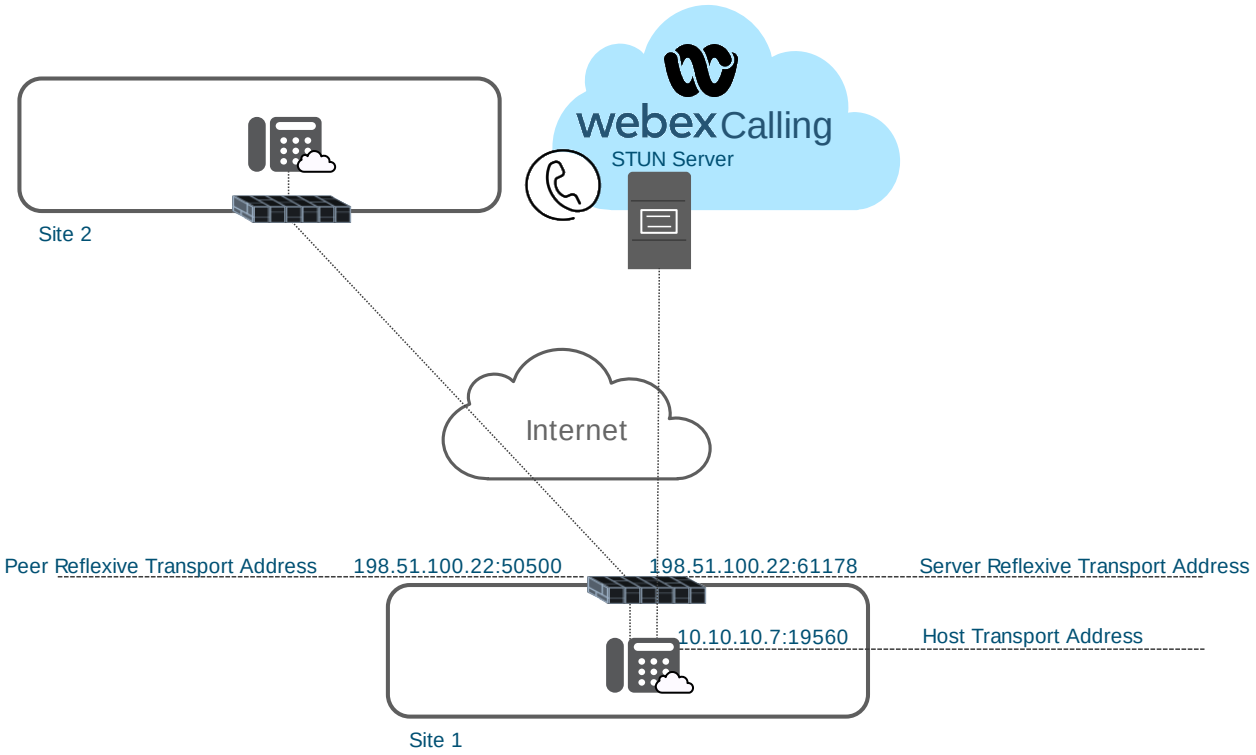
Candidates: Host, Server Reflexive and Peer reflexive are possible candidates to send media to an endpoint. During the connectivity check, one of them will be selected if ICE negotiation is successful

5-tuple: identifies a connection between two systems. It is identified by the transport protocol, source and destination IP, source and destination transport port

Selected pair: the transport address of the calling and the called endpoint selected through connectivity checks

Figure 24 shows the different candidate types:

Figure 24 Host, Server Reflexive, and Peer Reflexive candidates



For an endpoint communicating with systems outside of the firewall the host transport address (local IP address and port of the endpoint) gets translated to an external transport address by the firewall NAT. Figuring out its external transport address is key for an endpoint to establish bidirectional UDP communication with endpoints outside of the firewall. This can be achieved by using a STUN server. Figure 24 shows that for an endpoint sending packets using an internal transport address 10.10.10.7:19560 to two different hosts different external transport addresses can be created by the firewall's NAT, one server reflexive address for packets sent to the STUN server and one peer reflexive address for packets sent to the endpoint in site 2. If a NAT assigns different external transport addresses for communications with different external systems than this is called address dependent. If the external transport address not only depends on the destination IP address but also the destination port, then this is referred to as address and port dependent. If the same external transport address is used for all packets from the same internal transport address independent of destination IP or port, then this is called endpoint-independent NAT.

With endpoint-independent NAT the endpoint can determine its public transport address, the server reflexive transport address, for a given source IP and port by use of an external STUN server before initiating a connection with any external endpoint using this source IP and port. In the presence of address-dependent (or address and port-dependent) NAT the external transport address (peer reflexive transport address) cannot be determined a priori and instead must be learned from the remote endpoint. The procedures involved in learning the server or peer reflexive transport address using them for bidirectional media will be covered in the following sections.

Among the different options that ICE permits, Webex Calling supports the following protocols and functionalities:

- SIP SDP answer/offer model used to carry ICE information
- STUN Protocol
- STUN Server
- UDP-based media
- Automatic fallback to standard media path anchored on the Webex Calling access

Call setup with ICE can be divided into four phases:

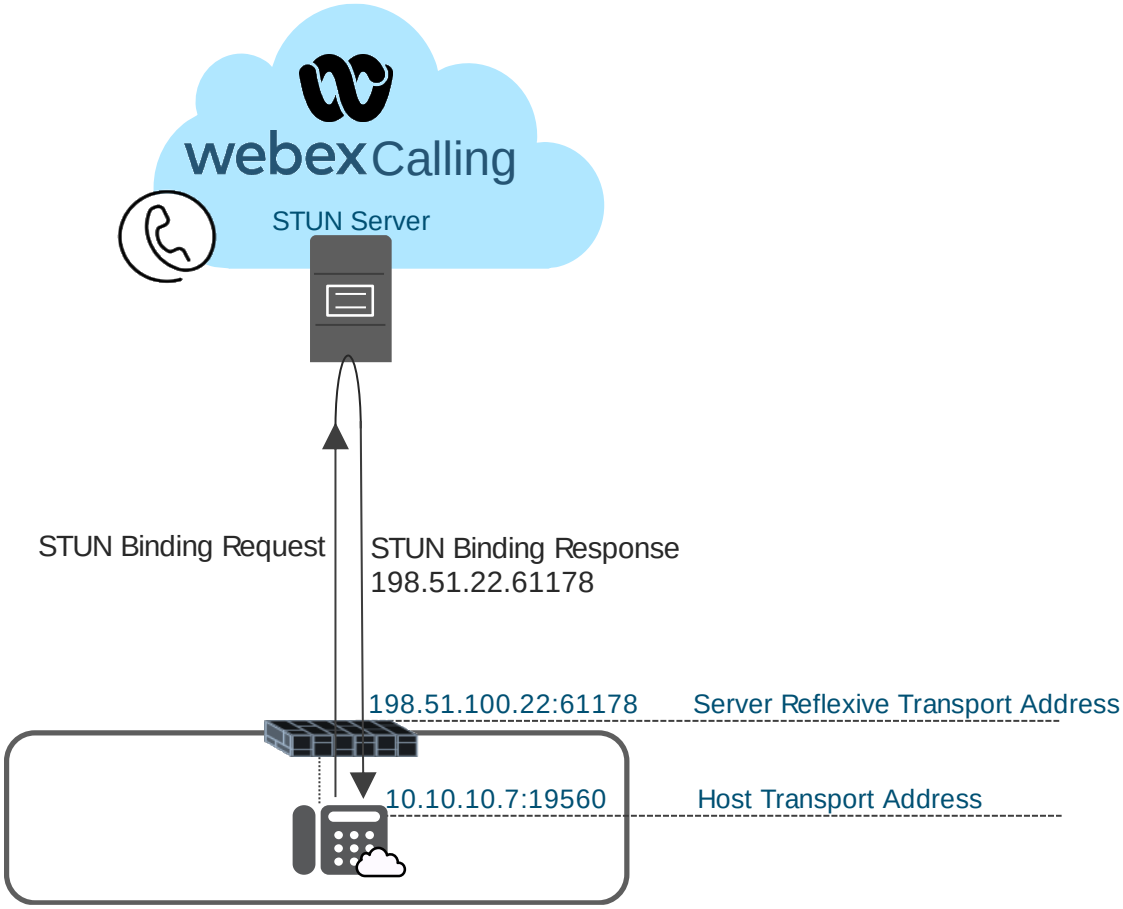
1. Candidate gathering: the endpoints talk to the STUN server to learn their server reflexive address before making or answering a call
2. Candidate exchange: the endpoints communicate their host and server reflexive candidates to the other party through SIP signaling
3. Connectivity checks: host, server reflexive and eventually discovered peer reflexive candidates are checked by each party to nominate a selected pair, that will be used to send and receive the media
4. Media path optimization: the selected pair is advertised through SIP signaling and media starts flowing using the selected pair

Candidates Gathering

Before initiating communication with another peer, the ICE client (Webex app or Webex Calling phone) determines its server reflexive address using a STUN server by sending a UDP STUN binding request packet to the STUN server on port 5004 from the private transport address the client intends to use for the communication with the other peer. The STUN server responds to that request with a STUN binding response containing the server reflexive address which is the external IP address and port as seen by the STUN server. The endpoint uses this STUN binding exchange for each transport address to be used. At least two server reflexive transport addresses need to be determined, one for SRTP and one for Real Time Control Protocol (RTCP).

This process is shown in Figure 25.

Figure 25 Candidates gathering through a STUN server



At the end of this process, the endpoint has the list of candidates that could be used to send and receive media. As an example, as an audio call uses a SRTP and a RTCP UDP port, the SIP offer would include 4 different candidates:

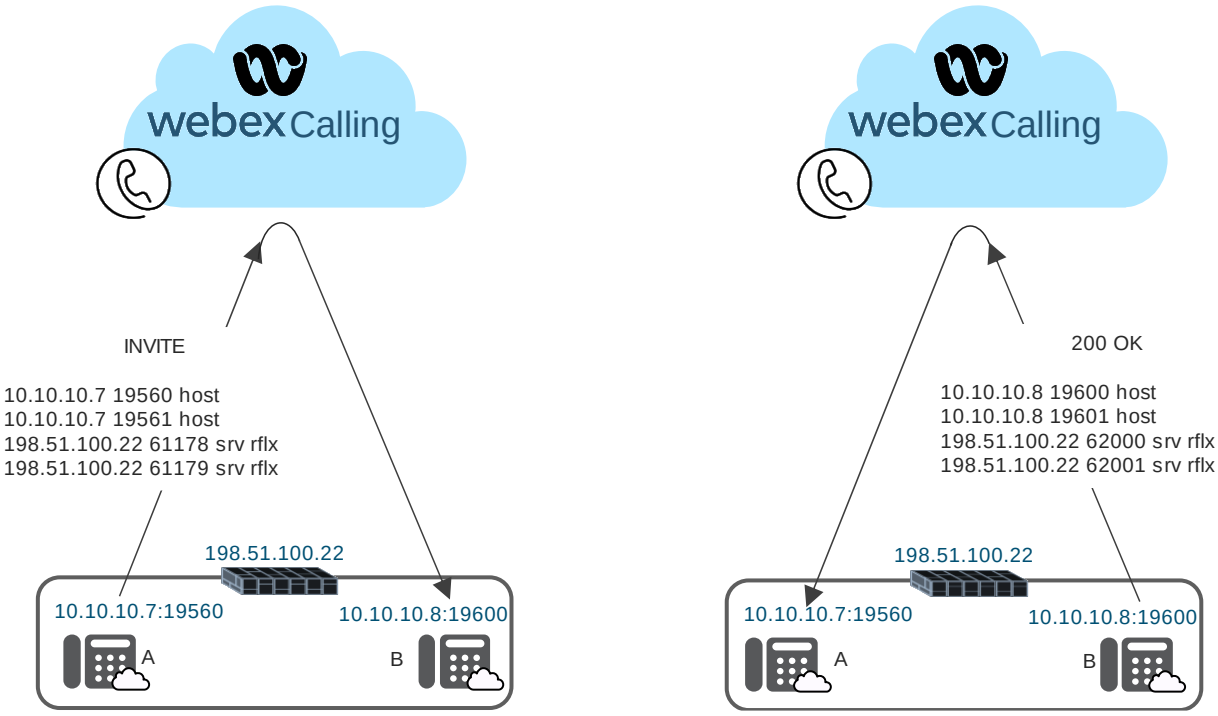
- Host transport address for media (SRTP)
- Host transport address for RTCP
- Server reflexive address for SRTP
- Server reflexive address for RTCP

Candidates Exchange

Gathered candidates are sent in the initial SDP offer in the SIP INVITE message.

The message is sent to Webex Calling and relayed to the called device. This process is shown in the left side of Figure 26.

Figure 26 Candidates exchange through SIP signaling



In response to the INVITE the called device returns its candidates in the SDP answer of the SIP response as shown in Figure 25. The process to obtain the candidates on calling and called device is identical.

At this point, both endpoints have the candidates from each side and connectivity check can start.

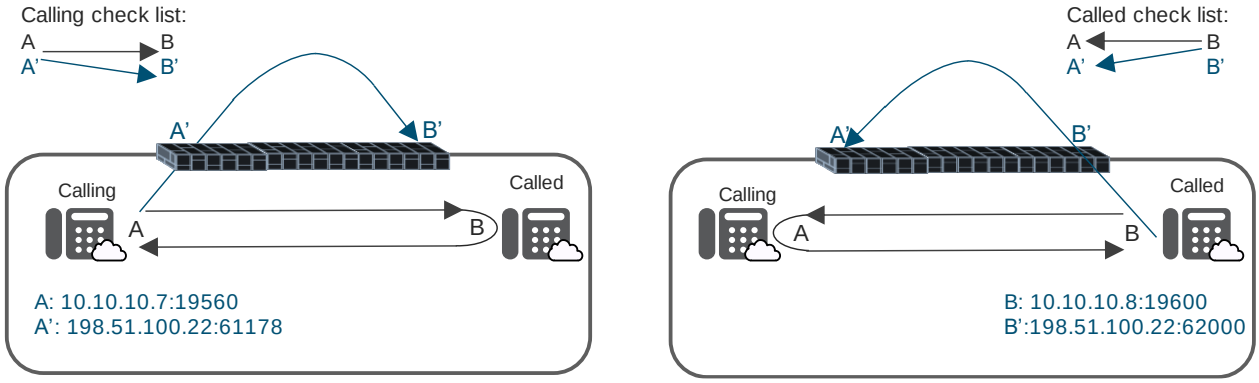
In this example each device receives the following addresses corresponding to the other side:	Called addresses received by the calling endpoint	Calling addresses received by the called endpoint
Host transport address for SRTP	10.10.10.8:19600	10.10.10.7:19560
Host transport address for RTCP	10.10.10.8:19601	10.10.10.7:19561
Server reflexive address for SRTP	198.51.100.22:62000	198.51.100.22:61178
Server reflexive address for RTCP	198.51.100.22:62001	198.51.100.22:61179

Connectivity Checks

Next, candidates are ordered by each endpoint. Connectivity checks using STUN bind requests and responses are performed using the same transport addresses that will be used to send the media. Connectivity checks are bidirectional: every endpoint sends and responds to connectivity checks. A candidate pair is considered valid when connectivity checks are successful in both directions.

In the following example A and B are the host transport addresses, and A' and B' are the server reflexive addresses, respectively for the calling and the called party. The connectivity checks between A and B are successful because they are in the same LAN, as Figure 27 shows. Connectivity checks between A and B' and B' to A might or might not work depending on the firewall operation. Only if the firewall allows access to internal services via the public IP address from inside the local network, then the connectivity checks involving A' and B' will succeed. This firewall behavior is called NAT hair-pinning, NAT reflection, or NAT loopback.

Figure 27 Connectivity checks between two endpoints in the same network

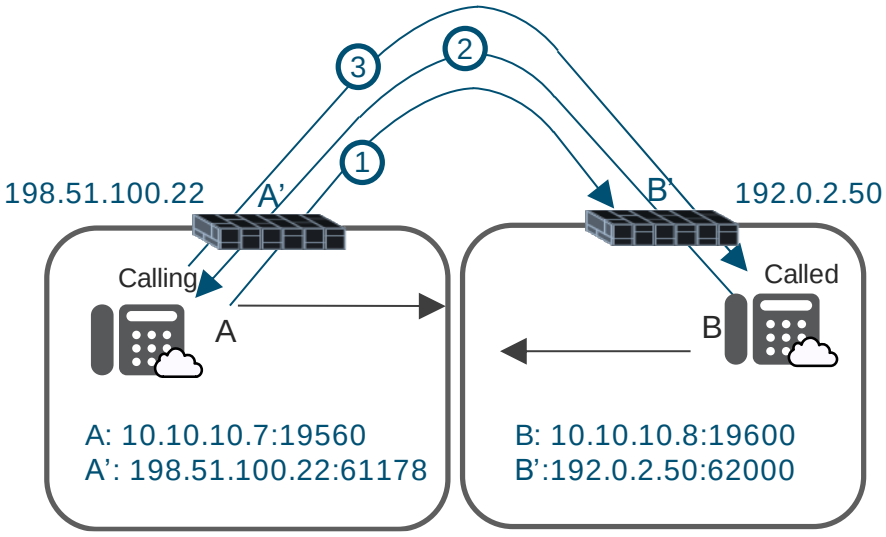


If the firewall is not configured for NAT reflection, then the only working pair is A and B, and is referred to in this document as [A, B]. This notation means that connectivity checks between A to B and B to A have been successful.

If NAT reflection is enabled, the working pairs are [A, B] and [A', B']. Note that [A, B'] and [A', B] in this example match [A', B'], as A is uniquely translated by NAT to A', and B to B'. This is true for endpoint-independent NAT. In the presence of address-dependent NAT the server reflexive candidate as learned from the STUN server is useless because a different peer reflexive transport address is chosen by the firewall for the communication with peers other than the STUN server and packets hitting the server reflexive address from the outside from a transport address other than the STUN server will not be allowed into the network by the firewall. ICE connectivity checks for external transport addresses can only succeed if at least on one side endpoint-independent NAT is used so that one endpoint can obtain a valid server reflexive address. With address-dependent NAT involved the endpoint behind the address-dependent NAT learns its peer reflexive address from the STUN Binding response received from the other endpoint.

In the next example the two endpoints are in disjointed networks. In this case connectivity checks between A to B and B to A fail. Both endpoints also try connectivity to the server reflexive address, as shown in Figure 28.

Figure 28 Connectivity checks between two endpoints in disjointed networks



The first STUN connectivity check message towards a server reflexive address in this example is sent by A towards B' (step 1). Because this is a new inbound connection from the called firewall perspective, this packet is blocked. However, when the calling device sends the STUN message to B', it opens a pinhole in the firewall, and this allows to receive return traffic. Consequently, the STUN connectivity check message between B and A' succeeds (step 2), because it uses the same 5-tuple that has been used from A to B', and consequently it is recognized as return traffic by the firewall.

The STUN message sent by B to A has created a pinhole in the called firewall; and for this reason, a new connectivity check sent by A towards B' now succeeds.

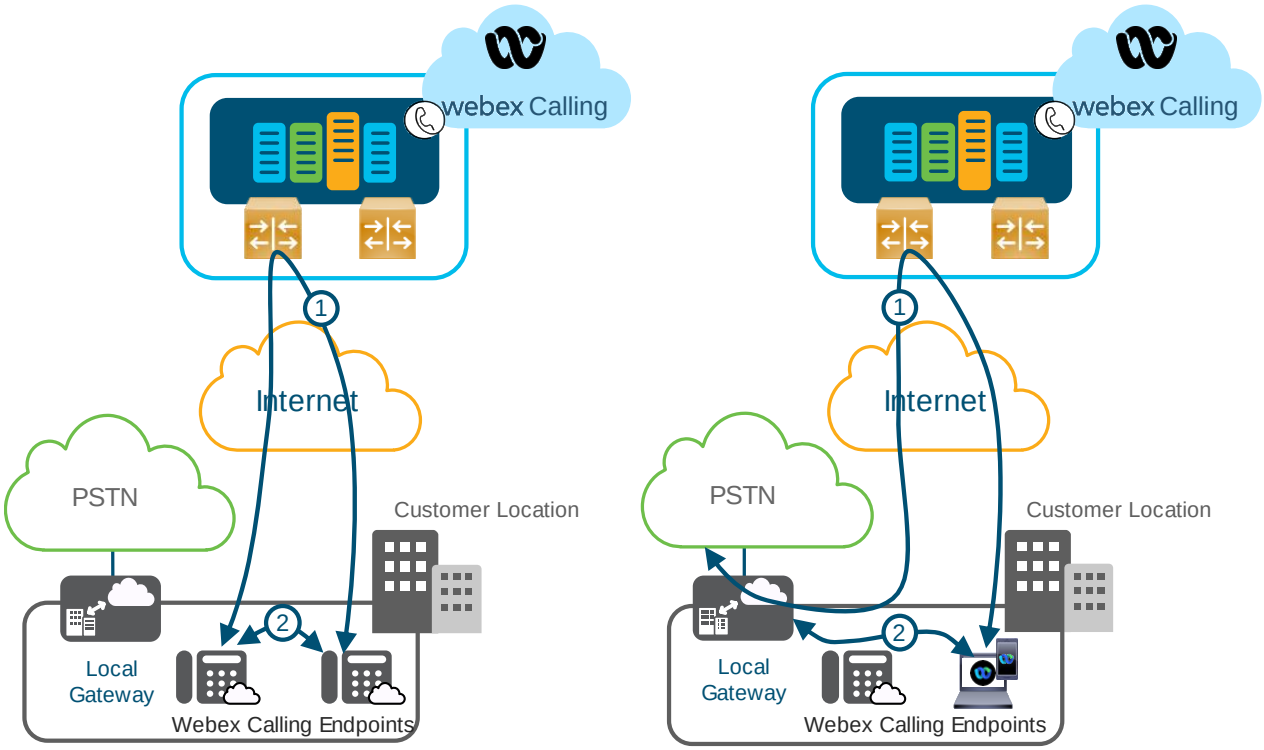
This is the reason why connectivity checks are performed multiple times; if the very first packet is discarded by the firewall, subsequent packets might be allowed due to firewall pinholes being opened outbound in each direction.

As result of the connectivity checks both endpoints now have a list of one or more validated candidates. To select the best candidate for media path optimization the identified candidates are prioritized based on the candidate type (host, peer reflexive, server reflexive in that order) and the best pair is negotiated.

Media Path Optimization

Immediately after the call setup succeeds with a SIP 200 OK message media starts to flow. This happens before the connectivity checks are initiated, and therefore the media takes the standard path through Webex Calling. Figure 29 shows this initial media path in step 1, for a phone-to-phone and phone-to-Local Gateway scenario.

Figure 29 Media path change following a successful ICE negotiation



After the STUN connectivity checks are proven to be successful, media must be re-negotiated via a new signaling phase. The ICE controlling agent – in many cases it corresponds to the calling endpoint – sends a re-INVITE with the selected candidate include in the SDP body of the SIP message. When Webex Calling receives the new INVITE, by checking the SDP concludes that ICE negotiation has been successful and does not add its own transport address to the SDP and instead just relays the SDP with the candidate received from the endpoint. Equivalently the other endpoint will return its selected candidate as part of the SDP media re-negotiation. Consequently, both endpoints switch to using the negotiated host, server reflexive, or peer reflexive transport addresses so that media between the endpoints will not be anchored on the Webex Calling access anymore. The media re-negotiation is shown in step 2.

If ICE negotiation is not successful, the media will be relayed by the Webex Calling network, as shown in step 1, without moving to step 2.

Local Gateway ICE Lite implementation

The Local Gateway supports a “lite” implementation, defined as an option in the ICE specification.

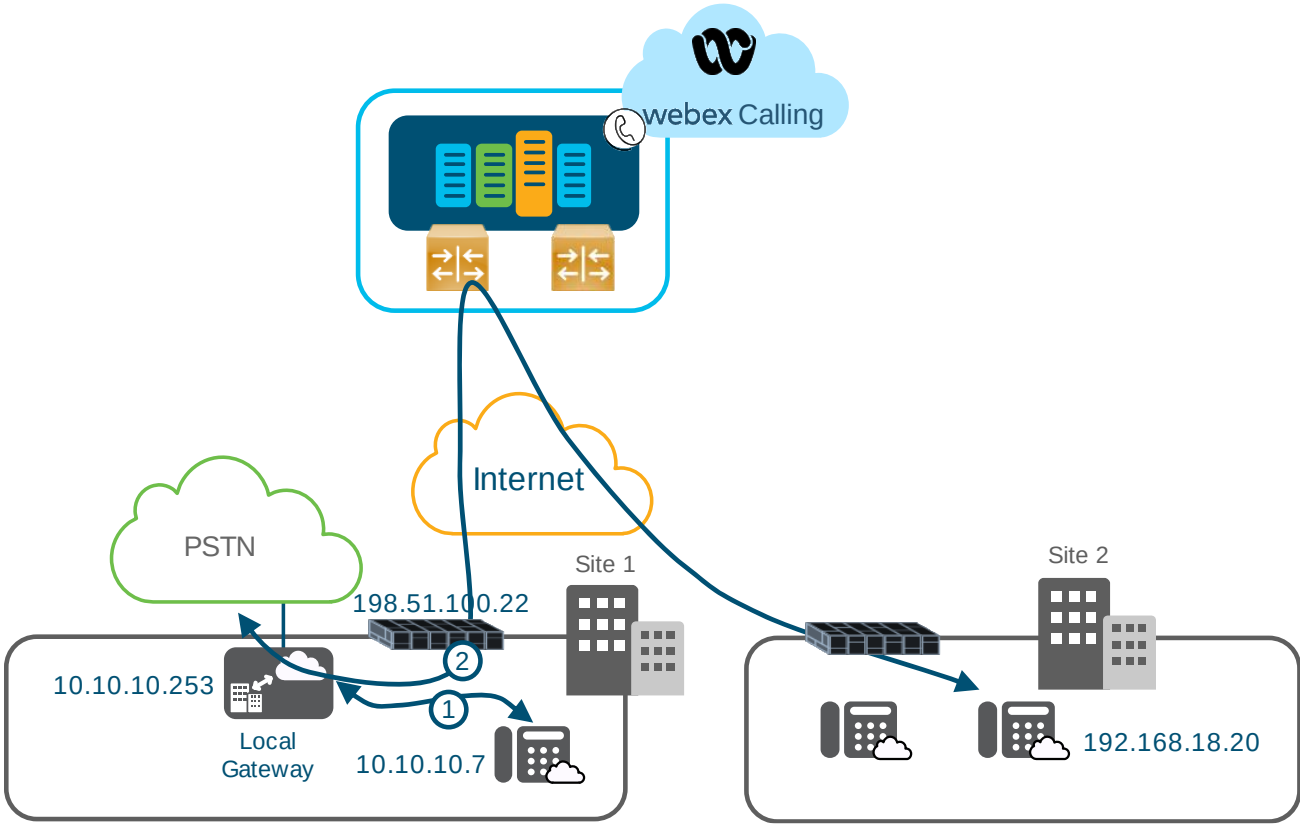
A device supporting ICE lite has the following major differences from a full ICE device:

- It does not collect candidates through STUN procedures. Therefore, only host candidates are advertised to the remote party
- It does not initiate connectivity checks, but it responds to connectivity checks
- In case the system implementing ICE lite has multiple interfaces, only one interface at most can be advertised in the SDP

The implications of ICE-lite implementation on the Local Gateway are the following:

- Because the Local Gateway with lite implementation does not collect server reflexive candidates, it is not possible to access the Local Gateway from a disjointed network, separated by NAT. This is shown in Figure 30.

Figure 30 Media Path Optimization with a Local Gateway



This illustration shows an endpoint with IP 10.10.10.7 in Site 1 with media sent direct to the Local Gateway. In this case media optimization is possible because the Local Gateway local transport address is not translated by NAT in the path between the endpoint and the Local Gateway. If instead the Local Gateway is only visible through its NAT translated address, such as in the case of an endpoint in Site 2, the media will flow through the Webex Calling network.

The Local Gateway has been designed in order not to require any inbound port opening in the firewall. For this reason, it can be deployed in the internal network. When it is deployed in the DMZ with dual interface, and ICE lite is configured in the outbound dial-peer towards Webex Calling, the Local Gateway includes the host candidates of the external interface. If the security administrators do not allow the internal endpoints to access the external interface of the Local Gateway, then PSTN media will be sent through Webex Calling to the Local Gateway.

ICE media path optimization on CUBE only works for IP-to-IP call legs. IP-to-PSTN call legs (between Webex Calling and a PSTN interface on CUBE) do not benefit from ICE media path optimization.

Intrusion Protection System Requirements

Because both signaling and media are encrypted, Cisco recommends that both of them are allowed to flow transparently and uninspected to Local Gateways and IP phones.

IPS inspection can be performed after the traffic reaches the Local Gateway and before it is sent without any manipulation to the Unified CM or endpoint. IPS should be positioned between the Local Gateway and the UCM. For inspection to work, it is required that traffic between Local Gateway and Cisco Unified CM or endpoint is sent in the clear. In this case, the Local Gateway will decrypt the traffic from Webex Calling before sending SIP signaling to Unified CM and RTP media to the destination endpoint or gateway.

If the IPS performs actions that prevent normal operation of the whole system, for example, it identifies legitimate traffic as malicious, the IPS inspection for traffic between the Local Gateway and Unified CM or endpoints should be disabled.

Codec Selection

Webex Calling optimizes audio call quality using the Opus codec, which is supported by most clients on the platform. Opus is supported by the Webex App as well as the MPP phones. It is also supported by the Local Gateway as an end-to-end codec without the ability to do audio codec transcoding, that is Opus to other codecs such as G.711 is not possible. However, it is currently not supported by analog telephone adapters (ATAs) and DECT phones. It is also not supported by most PSTN providers, and hence, G711 is the recommended option for PSTN and UCM interconnect calls.

To ensure a consistent call quality experience and codec negotiation, it is recommended to configure only G.711 codec on the Local Gateway to ensure all UCM and PSTN interconnect calls with Webex Calling select G.711 as the codec. If the IP PSTN provider only supports G.729 and Local Gateway is being used to provide PSTN interconnect for a Webex Calling deployment, then it is recommended to add G.729 as well along with G.711 in the Local Gateway's voice class codec configuration.

Webex App and Cisco phones 8800 series optimize the quality by using H.264 codec for video. The H.264 codec is supported by the Webex App and Cisco phone 8800 series for optimized video call quality.

For all other call flows, the Opus codec is supported as shown below:

- Webex App (desktop) ←→ Webex App (desktop)
- Multiplatform Phone ←→ Multiplatform Phone
- Multiplatform Phone ←→ Webex App (desktop)
- Multiplatform Phone ←→ Auto Attendant
- Multiplatform Phone ←→ Voicemail
- Webex App (desktop) ←→ Auto Attendant
- Webex App (desktop) ←→ Voicemail

Bandwidth Considerations

To determine the bandwidth required on the internet access for Webex Calling, the number of concurrent call legs and the codec used for each call leg needs to be considered.

Table 7 shows the call types available with a Webex Calling deployment along with the codec and maximum bandwidth required for each call type. The required audio call bandwidth for each call type can be calculated using the following general formula:

*Number of expected concurrent calls * Bandwidth per call based on codec = Total bandwidth.*

Table 7 Webex Calling Call Type Bandwidth Calculations

Call Types	Codec - Bandwidth	Total bandwidth
Webex App / MPP Phone → Webex App	Opus - 40 kbps	Number of concurrent calls * 40 kbps
Webex App / MPP Phone → MPP Phone	Opus – 40 kbps	Number of concurrent calls * 40 kbps
Webex App / MPP Phone → PSTN via LGW	G.711 – 80 kbps	Number of concurrent calls * 80 kbps
Webex App / MPP Phone → PSTN via CCP	G.711 – 80 kbps	Number of concurrent calls * 80 kbps
Webex App / MPP Phone → Enterprise via LGW	G.722 – 80 kbps	Number of concurrent calls * 80 kbps
Webex App / MPP Phone → Webex Calling Voicemail	Opus – 40 kbps	Number of concurrent calls * 40 kbps

By summing the concurrent required network throughput per call type, the total potential bandwidth requirement for a specific site can be determined.

All call legs are always anchored on the Webex Calling access SBCs for the signaling portion, while the media could be direct if ICE is successful. To determine the required internet bandwidth for any given Webex Calling location not only the inter-location calls need to be considered, but also intra-location calls and calls to and from a Local Gateway at that location. For example, an intra-site call between two MPPs would need 2 x 40 kbps full duplex on the location's internet access.

By summing the concurrent required network throughput per call type, the total potential bandwidth requirement for a specific site can be determined.

Table 8 shows an example of a complete bandwidth calculation exercise assuming that all devices are in the same location.

Table 8 Webex Calling Bandwidth Calculation Example

Call Types	Number of Concurrent Calls	Total Bandwidth
Webex App / MPP Phone → Webex App	15	2 * 15 * 40 kbps = 1,200 kbps
Webex App / MPP Phone → MPP Phone	15	2 * 15 * 40 kbps = 1,200 kbps
Webex App / MPP Phone → PSTN via Local Gateway	50	2 * 50 * 80 kbps = 8,000 kbps
Webex App / MPP Phone → PSTN via Cloud Connected PSTN	0	0 * 80 Kbps
Webex App / MPP Phone → Enterprise via Local Gateway	15	2 * 15 * 80 kbps = 2,400 kbps
Webex App / MPP Phone → Webex Calling Voicemail	5	5 * 40 kbps = 200 kbps
TOTAL CALLS / BANDWIDTH	**100 calls**	**12,000 kbps / 12 Mbps**

For a video call bandwidth, the Webex App and the 8800 MPP phones support H.264 video with a maximum resolution of 720p with an average bandwidth usage of 1,500 kbps per call. However, the amount of bandwidth consumed at any point during the call will fluctuate based on the variable bit rate inherent in video communications.

Table 9 Webex Calling Video Call Type Bandwidth Calculations

Video Call Type	Codec - Bandwidth	Total bandwidth
Webex App / MPP Phone → Webex App / MPP Phone	H.264 - 1,500 kbps	Number of concurrent calls * 1,500 kbps

Table 10 Webex Calling Video Bandwidth Calculation in the same location

Video Call Type	Number of Concurrent Video Calls	Total Bandwidth
Webex App / MPP Phone → Webex App / MPP Phone	10	2 * 10 * 1,500 kbps = 30,000 kbps
TOTAL CALLS / BANDWIDTH	**10 calls**	**30,000 kbps / 30 Mbps**

Note: The bandwidth calculations above are based on net codec bitrate. When sizing the internet access signaling and L3 overhead need to be considered.

If ICE and ICE Lite are enabled, then the bandwidth might be consumed on the LAN/WAN network only. However it is worth to point out that in the very initial setup of an ICE call the media streams are sent to Webex for few seconds, just before the path optimization happens. So the number of concurrent calls is still important to determine the total bandwidth towards Webex. However, the number of concurrent calls would be calculated over a much shorter period than in the non-ICE case, thus resulting in a reduced number. As an example, if a company determines that in the peak hour there are 100 concurrent calls towards the Internet without ICE for a total of 4 Mbps, then with ICE this number reduces because call optimization happens after few seconds, thus releasing the Internet bandwidth.

Directory Integration

Webex Calling directory is shown as part of the "User" menu in Control Hub. There are several ways to import users into Control Hub:

- Synchronize users from on-premises Active Directory via Directory Connector
- Synchronize users from cloud-based directories such as Okta and Azure AD
- Add users manually in Control Hub or via CSV file
- Add users via People API (this option requires API usage as part of a customized provisioning system)

The Cisco Directory Connector synchronizes directories with the Webex cloud. This allows administrators to maintain user accounts and data in the Active Directory and on-premises changes are automatically replicated to the cloud.

Cloud-based directories such as Azure AD or Okta do not require a Directory Connector. User sync is configured on both Control Hub and the cloud-based directory.

Users can also be manually provisioned by importing a CSV file via the Control Hub interface.

Documentation related to the Cisco Directory Connector is located at https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cloudCollaboration/spark/hybridservices/directoryconnector/cmgt_b_directory-connector-guide-admins.html.

Users can add contacts and search the user directory from the end-user portal at the https://settings.webex.com page.

Dial Plan

When a user is enabled for Webex Calling, they are assigned to a location and receive an extension within the location. The user can optionally, also be assigned a phone number (DID) from a previously defined pool of available phone numbers. The phone number pool is defined in Webex Control Hub. All phone numbers must be allocated via the PSTN provider and routed to the correct customer PSTN trunk or allocated via the cloud PSTN provider (Cisco or Cloud Connected PSTN provider) to establish reachability for these phone numbers.

PSTN Destinations

To dial PSTN destinations, a Webex Calling user can use the common PSTN dialing practice of the country the location is assigned to. For example, from a US location, national PSTN destinations can be dialed as 10 digits, “1” followed by 10 digits, or by dialing 7 digits. When dialing 7 digits Webex Calling automatically prepends the dialed digits with the NPA of the location’s main number. Similarly, international destinations can be dialed as “011” followed by the full E.164 number. Additionally, +E.164 dialing (“+” followed by an E.164 number) can be used by Webex Calling users.

PSTN Access Code

An optional PSTN outbound dial digit can be defined for each location. This is called a “PSTN access code”. A PSTN access code is typically used in enterprise environments to avoid overlaps with other on-net dialing habits. In the US, “9” is commonly used as the PSTN access code.

For example, if “9” is defined as a PSTN access code for a US location, then to dial a national US destination, the user dials “9” followed by 10 digits, “91” followed by 10 digits, or 7 digits (for numbers within the same NPA). An international destination is dialed with “9011” followed by a full E.164 number.

Abbreviated On-net Dialing

For abbreviated on-net dialing, a routing prefix length, an internal routing prefix, and the internal extension length must be configured for each Webex Calling customer.

- Routing prefix length — defines the length for all routing prefixes to be configured for each location
- Internal routing prefix — is the common first digit for all location routing prefixes
- Internal extension length — defines the extension length to be used in each location

With these three settings, a universal abbreviated inter-site on-net dialing habit is defined in the form: <internal routing prefix>-<site code>-extension.

If for example, an internal routing prefix of “8”, a routing prefix length of four and an extension length of four are configured then all on-net destinations can be dialed in the form “8-XXX-XXXX”. Here the leading “8” followed by three digits is the location’s prefix and the last four digits are the extension defined within the location.

Note that the configured internal routing prefix length includes the leading routing prefix. To use three-digit site codes, the routing prefix length must be set to four.

If the routing prefix is set to “8” and the routing prefix length is set to four, always include the leading “8” in the routing prefix when defining location routing prefixes

When configuring the location routing prefix, a warning; “Enter and save a routing prefix that aligns with global Call Settings steering digit 8 to reduce delays in dialing.” displays to remind the administrator of this requirement.

Table 11 Abbreviated on-net, inter-site dialing with a steering digit of 8 and routing prefix length and extension length both set to four

Site	Site Code	Location Prefix	Extension Range	Abbreviated On-Net Dialing
SJC	140	8140	4000-4999	81404XXX
NYC	121	8121	4000-4999	81214XXX
RTP	191	8191	1000-1999	81911XXX

Table 11 shows three example sites configured with a site code, an internal routing prefix of “8”, a routing prefix length of four and an extension length of four. It also shows the resulting, abbreviated on-net, inter-site dialing to reach each site’s destinations.

Even though SJC and NYC use the same extension range, the abbreviated on-net, inter-site dialing habit to reach destinations in these sites is still unique. The site codes are all three digits long and together with the steering digit “8”, result in four-digit location prefixes.

Defining the above fixed-length structure for abbreviated on-net, inter-site dialing enables Webex Calling to push a dial plan to the Webex Calling devices. When off-hook, the phone recognizes this abbreviated on-net, inter-site dialing habit and immediately stop collecting digits and sends the dialed digits to Webex Calling for analysis and call routing.

This fixed-length number structure for abbreviated on-net, inter-site dialing helps avoid inter-digit timeout and improves the overall user experience.

To enable calling from Unified CM to on-premises Unified CM and correct call type classification at least one dial plan needs to be configured in Webex Calling and for all enterprise and +E.164 number ranges homed on Unified CM the respective +E.164 and ESN patterns need to be configured in that dial plan. This not only guarantees correct routing of Unified CM destinations to Unified CM but also makes sure that for a call received by Webex Calling from a trunk that call is correctly classified as premises or PSTN call based on a caller ID match.

Integrated Audio

Integrated audio for Webex Calling allows organizations with both a Webex Meetings and Webex Calling subscription to take advantage of an optimized call routing path to Webex Meetings audio. When using the Call-in or Callback options to join a Webex meeting, calls remain within Webex, saving on organization costs by reducing the number of calls that route via PSTN. For design details please refer to the Preferred architecture for Webex Edge Audio https://www.cisco.com/c/dam/en/us/td/docs/solutions/PA/WbxEdge/PAEdgeAudio.pdf.

Service Assurance

Service Assurance refers to the suite of tools that help customers, partners, and Cisco to successfully deploy and manage Webex Calling deployments.

CScan can be used for troubleshooting initial setups or for monitoring the health of Webex Calling calls.

CScan

CScan is the network readiness tool designed for Webex Calling. Customers or partners can use CScan to test their network connection via Internet.

To test the connection, just go to cscan.webex.com and select the location (one of the Webex Calling data centers) that the user resides in (closest to the user). Users can do a basic test which would test the internet connectivity (provide information on latency, download and upload speeds, and the ports that are needed for the Webex Calling). The advanced diagnostic option provides additional details on the QoS parameters such as Jitter, packet loss and latency.

For more information on using CScan, refer to the Webex help page at https://help.webex.com/en-us/y27bej/Use-CScan-to-Test-Webex-Calling-Network-Quality.

Analytics

Analytics in Webex Control Hub

Administrators can use the Analytics page in Webex Control Hub to see up to 13 months of historical Webex Calling data. Administrators also have access to 13 months of data for calls based in the Webex app if the organization has Pro Pack. If the organization doesn't have Pro Pack, administrators have access to 3 months of data for calls based in the Webex app.

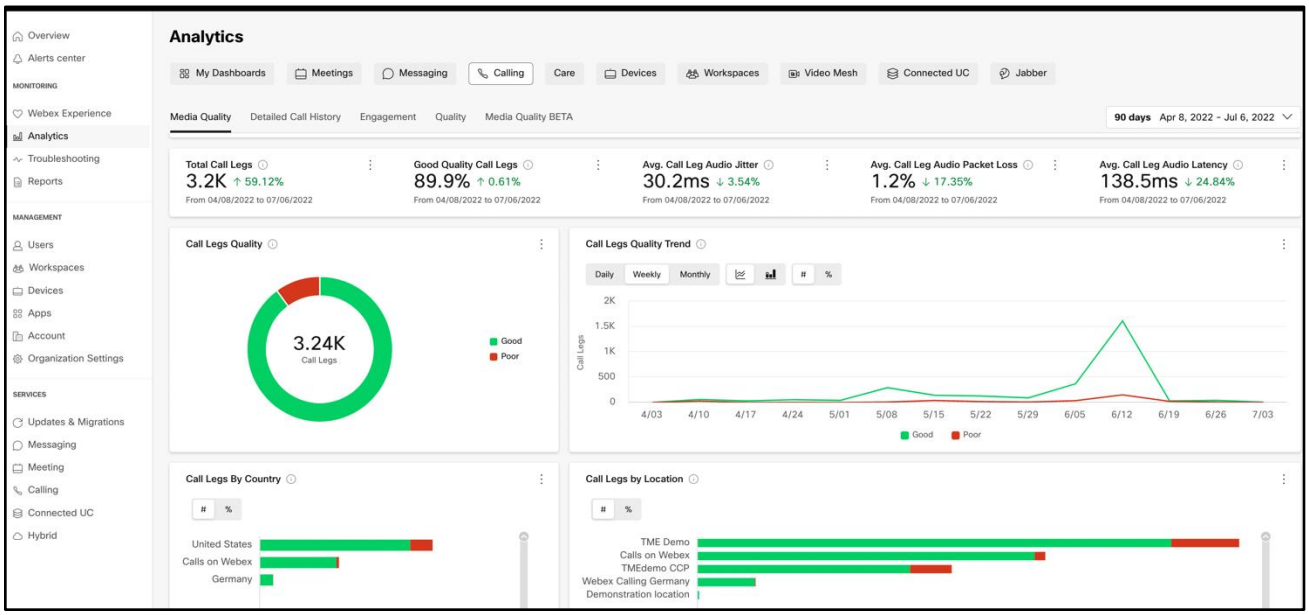
Cisco keeps historical data for calls involving Webex Calling desk phones, Webex App apps (desktop and mobile), and the Webex Calling App (desktop and mobile).

Media Quality Analytics

Media Quality dashboard in Control Hub makes it easy to manage Webex Calling and Call on Webex call quality across your organization. High level key performance indicators (KPIs) give administrators a quick view of global call quality. The charts provide detailed views of this data by location, IP address, media type, connection type, codec, endpoint type, and IP phone model.

Data is also now updated near real-time. Administrators can see call quality data within 15 minutes of when a call ends.

Figure 31 Webex Calling Media Quality Analytics

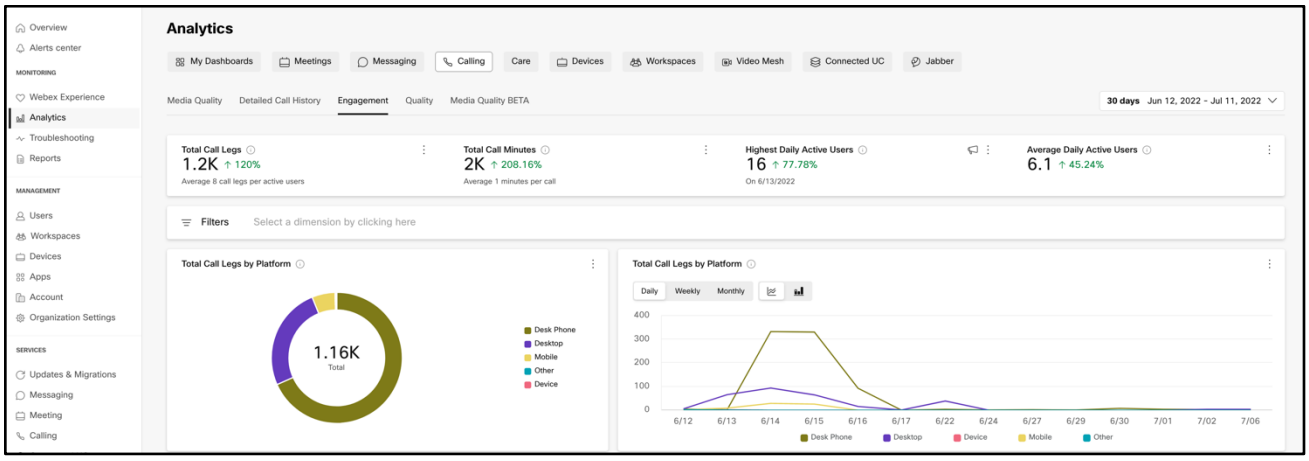


Administrators can also get details on Quality Metrics and Engagement Analytics for Webex Calling calls. They can view historical data of call usage and engagement including media quality records. The Webex Calling analytics are available in Control Hub under “Analytics and Calling”.

Engagement Analytics

The Engagement Analytics provides graphical information on Calls and Call Minutes, detailing all the point-to-point calls made within the organization.

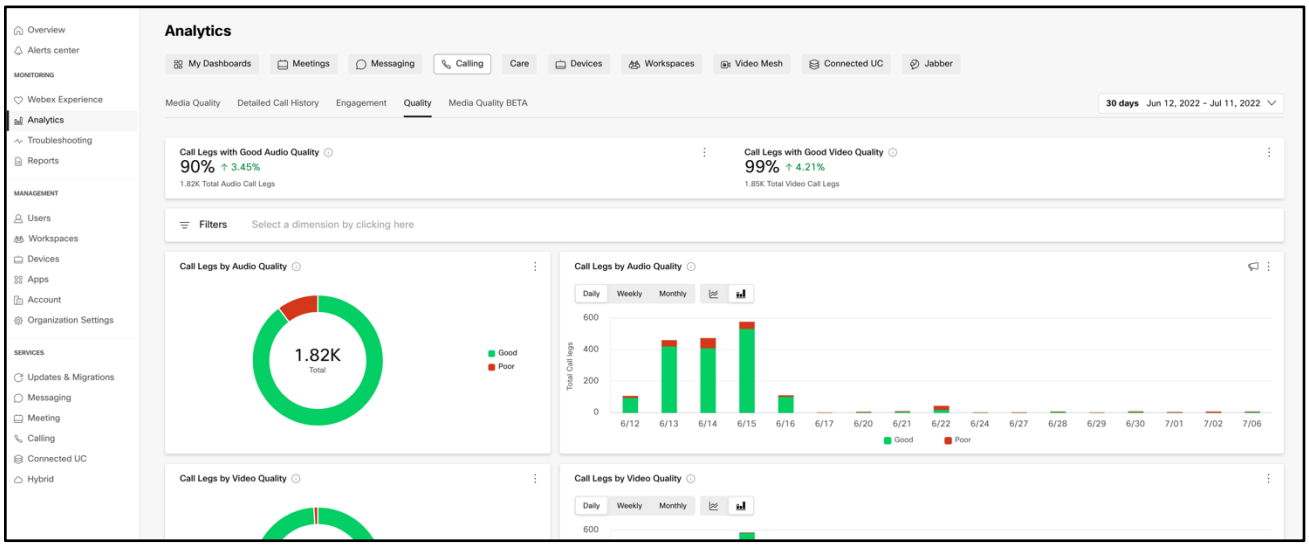
Figure 32 Webex Calling Engagement Analytics



Quality Analytics

The Quality Analytics tab allows the administrator to view records for each call and use sliders to filter calls based on quality statistics.

Figure 33 Webex Calling Quality Analytics



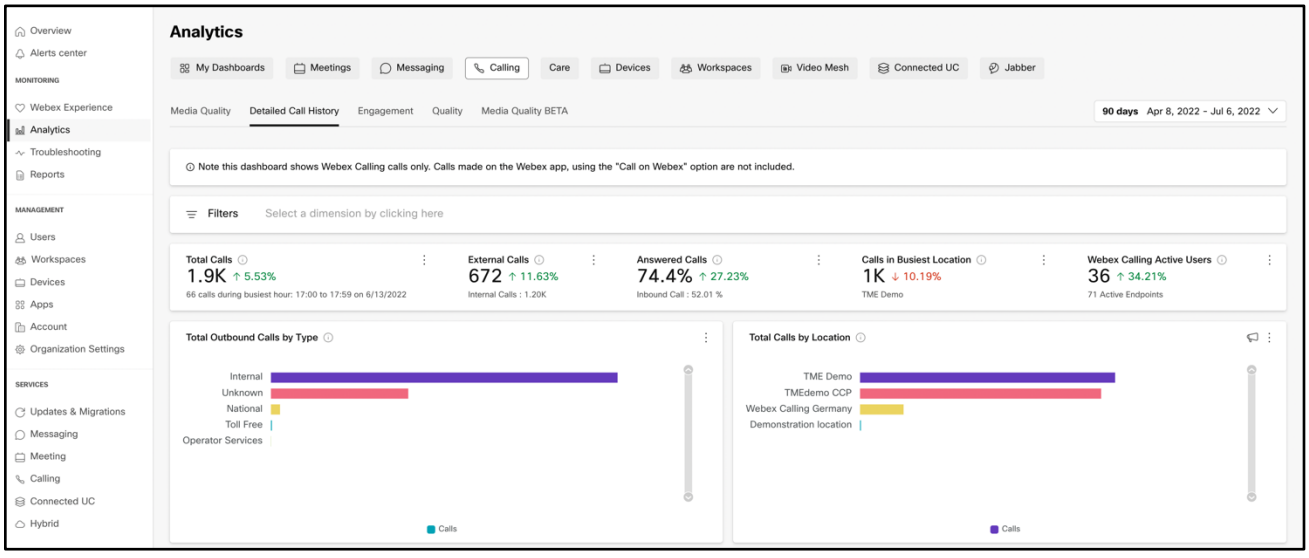
All statistics are directly collected from the devices/endpoints so that the QoS parameters (packet loss, jitter, latency) reflect the experience of the call from the end user's perspective.

Detailed Call History

Administrators can get the Call Detail Records by choosing the "Detailed call history" under Analytics tab in Webex Control Hub

Call Detail Records (CDRs) of up to 12 months from the current date is available for administrators to determine the calling behavior patterns for users.

Figure 34 Webex Calling Detailed Call History



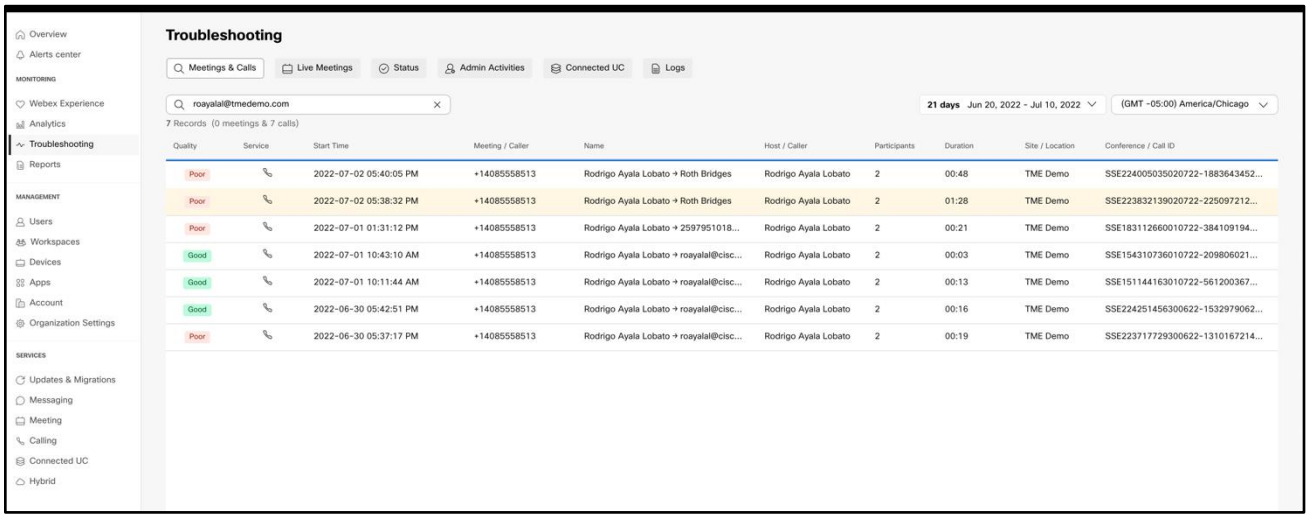
Troubleshooting

The troubleshooting view in Webex Calling from Control Hub allows administrators to troubleshoot media quality issue in a Webex call. Administrators can search for information related to the call, view its media statistics, identify where the issue occurred and resolve the problem.

Administrators can search using the following criteria to get a list of calls where a media session was utilized with at least one Webex Calling registered endpoint:

- Email IDs
- Phone numbers (exact string match)
- MAC address
- Call IDs

Figure 35 Webex Calling Troubleshooting



The media quality troubleshooting allows administrators to:

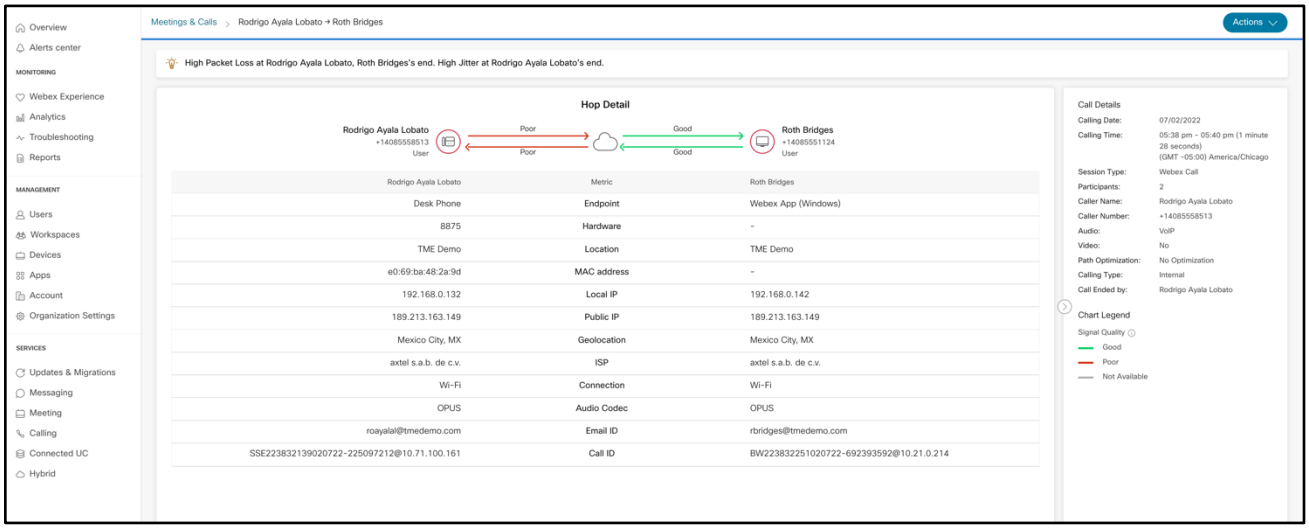
- View the end-to-end experience of the participants of the call.
- View a hop detail of the call.
- View if the media traverses through the Webex Calling cloud, or directly between the users (using Interactive Connectivity Establishment (ICE)).
- View calls for the past 21 days.
- Analyze the call quality metrics that impacted the experience of the user. For example, an administrator may observe high jitter on clients that are connected to Wi-Fi networks, but packet loss and latency may be acceptable.
- Detect if the issue is with the caller or the callee.

Supported Call Flows for Troubleshooting from Control Hub

The media quality reports are collected from the caller and callee endpoints and the media relay points. This allows a segmentation of the media experience to narrow down and identify whether the issue occurred at the:

- Caller or callee
- Media path to or from the Webex Calling cloud

Figure 36 Webex Calling Troubleshooting Hop Details



For more information on using Troubleshooting, refer to the Webex help page https://help.webex.com/en-us/article/frj1efb/Troubleshoot-Webex-Calling-Media-Quality-in-Control-Hub.

Webex Calling Dedicated Instance

Dedicated Instance is part of Cisco's Cloud Calling portfolio, powered by Cisco's collaboration technology—Webex Calling and Cisco Unified Communications Manager (Cisco UCM) based architecture. The service offers voice, video, messaging, voicemail, meeting, and mobility solutions with the features and benefits of Cisco IP phones, mobile devices, and desktop clients. Dedicated Instance is integrated with Webex Calling so that customers can take advantage of Webex platform experience along with preserving UC Manager experience. Customers with need for customised cloud-based enterprise grade features can deploy Dedicated Instance and provide rich Webex Calling features to their install base.

Dedicated Instance is hosted and operated by Cisco in North America, Europe, Asia Pacific and Australia regions in Webex data centers. Cisco Webex data centers are globally distributed and geo-redundant.

Dedicated Instance UC Applications

Dedicated Instance includes the following set of applications:

- Cisco Unified Communications Manager
- Cisco Unified Communications Manager IM & Presence Service
- Cisco Unity Connection
- Cisco Expressway
- Cisco Emergency Responder (AMER only)

Dedicated Instance is deployed with one cluster for all UC applications and can grow, as required, to accommodate additional users that have been purchased on the Webex Calling Subscription. Customers have flexibility to request additional clusters if additional capacity is required. This is done by opening a Service Request (SR).

Though Dedicated Instance applications are cloud-based solutions, configuration guidelines and best practices match the ones that have been deployed for on-premises scenarios. For best practices on UC Applications, refer to the collaboration Enterprise CVD at https://www.cisco.com/c/en/us/td/docs/solutions/CVD/Collaboration/enterprise/14/collbcvd/control.html.

Cisco creates the SIP trunk between multi-tenant and dedicated instance. Partners/Customers can bring in their existing dial plans and connect to the platform for internal dialing.

Cisco preconfigures trunks on Webex Calling multi-tenant and dedicated instance deployments. Partners and customers can plug in their existing dial plan with minimal configurations.

Multi-tenant and dedicated instance applications talk through gateways deployed in the cloud.

The following configuration is done (by Cisco) on both multi-tenant and dedicated instance deployments.

Multi-tenant:

- Location for interop
- Two redundant preconfigured SIP trunks to the Gateways in Regional DCs and register to Multi-tenant
- Route groups preconfigured to include the 2 trunks

Dedicated Instance:

- Provision Gateways in each regional Datacenter
- Preconfigured SIP trunks on Unified CM to Gateways
- RGs (route groups) including the 2 trunks to Gateways
- RL (route-list) to the respective RG

Note: The partner/customer should not delete the preconfigured trunks on either multi-tenant or dedicated instance deployments because this would affect the internal calling behavior. For more details on the base configuration for

dedicated instance, refer to https://help.webex.com/en-us/article/2vpf1/Dedicated-Instance-for-Webex-Calling---Base-Configuration.

Dedicated Instance Peering (Connectivity)

Peering is the interconnection between the Dedicated Instance network and a partner's or customer's network that enables communication with Dedicated Instance collaboration services. Customers onboard with either a partner or directly to Cisco with a direct connect option. It is a mandatory process performed once per geographical area.

Peering establishes:

- Partner management and operations network
- Partner's customer aggregation framework
- Customer access to their respective private Dedicated Instance
- Partner's supplemental services framework
- Cloud-based software services and managed services to the customer

After peering links are established, the partner can logically partition the traffic for one or more of their customers, their management network, and Internet access for over the top (OTT) clients or MRA clients.

After peering, the partner has access to the following services in the Dedicated Instance:

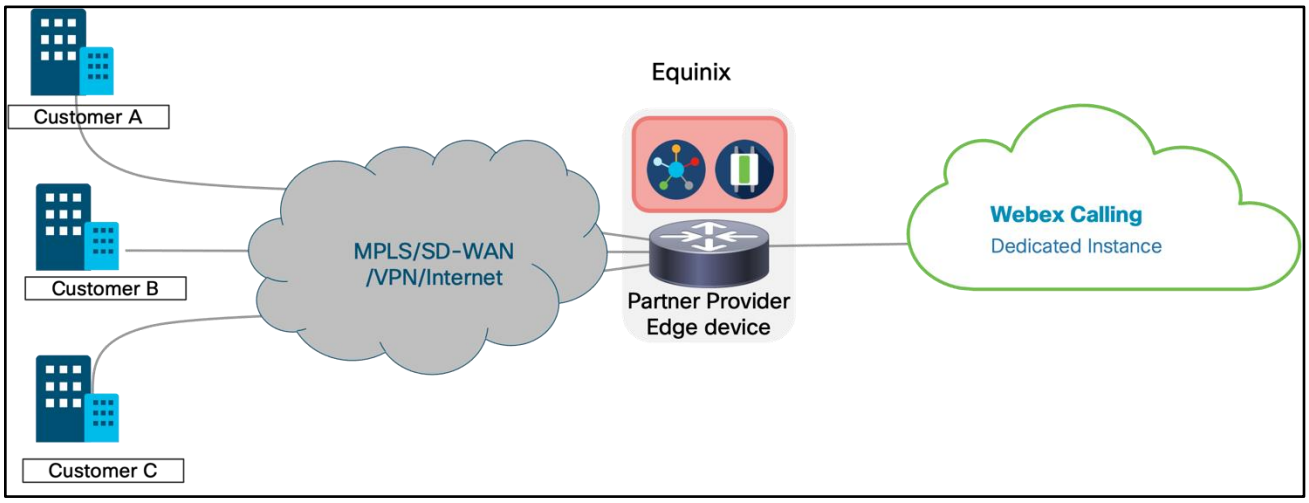
- Dedicated Instance application management interfaces
- Operational tools – Assurance portal, RTMT
- Provisioning tools (optional)
- Reports
- API access

Partner Connected Peering

In the Partner Connected peering model, the partner aggregates and connects their customers' WAN to Dedicated Instance in each geographic region where they offer service.

The Partner Connected model requires Equinix presence and Direct connection in each physical location with 1 GB or 10 GB single-mode fiber. Partner equipment must support Layer 3 sub-interfaces with dot1q tagging for separate VRFs. Partners can deploy one or two devices (a layer 3 WAN router for example) per location. The partner aggregates all connections provided (customer, management, Internet, PSTN). The partner also defines the customer network access options (MPLS, SD-WAN, VPN, or Internet).

Figure 37 Dedicated Instance Partner Connect



Webex Edge Connect

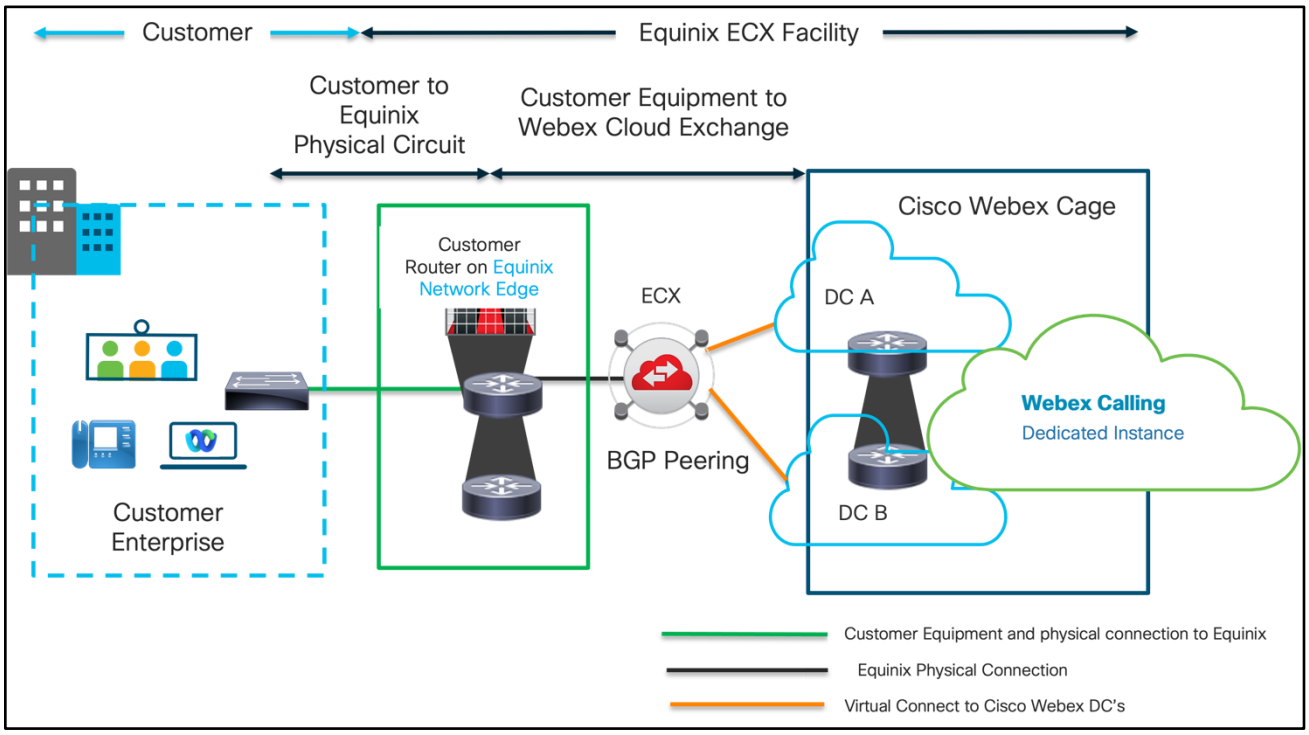
Webex Edge Connect peering is a private point-to-point link that bypasses the public Internet and connects your network to a Cisco Webex data center location via Equinix. Customers connect dedicated, managed, redundant IP links from the customer WAN to Dedicated Instance (calling workload only). Customers connect their WAN to the Webex backbone via the Equinix cloud exchange. Webex Edge Connect delivers direct connectivity and dedicated high-speed bandwidth to Webex cloud services.

Peering is done via Equinix Cloud Exchange (ECX) fabric location with bandwidth options ranging from 200 MB to 10 GB.

Equinix handles network flow from the peering point to the Dedicated Instance data center. Equinix charges for the physical connection and each virtual connection.

For more information about Webex Edge Connect, refer to the Webex Edge Connect Preferred Architecture document located at https://www.cisco.com/c/dam/en/us/td/docs/solutions/PA/EdgeConnect/PA_Edge_Connect_Design.pdf.

Figure 38 Customer Connected peering: Webex Edge Connect



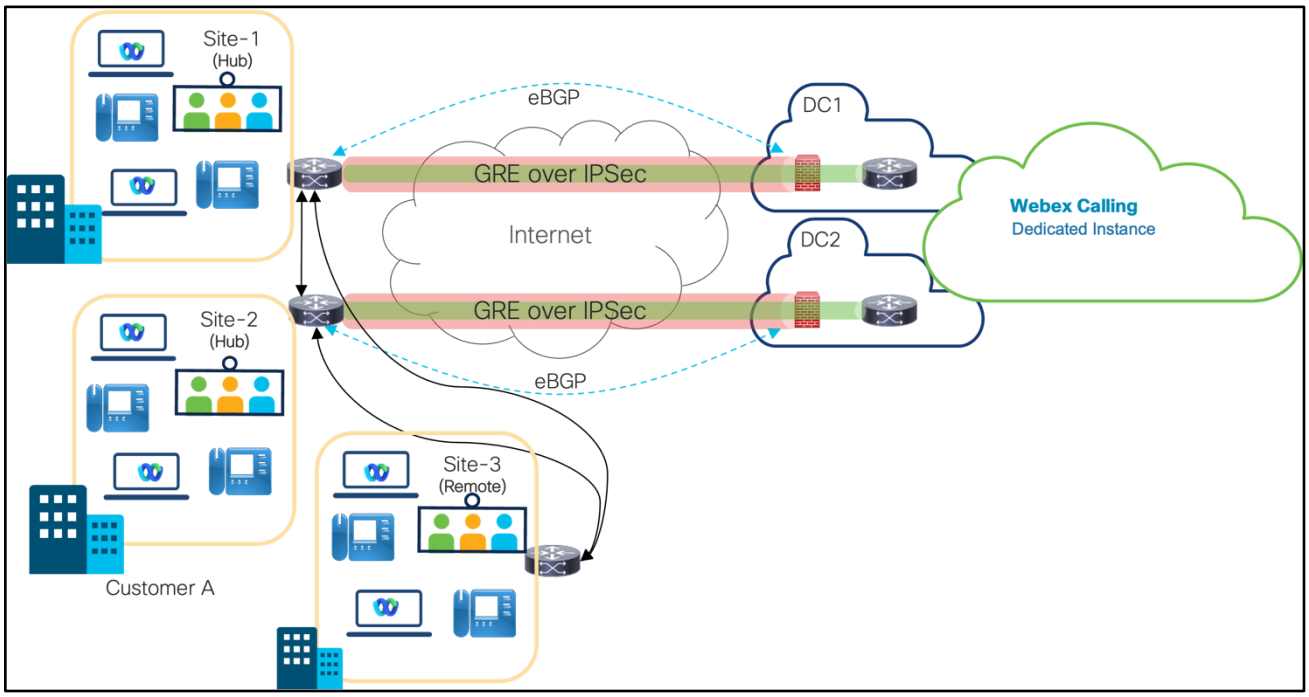
As shown in Figure 38 customer connects to the Cisco Webex backbone via Equinix and ECX Fabric. The customer would have a physical connection from their network to Equinix and then two virtual circuits each per region connecting to the Cisco data centers.

Virtual Connect (VPN)

Virtual Connect is an additional add-on option for Cloud Connectivity to Dedicated Instance. Virtual Connect enables Customers to securely extend their Private Network over the internet using point-to-point IP VPN Tunnels. This connectivity option provides a quick establishment of Private Network connection by using the existing Customer Premise Equipment (CPE) and internet connectivity.

Virtual Connect has a bandwidth limit of 250 Mbps per tunnel and is recommended for smaller deployments. Since two point-to-point VPN tunnels are used all traffic to the cloud has to go through the customer headend CPE, and therefore it may not be suitable where there are a lot of remote sites. However, in most cases where the user is remote, mobile and remote access can be deployed, thus reducing the impact of the bandwidth in the virtual connect circuits.

Figure Customer Connected peering: Virtual Connect



As shown in Figure , two hub sites are recommended for better redundancy, but one Hub site with two tunnels is also a supported deployment model.

For calls between endpoints registered with Dedicated Instance (on-prem-on-prem), the media will flow within the Customer's LAN/WAN.

The media will flow through the tunnel for Dedicated Instance on prem endpoints in following scenarios

- CCP PSTN (Leveraging the CCP PSTN with Webex Calling)
- WxC to Dedicated Instance endpoints calling and vice versa
- Integrated Audio leveraging the Webex meetings On-net SIP trunk

Besides, any call between WxC and Dedicated Instance always flows through the Webex datacenters.

For more information about Virtual Connect, refer to the Dedicated Instance-Virtual Connect document located at https://help.webex.com/en-us/article/v8jhcv/Dedicated-Instance-Virtual-Connect#Cisco_Concept.dita_a94968f5-46a9-4b47-9250-ce51169b33e3.

Dedicated Instance PSTN Options

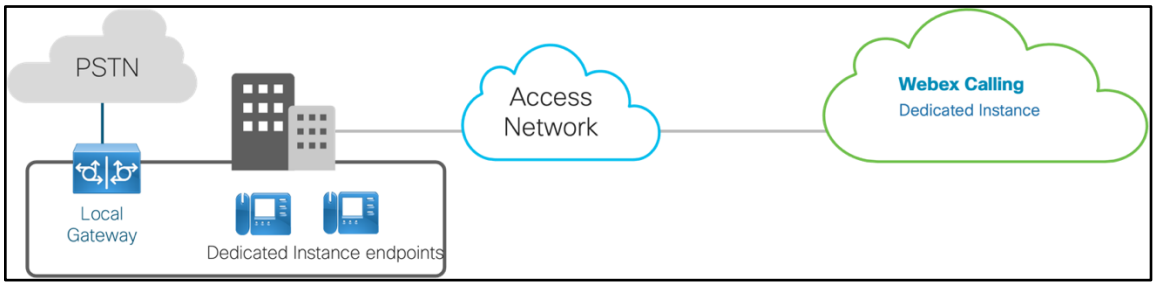
- Dedicated Instance customers can access the PSTN with one of the following options:
- Local Gateway located on Customer premises (Local Breakout)
- Partner Hosted Local Gateway (Bundled PSTN)
- CCPP (Cloud Connected PSTN provider)

Local Gateway (Local Break Out)

The local gateway function is deployed on the customer or partner premises. The local gateway registers with Dedicated Instance as a trunk and routes all calls between the PSTN and Dedicated Instance.

A local gateway connects to the PSTN either directly, by terminating a PSTN trunk (TDM or IP) on the same router, or by connecting to an existing PSTN gateway via SIP trunk. The local gateway function running on a Cisco voice gateway or Cisco Unified Border Element (CUBE) enterprise routes calls to the PSTN.

Figure Dedicated customer local gateway deployment



While combining PSTN access and the connection to Dedicated Instance requires less hardware (with CUBE) to be installed and maintained on the customer's network, implementing both functions on separate devices can be a preferred option in cases where an existing PSTN gateway is used after the migration of a site to Dedicated Instance.

Partner Hosted Local Gateway (Bundled PSTN)

Instead of deploying individual local gateways on each customer's network, a partner can also host a customer's local gateway in their own data center.

Figure 39 Partner hosted local gateway option

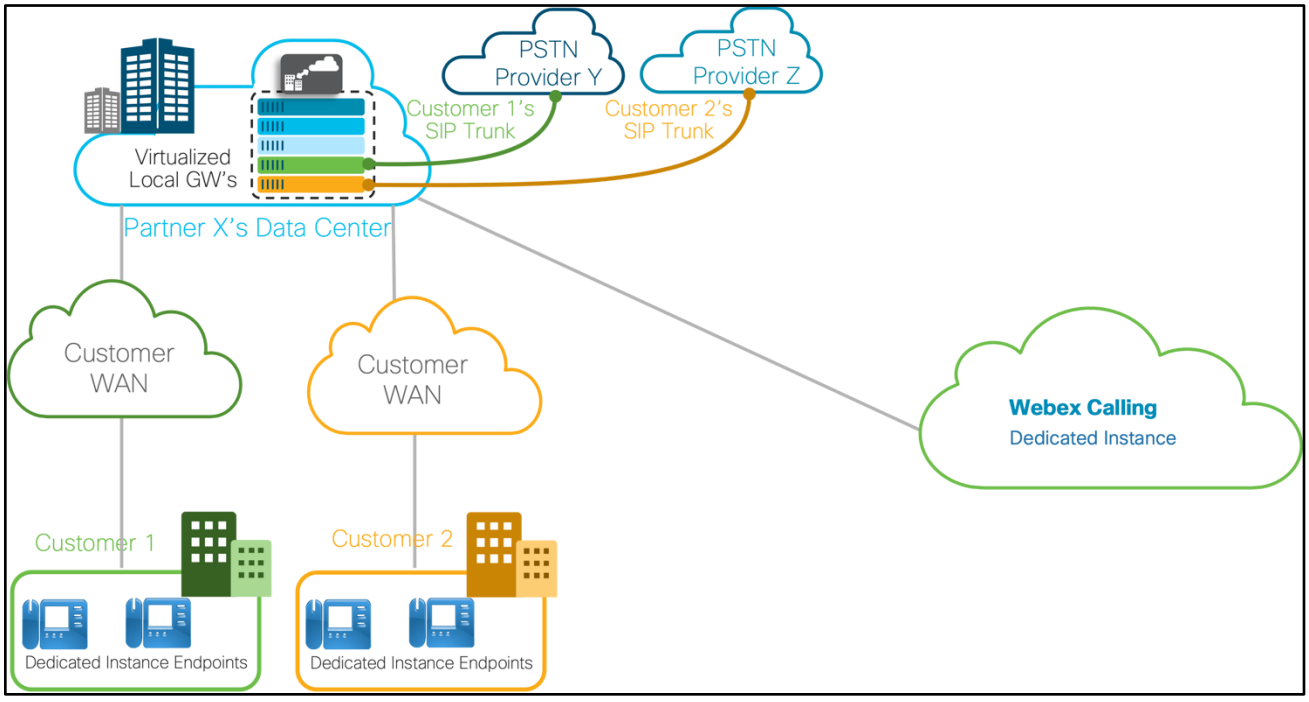
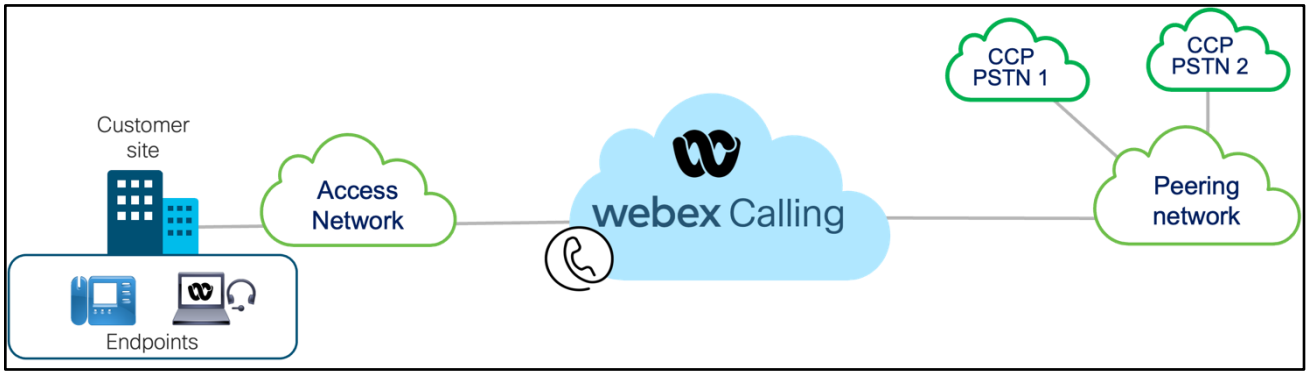


Figure 39 shows that the partner in this example will deploy local gateways in their data center. The dial-peer based call routing configurations of each individual customer can be combined on a single CUBE. Separation of traffic between customers is achieved by the proper dial peer routing configuration. This allows partners to determine the local gateway, map calls received from Dedicated Instance to customer specific PSTN trunks, and vice versa. This deployment model enables the partner to deploy, maintain, and operate local gateways for various customers more efficiently.

CCPP (Cloud Connected PSTN provider)

Cloud Connected PSTN (CCP) enables global cloud PSTN calling options for Dedicated Instance. Dedicated Instance leverages existing CCP partner peering with Webex Calling for this feature. To enable CCP for Dedicated Instance, administrators will have to configure the Route Lists on Webex Control Hub.

Figure 40 Cloud Connected PSTN



As shown in Figure 40, customers can use PSTN services from different CCP providers for their enterprise.

For more information about CCPP refer to the Cloud Connected PSTN document located at https://help.webex.com/en-us/article/nw3ygtq/Cloud-Connected-PSTN

Case Study 1a: Unified CM with Centralized Call Processing and multiple Webex Calling Locations

Adding Webex Calling to existing Unified CM installations provides a solution for when centralized call processing is not possible due to insufficient available WAN bandwidth or other logistical challenges.

This case study examines scenarios where Webex Calling locations are combined with a multi-site Unified CM deployment. This type of deployment is useful in transition scenarios such as moving smaller sites from Unified CM to Webex Calling locations.

Local Gateway is a required component to establish connectivity between Webex Calling and on-premises Unified CM. Using the Webex Calling dial plan routing logic this can be complemented by Cisco PSTN or Cloud Connected PSTN to provide PSTN services for Webex Calling users. If no cloud PSTN is present, then the Local Gateway is used both for premises and PSTN calls.

Although Local Gateway can be deployed stand-alone, this case study focuses on Cisco Unified CM deployment integration.

Figure 41 Stand-alone Local Gateway Deployment

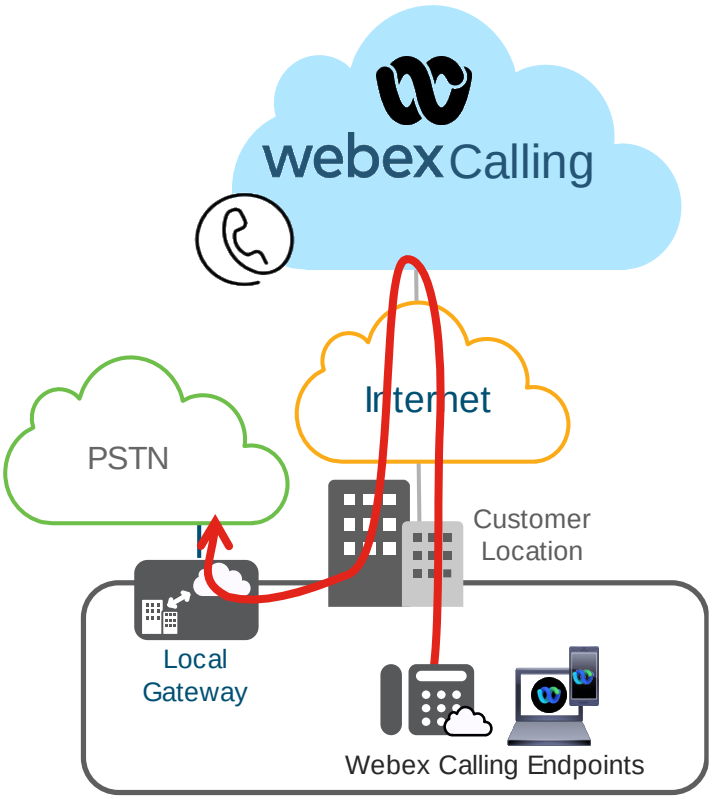


Figure 41 shows a single-location Webex Calling deployment with Local Gateway. PSTN calls originating from Webex Calling endpoints in this case are sent to the Local Gateway which provides access to the PSTN.

If phones registered to Unified CM deployed as on-premises call control require direct dialing to Webex Calling locations, a Local Gateway integration is required.

Figure Local Gateway combined with a dedicated PSTN Gateway

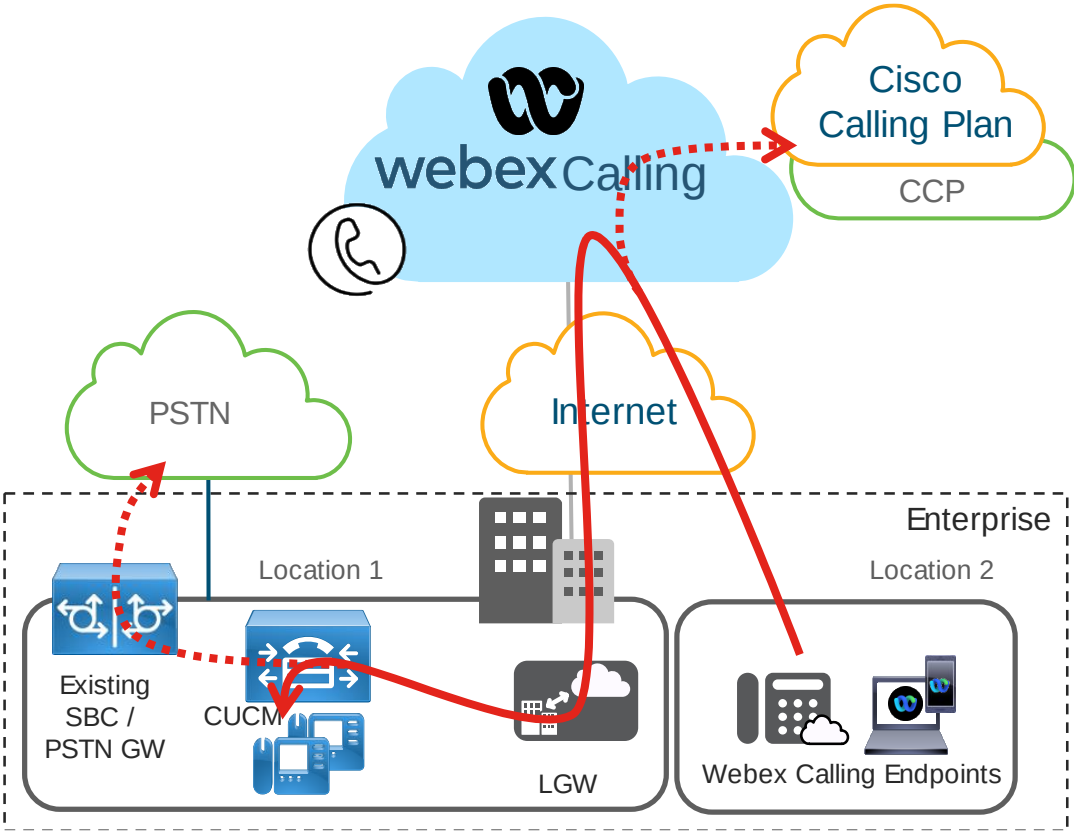


Figure shows that calls originating from Webex Calling endpoints are sent to the Local Gateway based on Webex Calling dial plan routing logic. The Local Gateway then sends the calls on to Unified CM. The enterprise dial plan provisioned on Unified CM determines whether the call needs to be extended to an endpoint registered to Unified CM or to the PSTN via the existing PSTN GW infrastructure. Routing to the PSTN is only an option if the Webex Calling location is configured for premises-based PSTN. If instead cloud PSTN is used, then PSTN calls from Webex Calling users are sent to the cloud PSTN choice configured for the calling user's location.

Figure 42 Local Gateway co-resident with PSTN Gateway

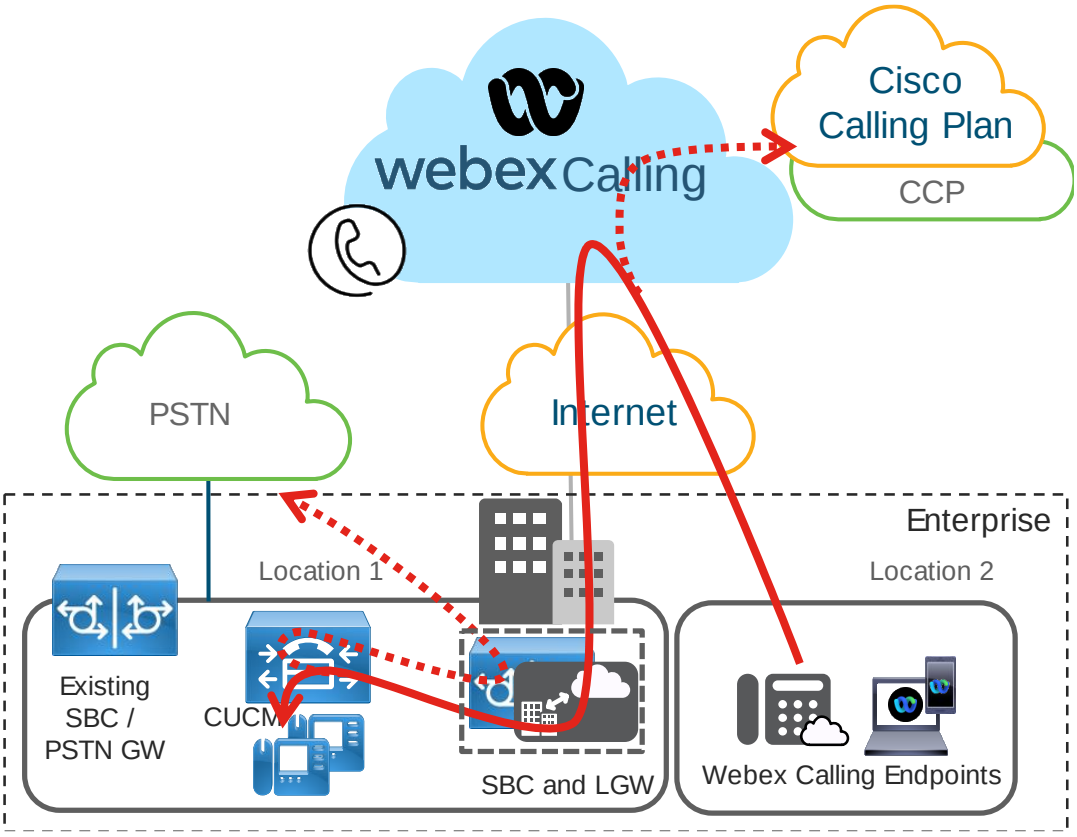


Figure 42 shows a variation of this deployment mode where the Local Gateway and the PSTN access function reside on the same device. The difference in this scenario is the configuration on the combined Local Gateway and PSTN Gateway device. Neither from the perspective of the Webex Calling nor the Unified CM configuration, there is any major difference to the previous scenario. Note that with this scenario, Unified CM receives two types of calls from the CUBE hosting PSTN access and Local Gateway functionality: calls from the PSTN and calls from Webex Calling. To allow for differentiated class of service to be applied to these call types, two SIP trunks should be configured between CUBE and Unified CM: one for each call type. To achieve this, different SIP listening ports must be configured for each SIP trunk on Unified CM in the SIP trunk's security profile.

Differentiation between PSTN calls and Webex Calling calls received from the Local Gateway is also required if the Webex Calling location is configured for cloud PSTN. While no PSTN calls would need to be forwarded back Unified CM from Webex Calling via the Local Gateway and back to the PSTN, Unified CM would still need to be able to apply differentiated class of service: calls originating from Webex calling would need access to remote on-net locations in multi-cluster deployments (for example via a connected Unified CM SME) while calls from the PSTN typically don't need this access.

Combining both functions on the same device allows for more cost-effective deployments.

Call Routing Considerations

Calls from Webex Calling to Unified CM

As described earlier in the architectural overview calls from Webex Calling users are routed to premises trunks based on Webex Calling dial plans. If premises based PSTN is used, then PSTN calls are also sent to the premises trunks. The

distinction between the two call types is up to Unified CM and depends on the enterprise dial plan provisioned on Unified CM.

Figure Dialing from Webex Calling to Unified CM

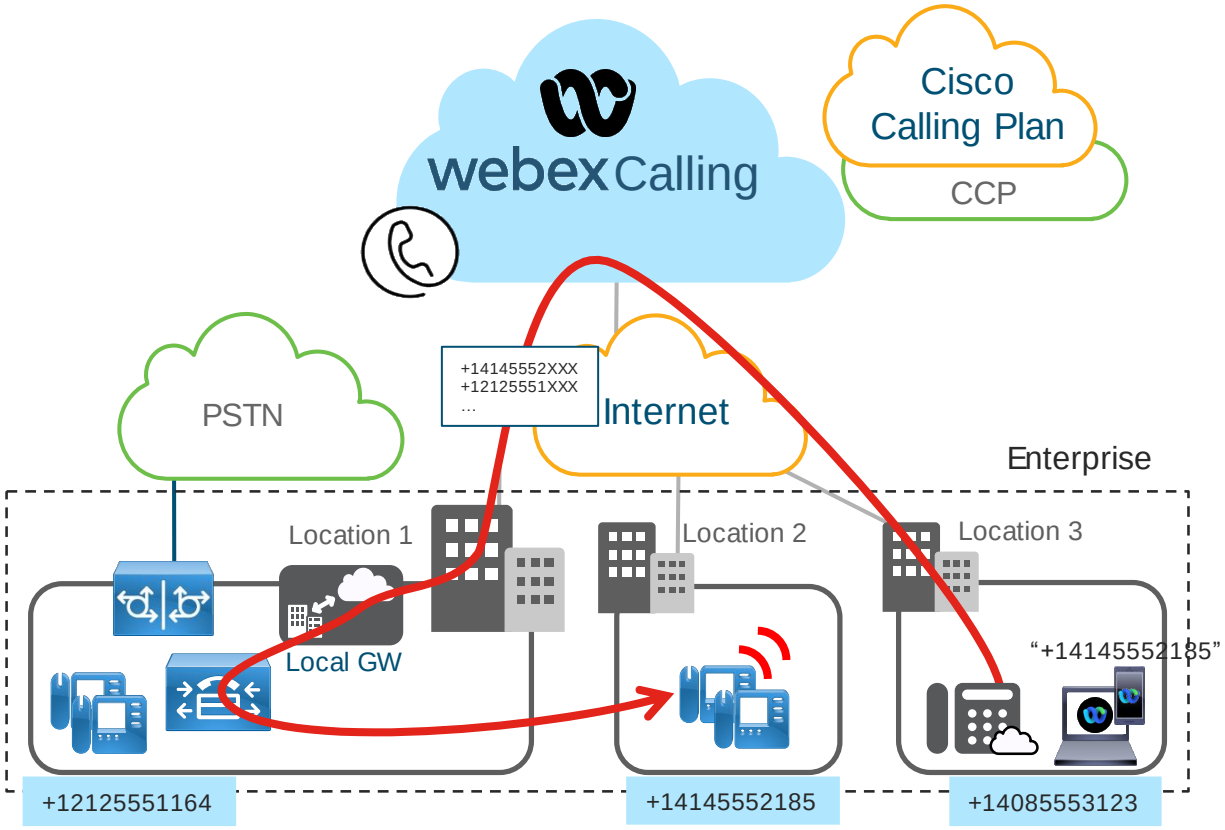


Figure shows an example of a multi-site Unified CM deployment with centralized call processing. A Webex Calling user is dialing a +E.164 number and the dialed number does not match any number provisioned for the customer in Webex Calling but the number matches a +E.164 pattern configured in a Webex dial plan. The call is therefore sent to the trunk or route group selected as the routing choice for that dial plan. The Local Gateway sends the call on to Unified CM. Call routing on the Local Gateway does not take the called address into consideration; routing is solely based on trunk attributes so that any call received from Unified CM is forwarded to Webex Calling and any call received from Webex Calling is forwarded to Unified CM. The called party number seen in the call leg from Webex Calling to Unified CM via the Local Gateway is the original dialed destination in +E.164 format. Unified CM references the configured dial plan and routes the call to a locally registered endpoint on which the called destination is provisioned as a directory number.

To also enable enterprise abbreviated on-net inter-site dialing from Webex Calling to Unified CM, the respective enterprise patterns need to be added to the Webex Calling dial plan.

Figure 43 Enterprise abbreviated dialing from Webex Calling

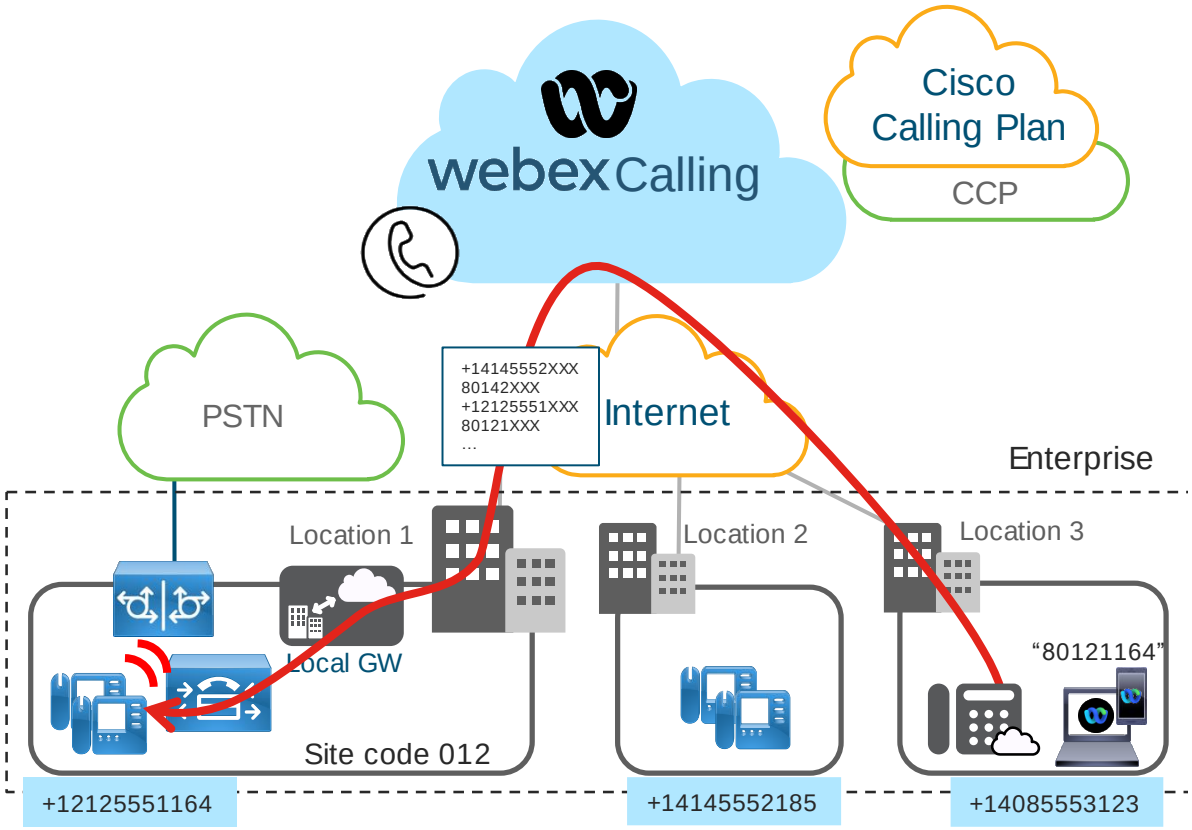


Figure 43 shows an enterprise with an enterprise-specific numbering plan using steering digit "8", four-digit routing prefixes, and four-digit extensions.

This enterprise-specific numbering plan can be used when dialing on a Webex Calling registered device. The dialed digit string, when evaluated in Webex Calling matches an enterprise pattern defined in a dial plan and as before the call is sent to the trunk or route group defined as the routing choice of the Webex Calling dial plan. The call is then sent on to Unified CM by the Local Gateway.

Unified CM implements abbreviated on-net inter-site dialing using the above enterprise specific numbering plan, the dial string "80121164" is mapped to the DID of an endpoint registered to Unified CM and the call is connected. To allow for enterprise-abbreviated dialing from Webex Calling to Unified CM, the appropriate settings (extension length, prefix length, and steering digit) must be configured and Unified CM must be configured to support this dialing habit when routing calls received from Webex Calling.

Figure 44 Calling a PSTN destination from Webex Calling

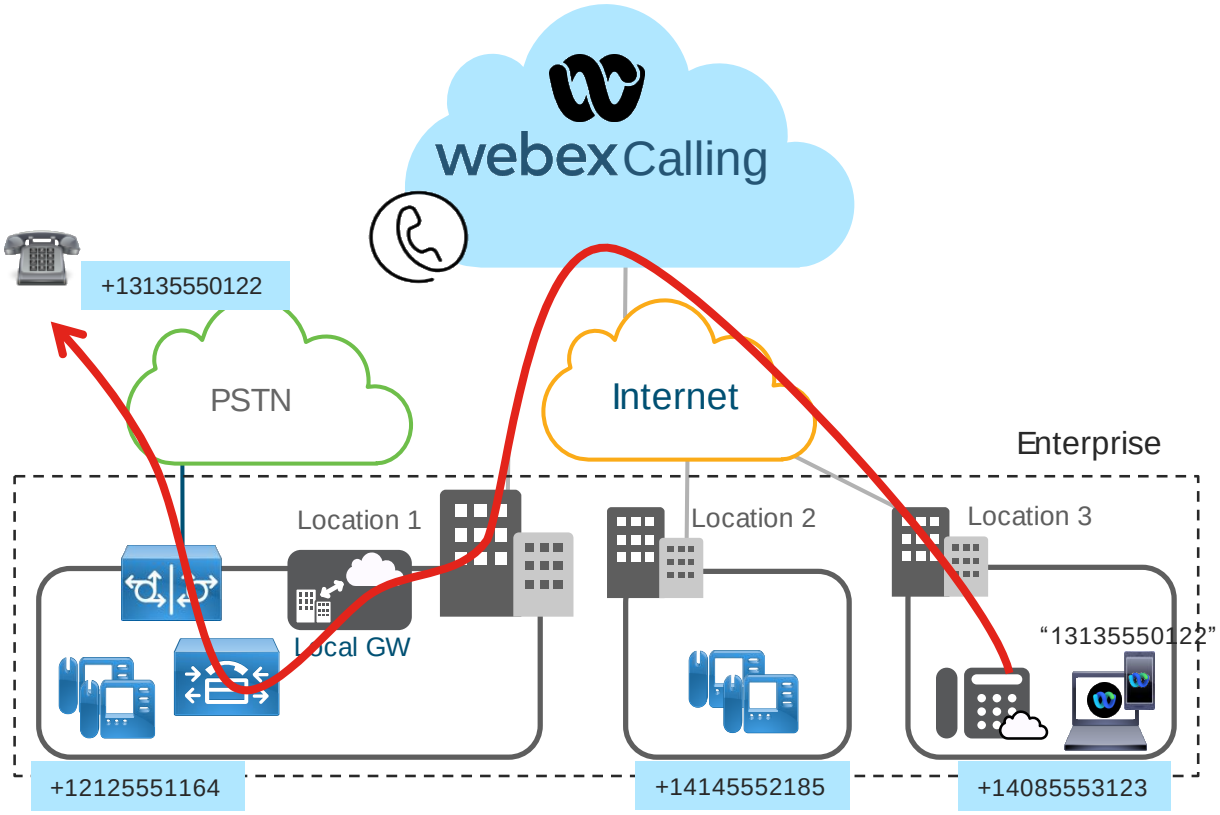


Figure 44 shows a PSTN destination dialed from a Webex Calling device in a location with premises based PSTN configured as the PSTN choice for that location. The call is considered off-net from Webex Calling's perspective and sent to the Local Gateway and then on to Unified CM. Unified CM does not locate an on-net match for the +E.164 address received from Webex Calling so it sends it on to the PSTN via the existing PSTN gateway controlled by Unified CM.

Calls from Unified CM to Webex Calling

To enable call routing from Unified CM to Webex Calling, a set of routes must be provisioned on Unified CM. This defines the set of +E.164 and enterprise numbering plan addresses in Webex Calling.

Figure 45 Calling from Unified CM to Webex Calling

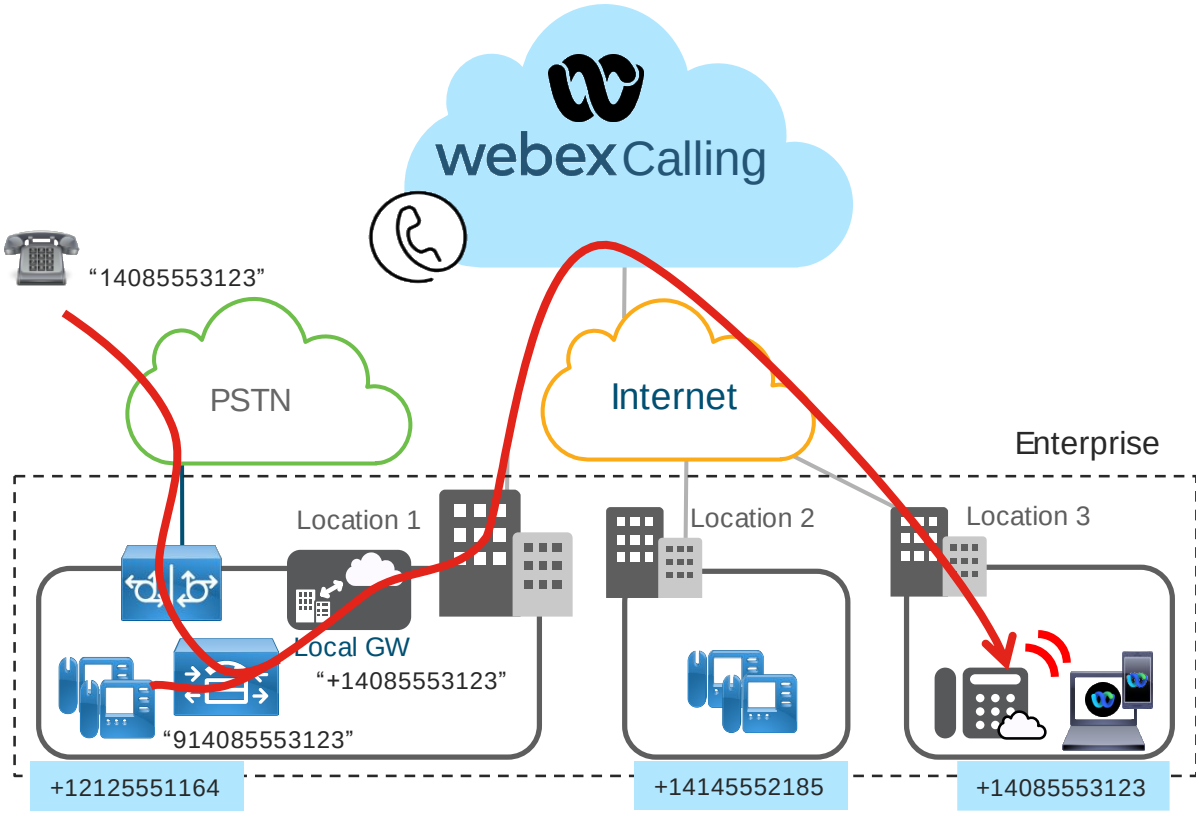


Figure 45 shows that with these routes in place, both depicted call scenarios are possible. If a PSTN caller calls a DID assigned to a Webex Calling device, the call is handed off to the enterprise via the enterprise's PSTN gateway, and sent to Unified CM. The called address of the call matches one of the Webex Calling routes provisioned in Unified CM and the call is sent to the Local Gateway. The called address must be in +E.164 format when sent to the Local Gateway. The Webex Calling routing logic routes the call to the intended Webex Calling device based on DID assignment. Calls from the PSTN should only be forwarded to Webex Calling by Unified CM if the Webex location uses premises-based PSTN or during a transition to cloud-based PSTN while the numbers haven't been ported to the new provider yet.

Calls originating from Unified CM registered endpoints destined for Webex Calling are subjected to the Unified CM provisioned dial plan. This dial plan usually allows users to use typical enterprise dialing habits for calling.

Called addresses for calls to Webex Calling can be in +E.164 format, enterprise format or extension format. If the called address is an extension (two to six digits) then Webex Calling uses the location of the trunk the call is received on as routing context to allow for disambiguation if the extensions in Webex Calling are not globally unique.

Allowing enterprise numbers or extensions as called addresses from Unified CM to Webex Calling enabled calls from Unified CM to Webex Calling users without a phone number.

Figure 46 Calling from Unified CM to Webex Calling

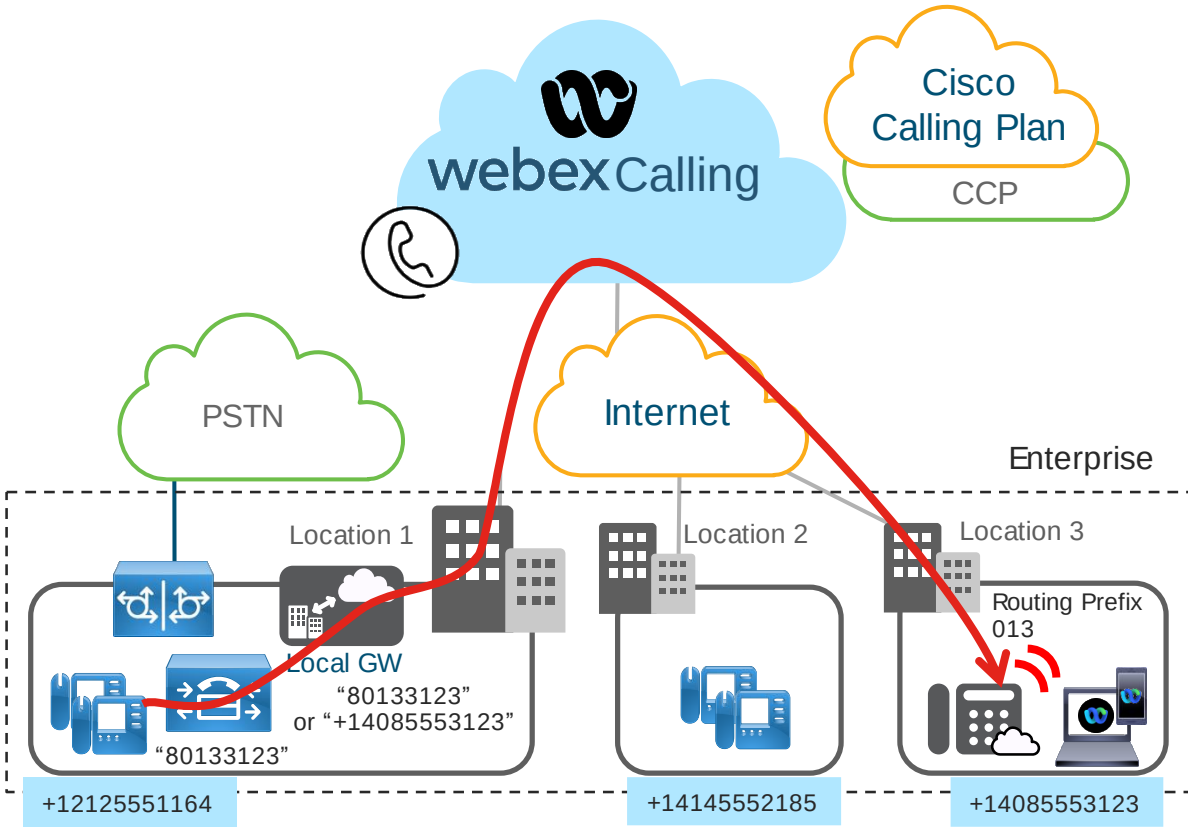


Figure 46 shows how using appropriate dialing normalization in Unified CM enables abbreviated on-net, inter-site dialing from devices registered to Unified CM to a Webex Calling location. The called address can either be normalized to +E.164 or sent without transformation. To normalize the called address to +E.164, Unified CM needs to know the ESN to +E.164 mapping for every Webex Calling location. When sending the called address unmodified then the enterprise numbering plan on Unified CM needs to be compatible with the enterprise numbering plan configured in Webex Calling. Not normalizing enterprise dialing to +E.164 for calls from Unified CM to Webex Calling enables calls to Webex Calling users without a phone number (extension only users).

Class of Service (CoS)

Strict CoS restrictions provide multiple benefits, including call loop avoidance and toll fraud prevention. When integrating Webex Calling Local Gateway with Unified CM's CoS, consider a CoS for each of a following:

- Unified CM registered devices
- Unified CM receiving calls from the PSTN
- Unified CM receiving calls from Webex Calling

CoS for Unified CM registered devices

Webex Calling destinations are added as new class of destinations to an existing CoS. Permission to call Webex Calling destinations is equivalent to permission to call on-premises (including inter-site) destinations.

If an enterprise dial-plan already implements an "(abbreviated) on-net inter-site" permission, then a partition is already provisioned on Unified CM and can be used to configure all the known on-net Webex Calling destinations in the same partition. If the "(abbreviated) on-net inter-site" permission does not yet exist yet then a new partition, for example "onNetRemote", must be provisioned. The Webex Calling destinations are added to this partition and it must be added to the appropriate calling search spaces.

CoS for calls coming from the PSTN or Webex Calling

Calls from the PSTN require access to all Webex Calling destinations only if premises based PSTN is used by Webex Calling locations or during the transition to cloud PSTN. This requires adding the above partition that holds all Webex Calling destinations to the calling search space used for incoming calls on the PSTN trunk. The access to Webex Calling destinations is in addition to any existing access.

Calls from the PSTN require access to Unified CM DIDs and Webex Calling DIDs. Calls from Webex Calling need access to Unified CM DIDs and PSTN destinations. The latter is only required if Webex Calling locations use premises based PSTN.

Figure 47 Differentiated CoS for calls from the PSTN and from Webex Calling

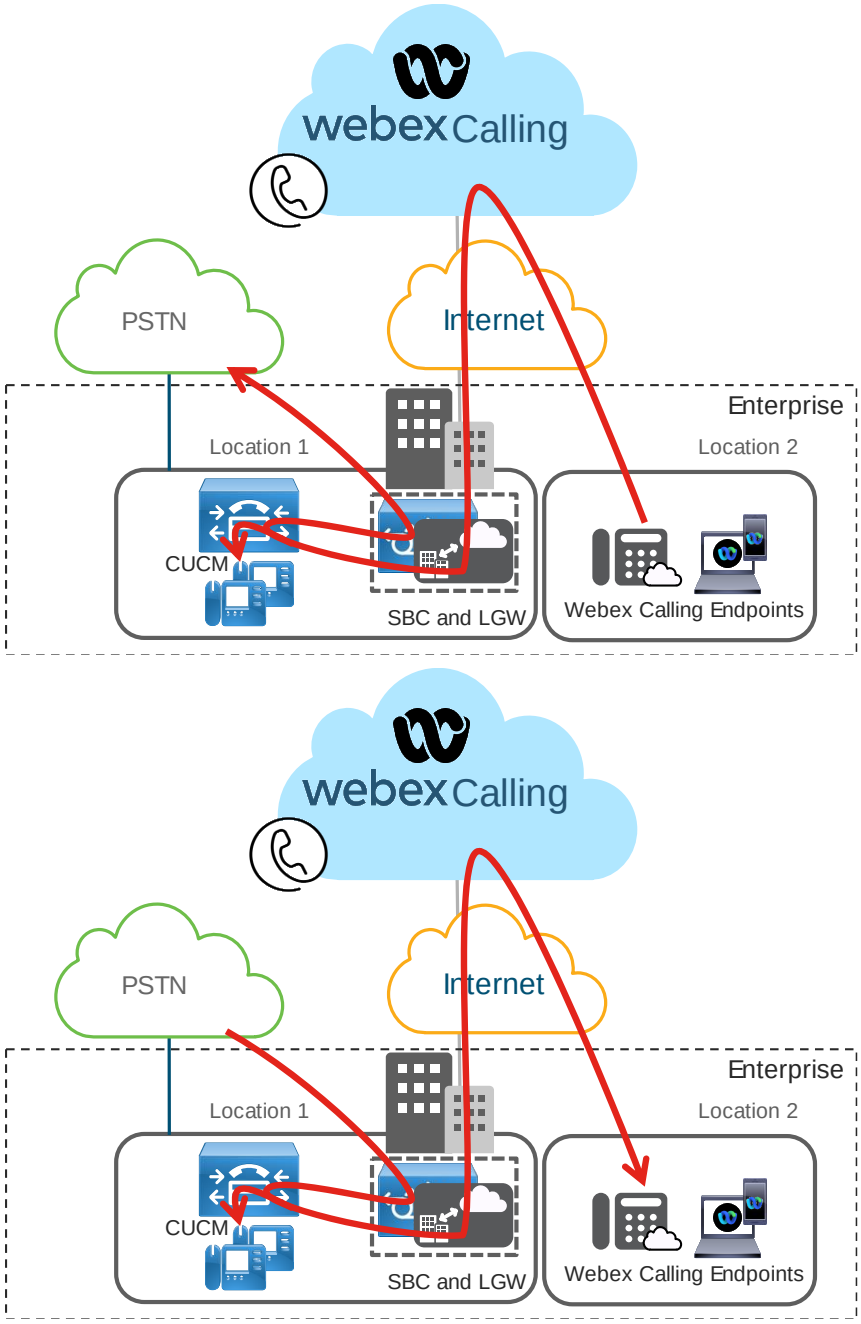


Figure 47 shows the two classes of service for calls from PSTN and Webex Calling. If the PSTN gateway functionality is co-located with the Local Gateway, then two trunks are required from the combined PSTN GW and Local Gateway to Unified CM: one for calls originating in the PSTN and one for calls originating in Webex Calling. This is driven by the requirement to apply differentiated calling search spaces per traffic type. Only with two incoming trunks on Unified CM this can easily be achieved by configuring the required calling search space for incoming calls on each trunk. Multiple trunks between Local Gateway and Unified CM can be configured by using different SIP listening ports for the two SIP trunks on Unified CM. The SIP listening ports are configured in the SIP trunk's security profile.

The dial plan in Unified CM is configured to use both trunks depending on the type of call: to PSTN or to Webex Calling. On the Local Gateway the trunk a call is received on is identified by matching on the port number in the topmost VIA header in the INVITE. This is achieved by combining `incoming uri via <some voice class>` on the incoming dial peer with `pattern :<UCM listening port>` in the `voice class uri <some voice class>` in the Local Gateway configuration.

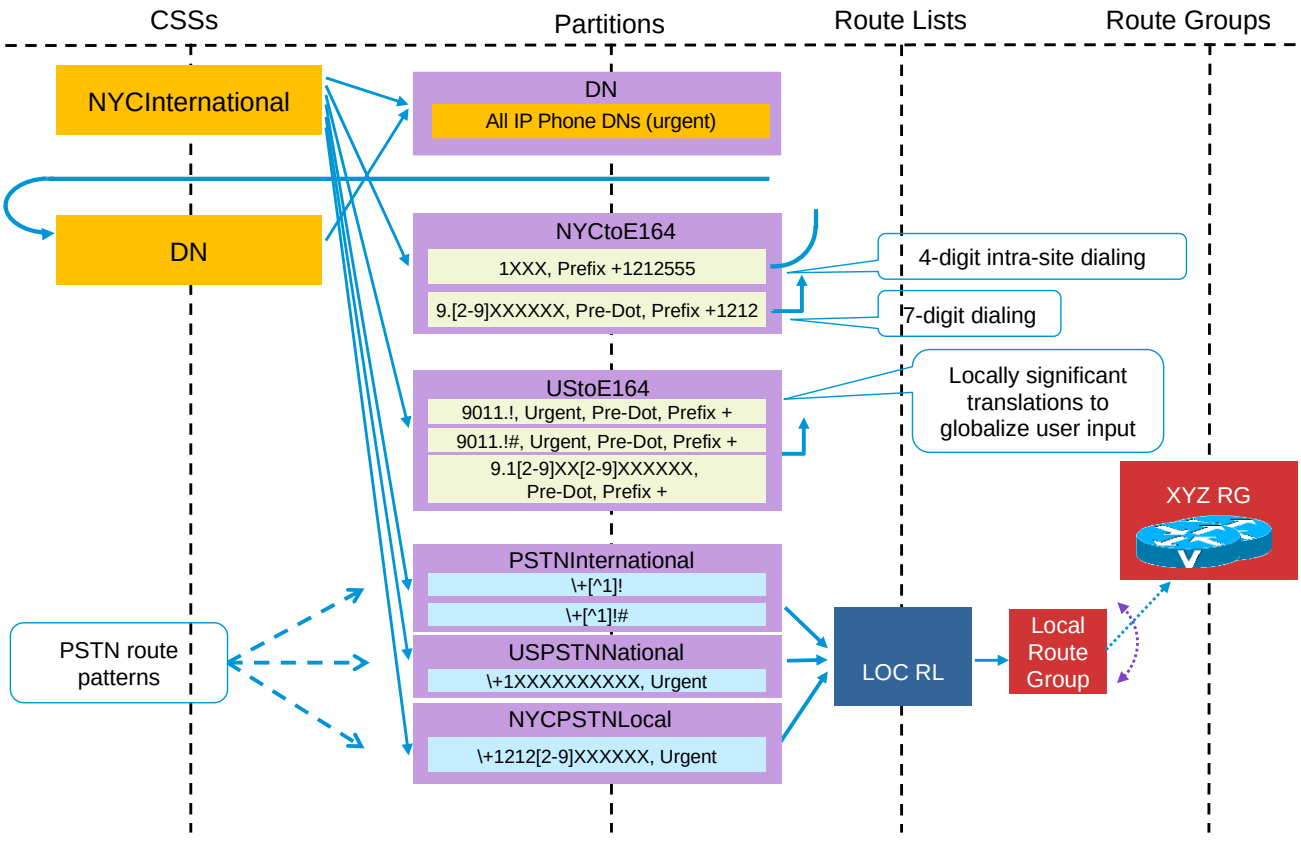
Unified CM Dial Plan Integration

This guide assumes an existing installation based on the best practices described in the latest version of the "Preferred Architecture for Cisco Collaboration On-Premises Deployments, CVD" available at https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/uc_system/design/guides/PAdocs.html.

The recommended dial plan design follows the design approach documented in the Dial Plan chapter of the latest version of the Cisco Collaboration System SRND available at the https://www.cisco.com/go/ucsrnd page.

The following information shows the configuration on Cisco Unified CM for Class of Service.

Figure 48 Dial Plan Recommendation



The previous figure shows an overview of the recommended dial plan design. Key characteristics of this dial plan design include:

- All directory numbers configured on Unified CM are in +E.164 format
- All directory numbers reside the same partition (DN) and are marked urgent
- Core routing is based on +E.164
- All non-+E.164 dialing habits. For example, abbreviated intra-site dialing and PSTN dialing using common dialing habits, are normalized (globalized) to +E.164 using dialing normalization translation patterns
- Dialing normalization translation patterns use translation pattern calling search space inheritance; they have the “Use Originator's Calling Search Space” option set.
- Class of service is implemented using site and class of service specific calling search spaces
- PSTN access capabilities (for example access to international PSTN destinations) are implemented by adding partitions with the respective +E.164 route patterns to the calling search space defining class of service

Figure 49 Adding Webex Calling Destinations to the Dial Plan

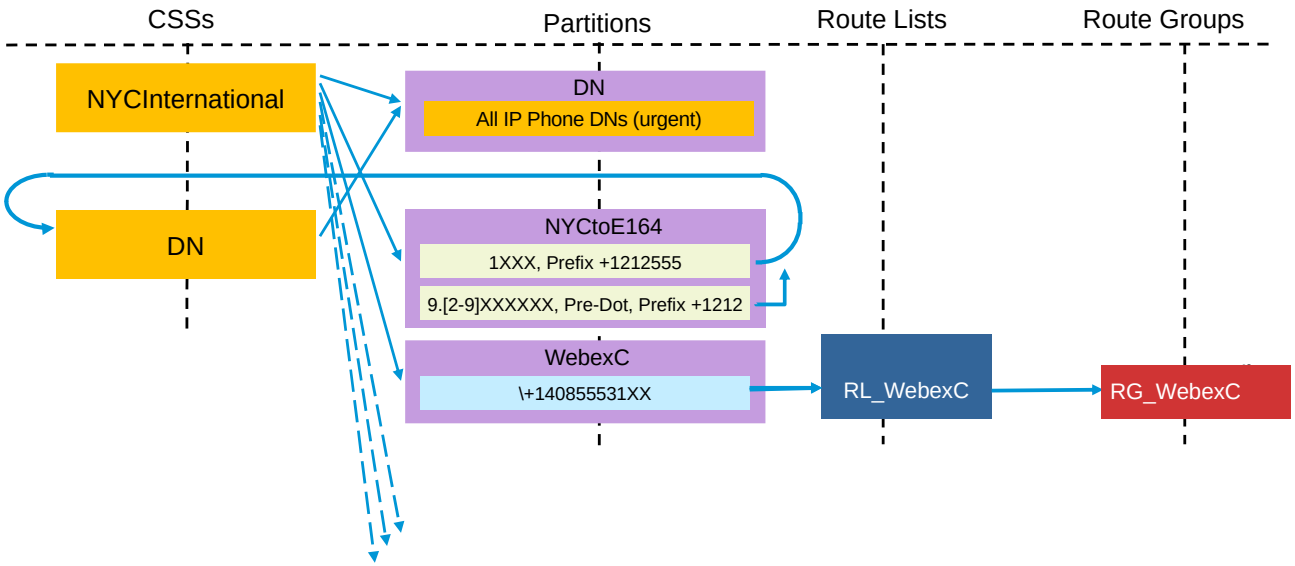


Figure 49 shows how to add reachability for Webex Calling destinations to this dial plan. A partition representing all Webex Calling destinations must to be created, “WebexC”, and a +E.164 route pattern for each DID range in Webex Calling is added to this partition. This route pattern references a route list with only one member: the route group with the SIP trunk to the Local Gateway for calls to Webex Calling. If multiple Local Gateways exist, then a Unified CM route group can be used for load balancing and redundancy. Because all dialed destinations are normalized to +E.164 either using dialing normalization translation patterns for calls originating from Unified CM registered endpoints or inbound called party transformations for calls originating from the PSTN this single set of +E.164 route patterns is enough to achieve reachability for destinations in Webex Calling independent of the dialing habit used.

If for example a user dials “914085553165” then the dialing normalization translation pattern in partition “UStoE164” normalizes this dial string to “+14085553165” which will then match the route pattern for a Webex Calling destination in partition “WebexC” so that Unified CM will ultimately send the call to the Local Gateway.

Figure 50 Adding Abbreviated Inter-Site Dialing

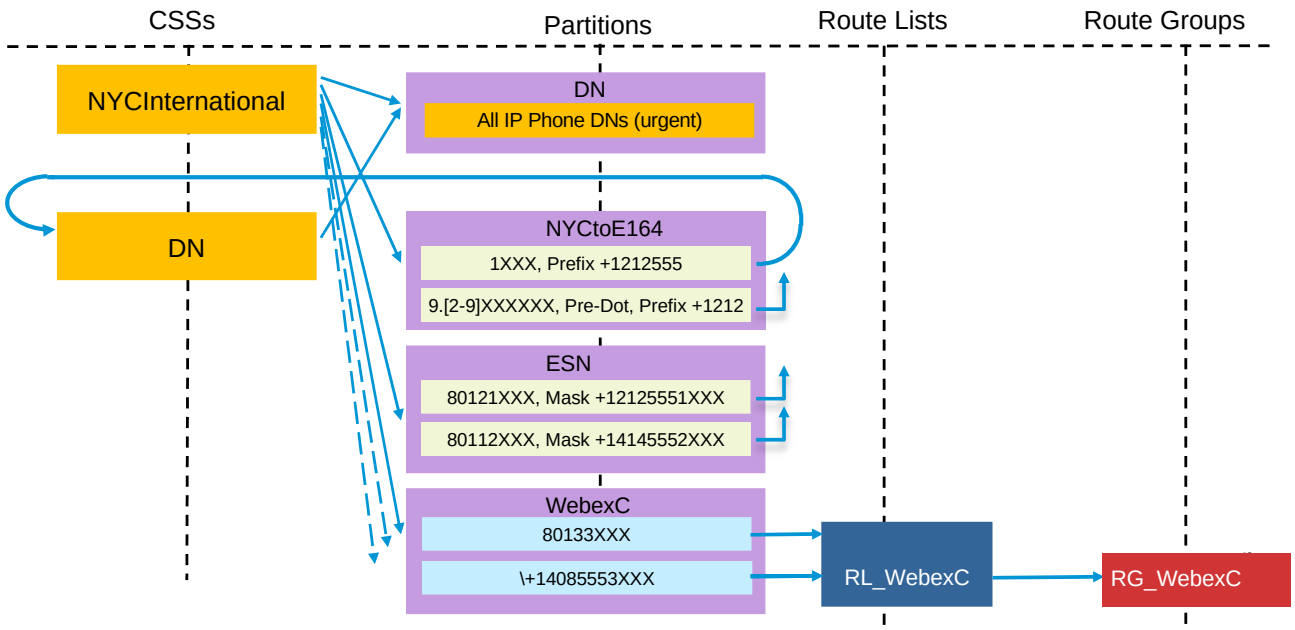


Figure 50 s**Error! Reference source not found.**hows the recommended way to add abbreviated inter-site dialing to the reference dial plan by adding dialing normalization translation patterns for all sites under the enterprise numbering plan to a dedicated partition (“ESN”, Enterprise Significant Numbers). These translation patterns intercept dial strings in the format of the enterprise numbering plan and normalize the dialed strings to +E.164.

Adding enterprise abbreviated dialing to Webex Calling destinations is achieved by adding the respective route pattern for the Webex Calling location to the “WebexC” partition (for example “80133XXX”). The called party is not transformed and instead sent to Webex Calling as is. This allows to use enterprise abbreviated dialing from Unified CM to extension only Webex Calling users.

Local Gateway deployment

At least one Local Gateway needs to be provisioned for the connection between Webex Calling and Unified CM. Multiple Local Gateways can be configured for scale and redundancy. The location the Local Gateway is assigned to is only relevant if the dialing context needs to be established for extension dialed calls from Unified CM to Webex Calling. If no partial migration occurs where some of the users of the same location are on Unified CM and some are on Webex Calling, extension dialing from Unified CM to Webex Calling is not a valid use case so that in that case the location association of the Local Gateways does not matter. The decision about the number of Local Gateways to deploy and where to deploy the Local Gateways is mainly driven by scale and reliability requirements.

In the simplest case, all Local Gateways are added to a single route group.

Webex Calling dial plan configuration

The global “Unknown Numbers Handling” parameter is set to “Standard behavior” so that routing to Unified CM and call classification is based on the dial plan patterns in an enterprise dial plan. The single route group with all Local Gateways is used as routing choice for that dial plan.

The global “Caller ID Format for Calls from and to On-premises” setting should be set to “ESN” to make sure that callback is possible from Unified CM for calls originating from Webex Calling

For each enterprise and +E.164 number range on Unified CM, the respective pattern is added to the dial plan.

On each Webex Calling location, the internal dialing routing prefix is configured in line with the end-to-end enterprise dial plan. It is essential to select a unique site prefix for each location. The same site prefix cannot exist on Webex Calling and on Unified CM at the same time. The “Enable routing unknown extensions to the Premises as internal calls” option is disabled.

Case Study 1b: Webex Calling with Dedicated Instance

This case study examines scenarios where Webex Calling locations are combined with Webex Calling Dedicated Instance. This type of deployment provides a Cisco Unified Communications Manager based stack of applications, hosted in Webex and part of the Webex Calling offering, dedicated to a single customer.

Dedicated Instance Call Flows

Dedicated Instance follows the standard approach of signaling via the cloud and media transmitted directly between endpoints. The media between endpoints will use the customer's network both within a site and between different sites.

The call flows described in this section include customer intra-site calls, customer inter-site calls, PSTN calls from Dedicated instance, and calls between endpoints registered to multi-tenant and dedicated instance platforms. Calls to external devices will flow through a PSTN connection via the customer WAN. If a customer has Mobile and Remote Access (MRA) configured, calls flow through the customer WAN.

Customer Intra-site Calls

Customer intra-site call signaling travels via the dedicated instance cloud. Media stays on-premises within the customer's site.

Figure 51 Dedicated Instance Customer Intra-site dialing

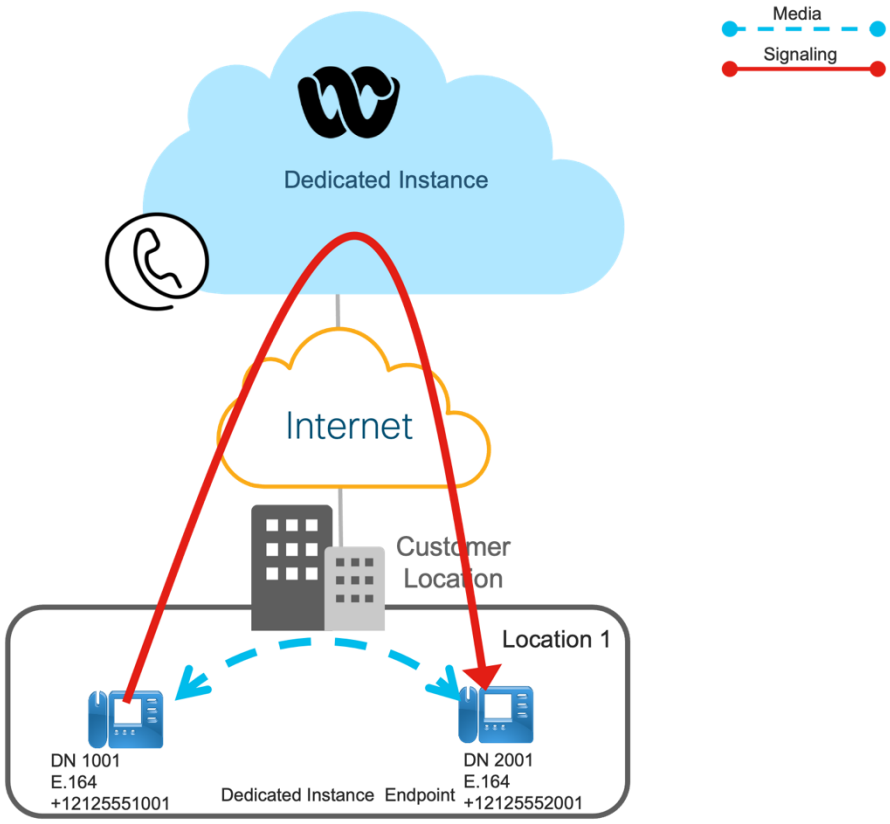


Figure 51 shows a single location Webex calling Dedicated instance deployment. Phones are registered to Dedicated instance. When 1001 dials 2001, call signalling traverses the Dedicated instance cloud. When call is answered by 2001, media is connected locally between phones.

Customer Inter-site Calls

Signalling for customer inter-site calls traverses the Dedicated Instance cloud, while media flows over the customer's WAN.

Figure 52 Dedicated Instance Customer Inter-site dialing

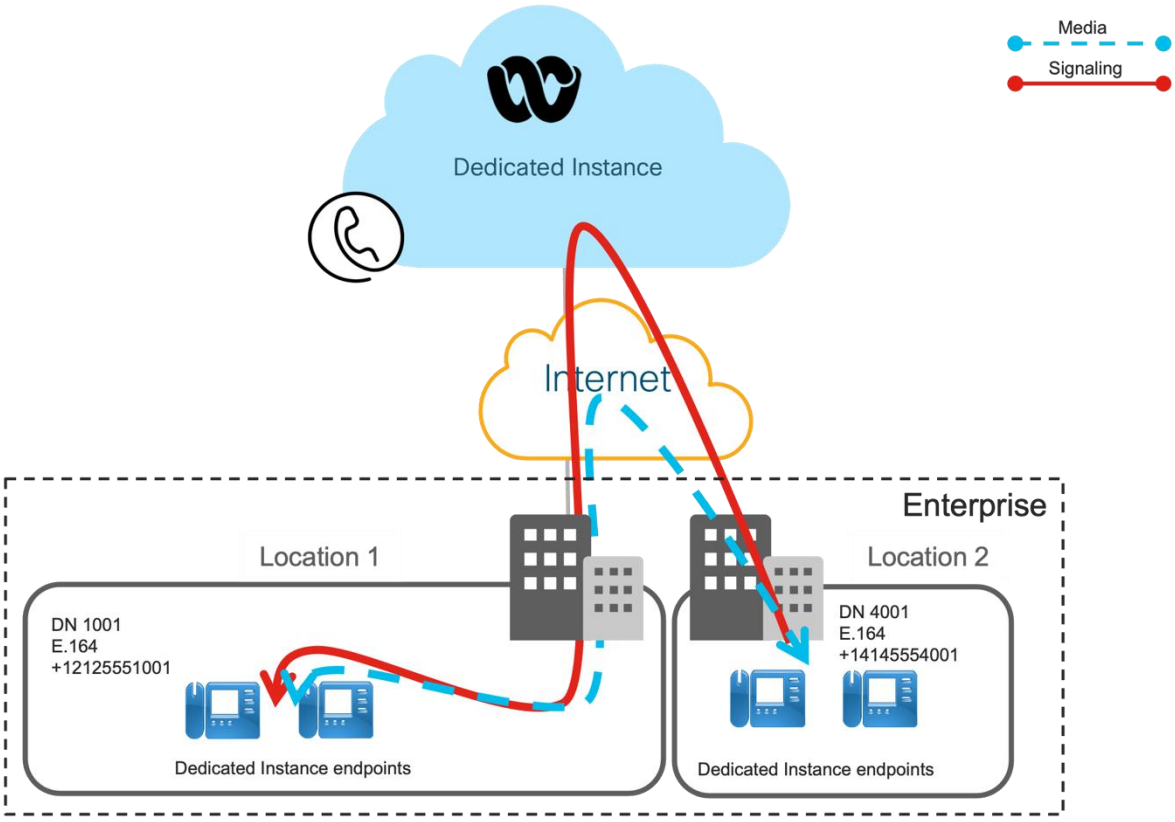


Figure 52 shows a multiple location of Webex calling Dedicated instance deployment. Both the location phones are registered with Dedicated instance. When phone 4001 at location 2 dial extension 1001 on location 1. It sends the call to dedicated instance cloud. The enterprise dial plan provisioned on Dedicated instance send the call to 1001. When 1001 answer the call, media will flow over the customer's WAN.

PSTN Calls from Dedicated Instance

Local Gateway is deployed on customer or partner premises. Local gateway is connected to the dedicated instance Unified CM via a SIP trunk and connect PSTN calls to the dedicated instance.

Figure 53 Dedicated Instance PSTN calls

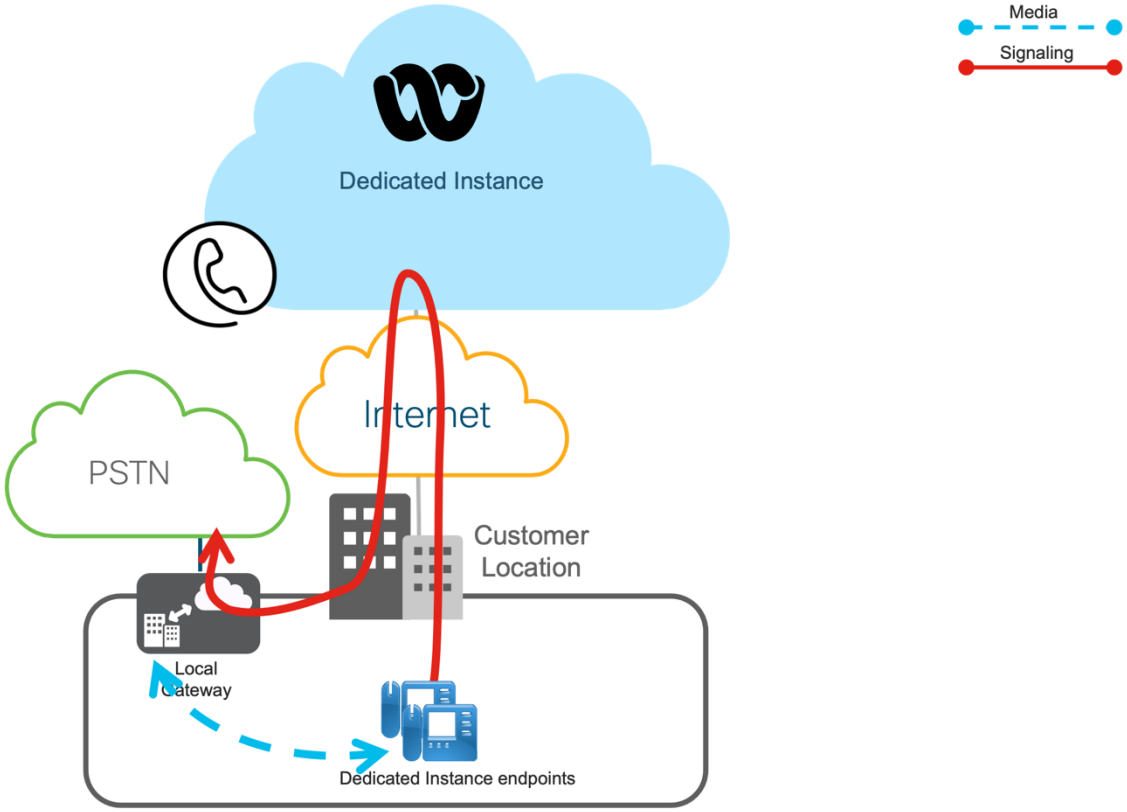


Figure 53 shows a single-location Webex Calling Dedicated instance deployment with Local Gateway. Calls originating from Webex Calling Dedicated Instance endpoints are sent to the Local Gateway based on Webex Calling Dedicated Instance dial plan routing logic. The Local Gateway then sends the calls on to PSTN. The enterprise dial plan provisioned on Webex calling Dedicated Instance determines whether the call needs to be extended to an endpoint registered to Webex Calling or to the PSTN via Local Gateway.

Dedicated Instance Endpoint to Multi-tenant Endpoint Calls

For the calls between dedicated instance registered endpoint and multi-tenant registered endpoint, signaling and media travels through Webex calling cloud.

Figure 54 Multi-tenant endpoint to Dedicated Instance calling

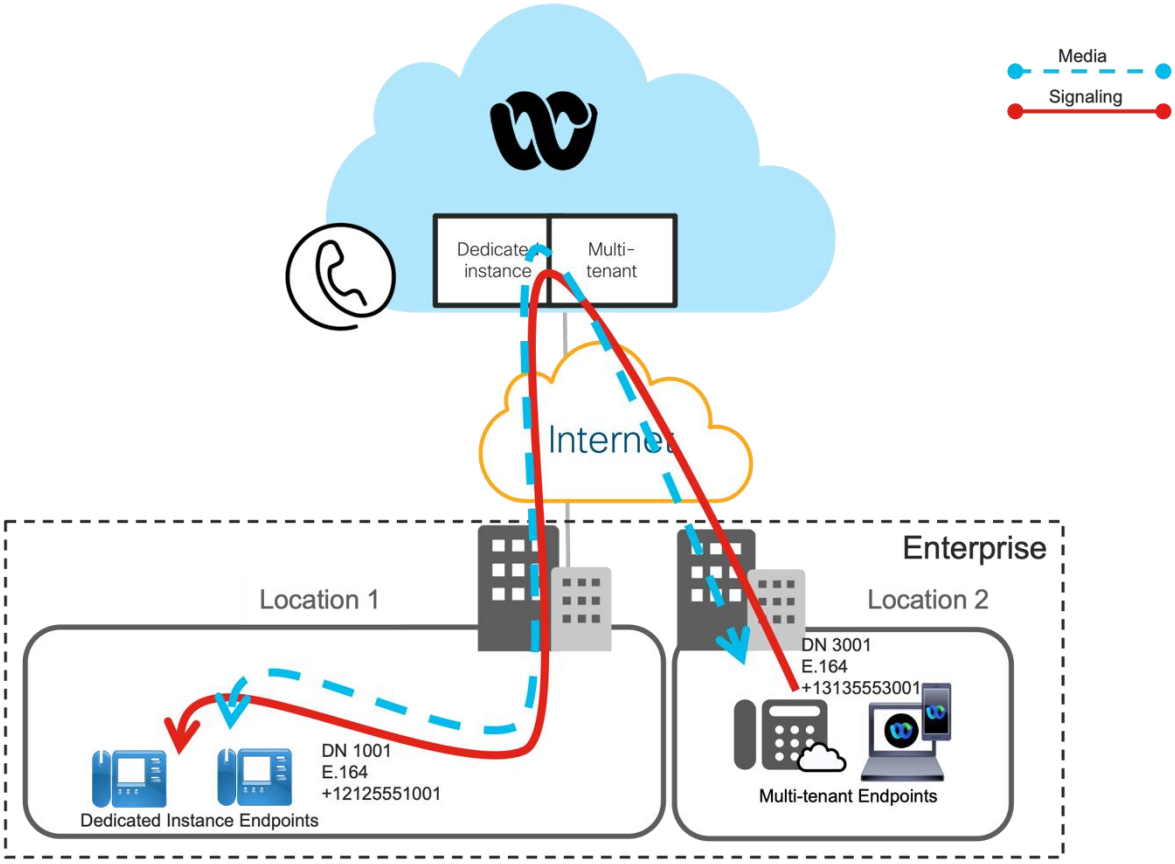


Figure 54 shows an example of a Multi-tenant endpoint calling a Dedicated instance registered endpoint. A Webex Calling user is dialing a number 1001 and the dialed number does not match any number provisioned for the customer in Multi-tenant but the number matches a dial plan pattern configured in a dial plan. The call is therefore sent to the trunk or route group selected as the routing choice for that dial plan. The Pre-configured Local Gateway sends the call on to Dedicated instance. Call routing on the pre-configured Local Gateway does not take the called address into consideration; routing is solely based on trunk attributes so that any call received from Dedicated instance is forwarded to Multi-tenant and any call received from Multi-tenant is forwarded to Dedicated instance. Dedicated instance references the configured dial plan and routes the call to a locally registered endpoint on which the called destination is provisioned as a directory number.

Call Routing Consideration

Call routing consideration for Webex calling Dedicated instance and Multi-tenant follows same concept and configuration for Unified CM with multiple Webex calling locations covered in case study 1a. When Webex calling Multi-tenant phone makes a call to phone in Dedicated instance, it uses dial plan to send the call to pre-configured trunk to reach phone on Dedicated instance. Partners and customers can plug in their existing dial plan with minimal configurations. Partners/Customers should not delete the preconfigured trunk on both the dedicated instance and multi-tenant deployments, which would affect the internal dialing and would require a Service Request to re-establish the connection. For more information on Base configuration refer to https://help.webex.com/en-us/article/2vpf1/Dedicated-Instance-for-Webex-Calling---Base-Configuration

Mobile and Remote Access Considerations

Mobile and Remote Access through Expressway can be deployed for remote users when needed. This has the benefit of reducing the impact on WAN traffic. As ICE is supported via mobile and remote access, in many cases two users over MRA would be able to have direct media. If some of mobile and remote access users have both a hardware endpoint and

a Webex app, then CTI-controlling the MRA phone from the Webex app is a supported scenario (Boardless CTI) even if Webex app uses a VPN-less connection to Dedicated Instance Unified CM. Not only users, but also small offices might benefit from this feature and use MRA instead of other connectivity options.

Dedicated Instance Dial Plan considerations

The dial plan is one of the key elements of a Webex Calling Dedicated instance, and an integral part of all call processing agents. Generally, the dial plan is responsible for instructing the call processing agent on how to route calls. Specifically, the dial plan performs the following main functions:

- Endpoint addressing: For destinations registered with the call processing agent, addresses are assigned to provide reachability. These internal destinations include all endpoints (such as IP phones, video endpoints, soft clients and analog endpoints) and applications (such as voicemail systems, auto attendants, and conferencing systems).
- Path selection: Depending on the calling device and the destination dialed, a path to the dialed destination is selected. If a secondary path is available, this path will also be considered if the primary path fails.
- Calling privileges: Different groups of devices can be assigned to different classes of service, by granting or denying access to certain destinations. For example, lobby phones might be allowed to reach only internal and local PSTN destinations, while executive phones could have unrestricted PSTN access.
- Manipulation of dialed destination: On the path from the dialing device to the dialed destination, the dial plan can apply manipulations to the dialed destination. For example, users in the US might dial 9011496901234 to reach a destination in the PSTN in Germany, while a user in France might be able to reach the same destination by dialing 000496901234. This dialed destination would need to be presented as 011496901234 to a PSTN trunk on a gateway in the US and as 00496901234 to a PSTN trunk on a gateway in France.
- Presentation of information about identities involved in the call: During session establishment and also while in a call, on both the calling and the called device, information about the other device is displayed. Depending on call state and direction, this includes calling, diverting, alerting, and connected party information. The dial plan can define mappings that influence the format and content of information displayed.

For more information on how dial plans can be implemented in Dedicated Instance, see the Cisco Enterprise Preferred architecture.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**