Design Scalable SIP Trunk Solution with vCUSP and CUBE

Contents

Introduction

Prerequisites

- Requirements
- Components Used

Problem

Solution

- Overview of Scalable SIP trunk solution with vCUSP and (v)CUBE
- Solution Description
  - Network Diagram - Base solution
  - Add SIP trunk Redundancy.

Introduction

This document describes a solution network design that enables scalable Session Initiation Protocol (SIP) trunks for enterprises and service providers. In this solution, a Cisco Unified SIP Proxy (CUSP) is used to federate incoming and outgoing calls over SIP trunks to a pool of Cisco Unified Border Element (CUBE) routers .

Contributed by Andres Salgado, Technical Marketing Engineer CUBE and Luis Ramirez Cisco TAC Engineer

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- CUBE
- TCP
- SIP
- UDP
- CUSP

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Problem

Load-share for multiple SIP trunk environments for deployments with multiple CUBE elements due to scalability, redundancy requirements from one or multiple SIP providers.

Solution

Overview of Scalable SIP trunk solution with vCUSP and (v)CUBE

Incoming SIP trunk signalling from a Service Provider is terminated on the CUSP. The CUSP distributes the calls to a pool of CUBE routers, which process call signaling and set up media sessions as required. SIP trunk call capacity can be scaled simply by an increase of the size of the (v)CUBE router pool. Thus the number of SIP trunks, as signified by the number of IP addresses for the signaling channel, can be minimized to just one.
A second CUSP with its associated SIP trunk can be added to the solution to introduce trunk redundancy and load balance. The service provider distributes calls over the two SIP trunks. In case of a fault with one CUSP, the service provider directs all calls to other SIP trunk, thus avoids service outages. This requires Options ping to be enabled from the Service Provider to monitor if the SIP trunk is UP.
In addition, the pool of CUBE routers increases the overall availability of the solution. Failure of any CUBE in the pool just reduces the call handle capacity of the solution, rather than to cause SIP trunk outages.
The CUSP incorporates policy engine features that allow for policy-based routing of calls such as time-of-day routing.
This design guide presents the architecture and components of the solution

Solution Description

This section describes the base scalable SIP trunk solution. The base solution provides scalable and load balance of SIP Trunks among CUBEs.
The base solution consists of the next elements:

•SIP trunk from the service provider.

•A CUSP

•Four CUBE routers. If incoming call demand grows, additional CUBEs can be added without required changes at the service provider or at the Cisco Unified Communications Manager

•The Cisco Unified Communications Manager

• The signaling path is represented by the blue line

•A media path for all elements, represented by the red line

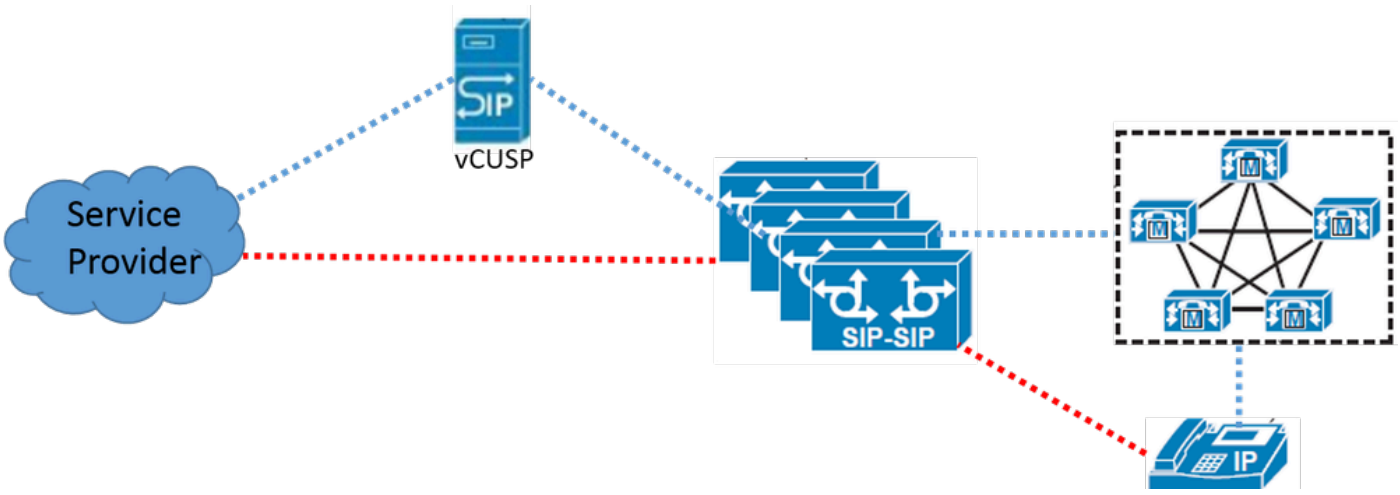
•Table-based routing supported by CUSP route tables

•Keepalive messages configured use the **server-group sip ping-options** command. The CUSP uses these messages to determine if a peer element is up or down, and if it determines that the element is down, it mark it as such and to stop calls to it. In this solution, the CUSP uses this command to test connections with Service Provider peers and the CUBE routers

CUBE routers can use the **voice-class sip options-keepalive** command to verify the status of peer elements. You can find out more about this command here:

This solution can be developed from a basic topology to a solution that has scaled to meet increased call volume and that has added failover, redundancy and routing to different service providers. You can have multiple service providers, multiple vCUSP and multiple (v)CUBEs in HA if required.

Network Diagram - Base solution



Add SIP trunk Redundancy.

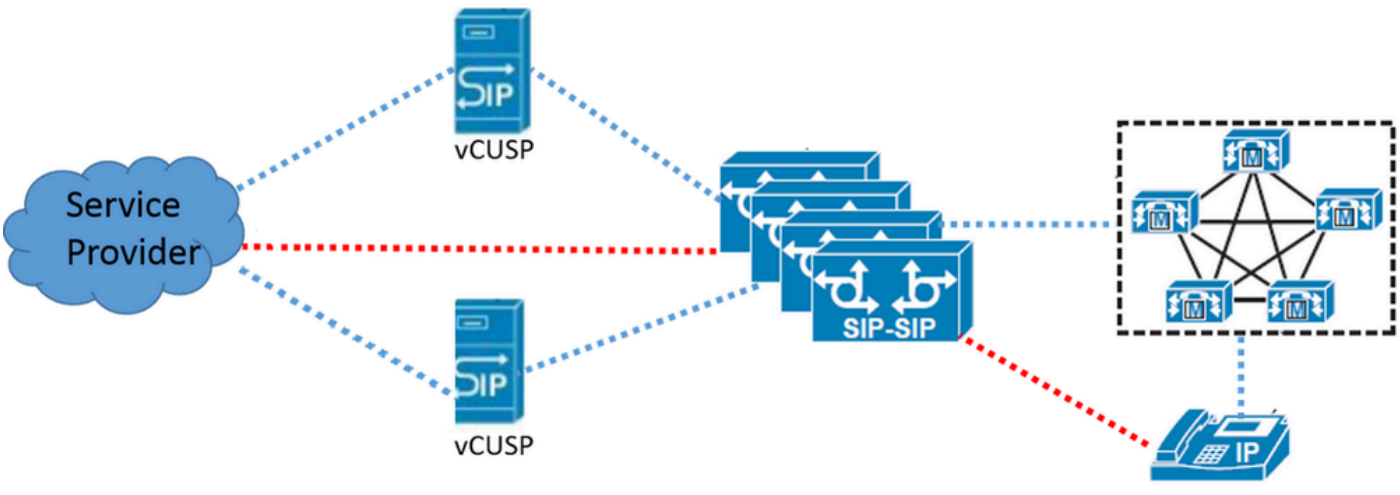
This image shows a redundant SIP trunk to the same service provider. Redundant SIP trunks ensure that SIP signaling can switch over to the secondary trunk if the primary trunk fail, and that new call requests can be handled. Redundancy can also be used for load balance.

This scenario adds these elements to the base solution topology:

•One additional SIP trunk to the service provider

•A CUSP

Topology for Redundant SIP Trunks from the Same Service Provider



there is a primary and a secondary CUSP. If the trunk with the primary fails, the service provider contacts the secondary CUSP.

Topology for a SIP Trunk from a Second Service Provider

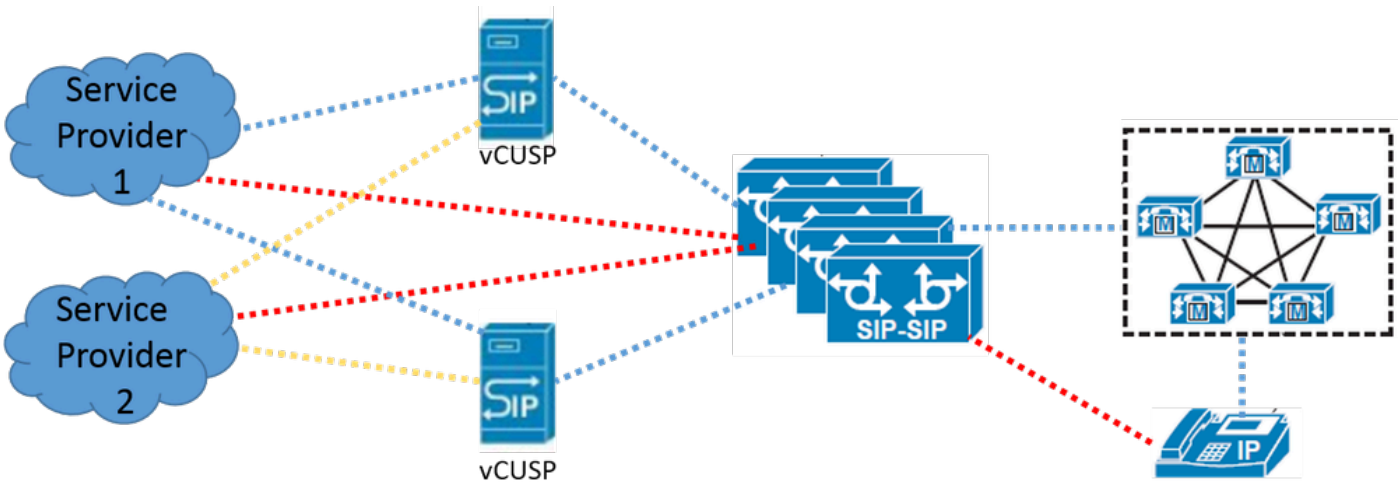


Image shows Service Provider 1 and its connections in light color to contrast with Service Provider 2. The figure shows that the Service Provider can load balance, Active-Active configuration with both CUSP. This can be accomplished by the service provider awareness of cusp1 and cusp2 IP addresses, If the attempt to reach cusp1 fails, the Service Provider routes to the cusp2 to take the additional load.

Routing policies configured on the CUSP can be used to control outbound calls to the service provider.

SIP trunk service providers can offer service plans that charge different call cost rates depend on the destination, time of day. When this is the case, you can route calls to the service provider accordingly to take advantage of the lowest rate.

CUBE-to-CUSP

DIfferent methods can be used to have the CUBE load balance among Cisco Unified SIP Proxies:

- A DNS SRV-based session target can be configured to allow the CUBE to use the priority of the DNS response
- Server Groups in Outbound Dial Peers on the CUBE. In order to effectively use this option, you must configure **voice-class sip options-keepalive profile** command to monitor the CUSP associated with the dial peer. If the CUSP is in down state, the server is marked down, and the CUBE can try the second CUSP without an attempt first the CUSP in down state

Related Information

- CUSP CLI Configfuration Guide
- SIP RFC
- **Technical Support & Documentation - Cisco Systems**