Application Note

Telecom
Dialogic® HMP
Media Server

Developing Higher Density Solutions with Dialogic® Host Media Processing Software

A Strategy for Load Balancing and Fault Handling

Executive Summary

Combining the SIP proxy and DNS-SRV load balancing and fault handling techniques provides a cost-effective way to develop higher density solutions using Dialogic® Host Media Processing (HMP) software. This application note describes the two techniques, explains how to combine them, and provides a test configuration for a solution using the combined techniques and Dialogic HMP software.

Table of Contents

Introduction

High-density IP-based media servers are in demand to provide multimedia services for a growing user base in both the service provider and enterprise market segments. Dialogic HMP software is a key media server building block for enabling rich multimedia applications that can be easily deployed in IP networks. Dialogic HMP software runs on standard Dialogic® Architecture-based platforms with either a rack mount or bladed form factor and supports a fixed density of multimedia resources per platform. To develop a higher density system, a cluster or server farm approach can be used in which multiple servers appear externally as a single server.

Distributed architectures also enable an environment where multiple media servers each cater to a specific application such as conferencing, IVR, or speech. The SIP proxy server can engage with the appropriate media server when a specific function is desired. For example, an incoming call for a conference can first be serviced by an IVR, which plays an announcement and collects necessary information from the caller such as conference bridge number and passcode. Once this information is authenticated, the IVR server notifies the SIP server to redirect the call to a conferencing server, which can be a completely separate, dedicated media server. This technique allows a robust and flexible architecture that can achieve high densities and high availability.

This document details a distributed architecture using AdvancedTCA single board computers running Dialogic HMP software to provide multimedia functions.

Load Balancing and Fault Handling Techniques

Several techniques can be employed for load balancing and fault handling. For media server scenarios that use IP signaling protocols such as SIP or H.323, the signaling traffic must be load balanced and distributed among multiple servers. These servers perform various media server tasks, such as playing an announcement, interacting with a caller using an IVR application, or processing speech. Some of these techniques are discussed briefly in this section.

Session Protocol

IP call control signaling using SIP or H.323 can establish media sessions between callers or endpoints and the media server. Other protocols such as MGCP, H.248, or Megaco may be used to control media streams between media servers and media gateways, but only SIP and H.323 will be discussed here.

Load balancing can be achieved when an application implements the functionality of a Gatekeeper in H.323 and a Proxy in SIP. The application intercepts the incoming H.323 or SIP messages, diverts them to appropriate "worker-servers" that do the processing needed, and modifies the messages so that proper routing takes place. A proprietary interface may be used to communicate with individual worker-servers, keep track of resource usage, and maintain a "live" or "heartbeat" status to help the application perform efficient load balancing. Other mechanisms that can be used to communicate between these servers include:

- SNMP RFC 1157
- SIP INFO requests
- RADIUS heartbeat

The comparison between H.323 and SIP in Table 1 shows why SIP is the protocol of choice for developing a load balancer.

H.323	SIP
Complex protocol with sub-protocols	Simple, easy-to-use protocol
Uses binary messages	Uses text messages
Has elaborate message set	Uses fewer messages
Difficult to debug	Simple to debug because it is text-based

Table 1. Comparison of H.323 and SIP Protocols

DNS-SRV-Based

Domain Name Service (DNS) is an integral part of any network, and network administrators are familiar with DNS A records. These records associate the name of a particular device (usually a PC) with the IP address of the network adapter for that computer. For example, every computer in a domain is given a name that does not change; however, the IP address of that computer may change. To ensure that the computers in a domain can communicate with each other, they must use their individual computer names instead of specific IP addresses.

In SIP communications using DNS-SRV records (RFC282), the name of a service is associated with the IP address of devices that provide those services. Typically, the records point to the IP addresses of SIP proxies that forward the SIP requests appropriately to various SIP endpoints. The syntax for DNS-SRV records is:

```
Service.Proto.Name TTL Class SRV
Priority Weight Port Target
```

Table 2 provides an explanation of the various arguments in the syntax.

Argument	Explanation
Service	Symbolic name of requested service (Example: SIP)
Proto	Most useful values are TCP and UDP
Name	Domain to which the record refers
TTL	Standard DNS term (optional)
Class	Standard DNS term (optional)
SRV	Standard DNS syntax token
Priority	Priority of target host. A client must attempt to contact the target host with the lowest priority number it can reach; target hosts with the same priority should be tried in an order defined by the weight field (see below). Range is 0-65535.
Weight	Load-balancing mechanism. When several target hosts have the same priority, a weight value signals the precedence that the host described in the record should have. Range is 1-65535. Domain administrators should use Weight 0 when no load balancing is required.
Port	Port of the service on the target host. Range is 0-65535.
Target	Domain name of target host

Table 2. DNS-SRV Record Syntax

Here is a sample record in which TTL and Class have been omitted:

```
_sip._tcp.dialogic.com SRV 10 1 5060
serv1.dialogic.com
```

In this example, the service name is "_sip._tcp.dialogic.com," the priority is 10, the weight is 1, the port is 5060, and the target is "serv1.dialogic.com."

Combining Techniques

Session-protocol load balancing uses an algorithm such as round robin or weighted, depending on the particular implementation. When SIP is used, SIP requests are forwarded to the media servers based on one of these algorithms. Although no specific drawbacks have been associated with this technique, additional processing is required to implement the load balancing algorithm.

If DNS-SRV-based load balancing is used, load balancing tasks move from the SIP server to the DNS, which is a standard part of any network. The SIP server is only required to support DNS SRV, and such support is already a requirement of the SIP standard RFC3261 for implementing redundancy and fault handling. Information about RFC3261 can be found at http://www.ietf.org/.

By combining session-protocol and DNS-SRV load balancing, the DNS can provide load balancing and redundancy for the Proxy or Gatekeeper, which can then provide more configurable and flexible load balancing and fault handling options to the media servers. A combined SIP proxy and DNS-SRV technique is discussed in this paper. When these two protocols are combined, the result is a robust load-balancing and fault-handling solution that can be used for building high-density SIP-based media server applications.

SIP and DNS-SRV-Based Solution

When the SIP proxy and DNS-SRV techniques are combined for fault handling, multiple servers are assigned for a domain and the DNS refers to them. When the primary server is down or not reachable, the DNS-SRV records can be used to access information about backup servers, and requests for service can be sent to any of these backup servers. Backup servers may be located in different geographic locations to minimize downtime caused by operational problems or natural disasters. Use of the backup servers is transparent to the client.

To achieve high availability, multiple DNS-balanced proxy servers can be used with multiple media servers having SIP proxy load balancing. However, to simplify the solution discussed in this section for demonstration purposes, only a single server with the SIP proxy and DNS SRV is used.

The solution uses the open source SIP Express Router (SER) as the load balancing SIP proxy. (For more information about SER, visit http://www.iptel.org/ser/.) SER is a high-performance, configurable SIP server that conforms to RFC3261, and it can act as SIP registrar, proxy, or redirect server.

Call Flow

Figure 1 shows a typical call flow among a SIP user agent, the load balancer, and a media server running Dialogic HMP software.

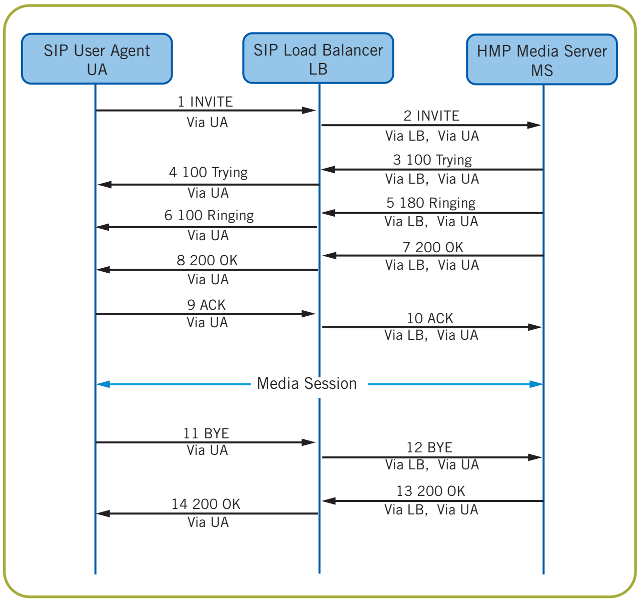

Figure 1. Call Flow

Test Configuration

Figure 2 shows the test configuration for a load-balancing and fault-handling solution that can be used for building high-density SIP-based media server applications with Dialogic HMP software. The solution has been tested using AdvancedTCA single board computers in an AdvancedTCA chassis.

Table 3 *(on page 5)* provides details about the test configuration in Figure 2.

Routing Logic

Routing logic has been modified from the default. Figure 3 shows the original routing logic and Figure 4 shows the modified routing logic.

```
Route()
{
        # forward to current uri now; use
stateful forwarding;
        # that works reliably even if we
forward from TCP to UDP

        if (!t_relay())
```

Figure 3. Default Routing Logic

```
Route()
{
        # forward to current uri now; use
stateful forwarding;
        # setup a failure route
        t_on_failure("1");
        t_relay();
}

failure_route[ 1]
{
        lookup("location");
```

Figure 4. Modified Routing Logic

A failure route has been added.

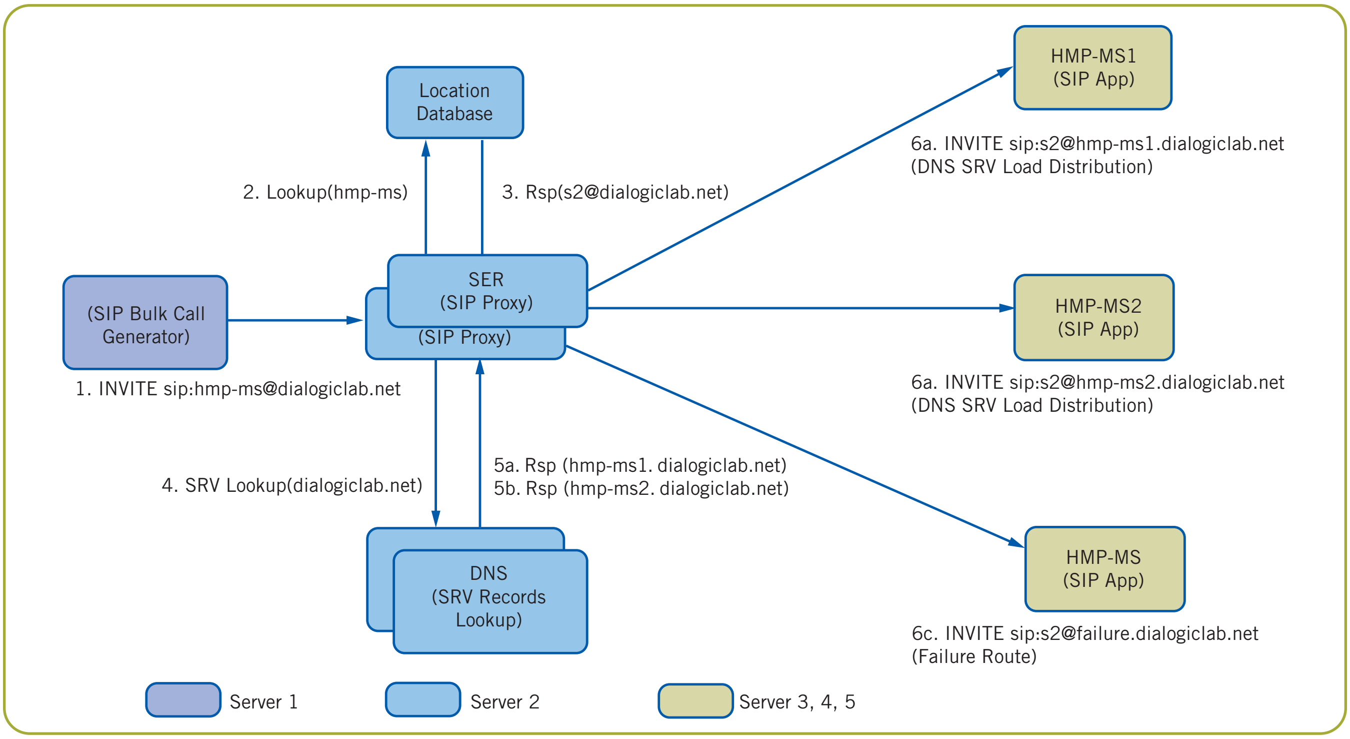
Location
Database
2. Lookup(hmp-ms)
3. Rsp(s2@dialogiclab.net)
HMP-MS1
(SIP App)
6a. INVITE sip:s2@hmp-ms1.dialogiclab.net
(DNS SRV Load Distribution)
(SIP Bulk Call
Generator)
SER
(SIP Proxy)
(SIP Proxy)
HMP-MS2
(SIP App)
1. INVITE sip:hmp-ms@dialogiclab.net
6a. INVITE sip:s2@hmp-ms2.dialogiclab.net
(DNS SRV Load Distribution)
4. SRV Lookup(dialogiclab.net)
5a. Rsp (hmp-ms1. dialogiclab.net)
5b. Rsp (hmp-ms2. dialogiclab.net)
DNS
(SRV Records
Lookup)
HMP-MS
(SIP App)
6c. INVITE sip:s2@failure.dialogiclab.net
(Failure Route)
Server 1
Server 2
Server 3, 4, 5

Figure 2. Test Configuration

Feature	Description
SBC	Dialogic® MPCBL0001N04 single board computer features a dual low voltage 2.2 GHz Intel Xeon processor and 1 GB RAM, 512 MB L2 cache, and 40GB HDD
OS	Red Hat Enterprise Linux (AS 3.0) Update 1 with 2.6.8.1 kernel
Server 1	Hosts the SIP bulk call generator SIPp, an open source test tool and traffic generator for the SIP protocol. For more information, visit http://sipp.sourceforge.net/.
Server 2	Hosts the SIP proxy SIP Express Router (SER). It also hosts the DNS server and the location database.
Server 3, 4, 5	Hosts the SIP Media server application running Dialogic HMP software. The application terminates an incoming SIP call and plays an announcement for five seconds.
HMP SW	Dialogic HMP Software Release 1.2 for Linux supports up to 240 G.711 media streams

Table 3. Test Configuration Details

DNS Configuration

The DNS entry must be modified if the calls for dialogiclab.net are to be distributed between hmp-ms1.dialogiclab.net and hmp-ms2.dialogiclab.net. The DNS entry for dialogiclab.net requires the following new A and SRV records:

```
Hmp-ms1    3600   IN   A    10.10.10.136
Hmp-ms2    3600   IN   A    10.10.10.155

_sip._udp.dialogiclab.net. 3600  IN SRV 0
1 5060  Hmp-ms1. dialogiclab.net.
_sip._udp.dialogiclab.net. 3600  IN SRV 0
1 5060  Hmp-ms2. dialogiclab.net.
```

Figure 5. Modified DNS Entry

Summary

A load balancing and fault handling approach that includes both the session protocol and DNS-SRV techniques can enable a high availability architecture using redundant SIP proxies, which alleviates the need to have redundant media servers and reduces costs. Since calls are spread across multiple servers, they are less likely to be lost or dropped due to hardware failures in a single server. A call can easily be redirected from a failed server to an active server.

Combining the SIP proxy with DNS-SRV techniques is an ideal way to develop higher density solutions using Dialogic HMP software.

Acronyms

AdvancedTCA	Advanced Telecom Computing Architecture
DNS-SRV	Domain Name Service-Service Record
HMP	Host Media Processing
IP	Internet Protocol
IVR	Interactive Voice Response
MGCP	Media Gateway Control Protocol
SBC	Single Board Computer
SER	SIP Express Router
SIP	Session initiation Protocol
TCP	Transmission Control Protocol
TTL	Transistor-Transistor Logic
UDP	User Datagram Protocol

For More Information

For more information about Dialogic Host Media Processing software, visit http://www.dialogic.com/products/ip_enabled/hmp_software.htm

To learn more, visit our site on the World Wide Web at **http://www.dialogic.com**

Dialogic Corporation
9800 Cavendish Blvd., 5th floor
Montreal, Quebec
CANADA H4M 2V9

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH PRODUCTS OF DIALOGIC CORPORATION OR ITS SUBSIDIARIES (“DIALOGIC”). NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN A SIGNED AGREEMENT BETWEEN YOU AND DIALOGIC, DIALOGIC ASSUMES NO LIABILITY WHATSOEVER, AND DIALOGIC DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF DIALOGIC PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY INTELLECTUAL PROPERTY RIGHT OF A THIRD PARTY.

Dialogic products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications.

Dialogic may make changes to specifications, product descriptions, and plans at any time, without notice.

Dialogic is a registered trademark of Dialogic. Dialogic's trademarks may be used publicly only with permission from Dialogic. Such permission may only be granted by Dialogic’s legal department at 9800 Cavendish Blvd., 5th Floor, Montreal, Quebec, Canada H4M 2V9. Any authorized use of Dialogic's trademarks will be subject to full respect of the trademark guidelines published by Dialogic from time to time and any use of Dialogic’s trademarks requires proper acknowledgement.

The names of actual companies and products mentioned herein are the trademarks of their respective owners. Dialogic encourages all users of its products to procure all necessary intellectual property licenses required to implement their concepts or applications, which licenses may vary from country to country.

This document discusses SIP Express Router, which is an open source product. Dialogic is neither responsible for your decision to use SIP Express Router (or any other open source product) in connection with Dialogic products including without limitation those referred to herein, nor is Dialogic responsible for any present or future effects such usage might have, including without limitation effects on your products, business, or intellectual property rights.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Dialogic products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.

05/07 9849-02