Oracle® Communications Diameter Signaling Router

IP Front End User Guide

Release 8.5.1
F51151-02
December 2022

ORACLE

Oracle Communications Diameter Signaling Router IP Front End User Guide, Release 8.5.1

F51151-02

Contents

Revision History

Release 8.5.1 - F51151-02 - December 2022

The following sections are updated for this release:

- Updated the description of Switch MAC Address Cache and Ping Feature in the following sections:
  - Traffic Distribution
  - Failure and Recovery Scenarios

Release 8.5.1 - F51151-01 - December 2022

No updates in this release.

List of Figures

List of Tables

1 Introduction

This *IP Front End (IPFE) User's Guide* and Help provide an overview of IPFE functions and procedures to use to configure IPFE. The contents include sections on the scope, audience, and organization of the documentation, and how to contact My Oracle Support for assistance.

Revision History

Release 8.5.1 - F51151-02 - December 2022

The following sections are updated for this release:

- Updated the description of Switch MAC Address Cache and Ping Feature in the following sections:
  - Traffic Distribution
  - Failure and Recovery Scenarios

Release 8.5.1 - F51151-01 - December 2022

No updates in this release.

Overview

The **IPFE** documentation provides information about **IPFE** functions, how to use the GUI and the following procedures to configure an **IPFE**:

- Specify **IPFE** Configuration Options
- Configure **IPFE** Target Sets

Scope and Audience

The IP Front End (IPFE) documentation is intended for anyone responsible for the configuration of the IPFE. Users of this guide must have a working knowledge of telecommunications, of network installations and the product that uses the IPFE functions.

Manual Organization

This manual is organized into the following chapters:

- Introduction contains general information about the IPFE help documentation, the organization of this manual, and how to get technical assistance.
- Introduction to IPFE provides information about the IPFE function.
- IPFE Configuration Options describes how to manage your IPFE configuration.
- IPFE Target Sets Configuration describes how to assign a list of application server IP address to a Target Set and associate the Target Set with an IPFE pair.

My Oracle Support

My Oracle Support (https://support.oracle.com) is your initial point of contact for all product support and training needs. A representative at Customer Access Support can assist you with My Oracle Support registration.

Call the Customer Access Support main number at 1-800-223-1711 (toll-free in the US), or call the Oracle Support hotline for your local country from the list at http://www.oracle.com/us/support/contact/index.html. When calling, make the selections in the sequence shown below on the Support telephone menu:

1. Select **2** for New Service Request
2. Select **3** for Hardware, Networking and Solaris Operating System Support
3. Select one of the following options:
    - For Technical issues such as creating a new Service Request (SR), select **1**
    - For Non-technical issues such as registration or assistance with My Oracle Support, select **2**

You are connected to a live agent who can assist you with My Oracle Support registration and opening a support ticket.

My Oracle Support is available 24 hours a day, 7 days a week, 365 days a year.

2
Introduction to IPFE

The IP Front End (IPFE) is a traffic distributor that transparently does the following:

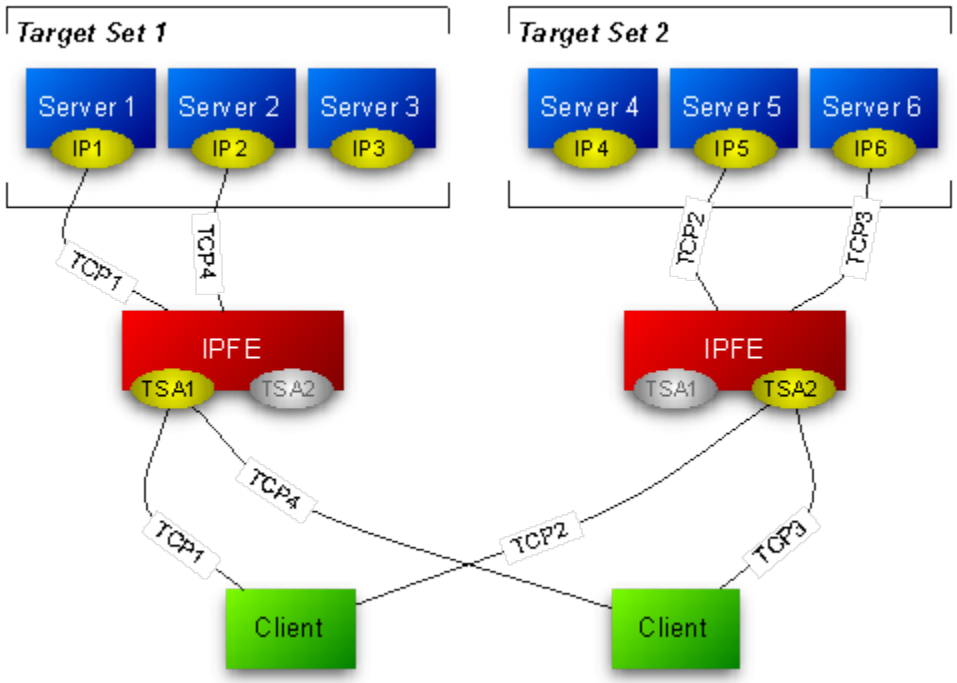
- Presents a routable IP address representing a set of up to 16 application servers to application clients. This reduces the number of addresses with which the clients need to be configured.
- Routes packets from the clients that establish new TCP or SCTP connections to selected application servers.
- Routes packets in existing TCP or SCTP connections to the correct servers for the connection.

IPFE Description

The IPFE acts as a specialized layer-3 router. The various servers to which the IPFE routes packets are divided into up to 16 groups, called Target Sets. Each of the target sets are assigned a shared Target Set Address, a publicly exposed service address.

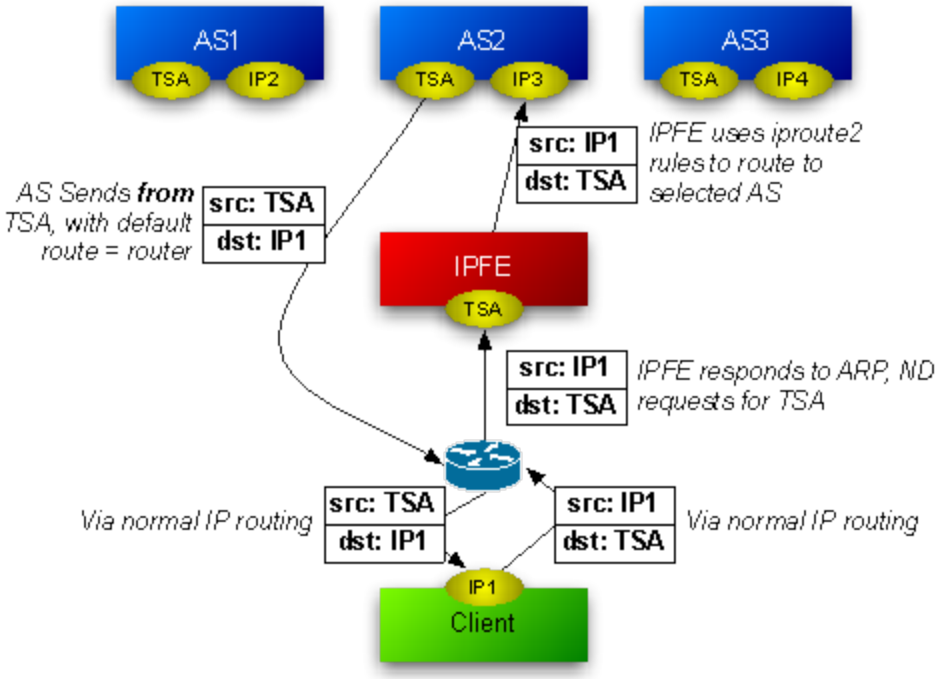
Figure 2-1 shows either two connections are maintained at all times, in active/active or active/standby, or that a single connection is maintained, with a backup address for clients to establish a connection, if the first connection fails.

Figure 2-1 IPFE Architecture



When the IPFE routes packets to application servers, it does not perform any rewriting of the packet. Figure 2-2 shows that neither the source IP address nor the destination IP address changes as it passes through the IPFE. The IPFE behaves as an IP router and does not act as a network address translator (NAT).

Figure 2-2 Packet Routing Through and Around the IPFE



Traffic Distribution

The IPFE is a packet-based load balancer that makes a large cluster accessible to incoming connections through a minimal number of IP addresses. These incoming connections can be TCP, unihomed SCTP, or multihomed SCTP. The IPFE distributes these connections among a list of target IP addresses by forwarding incoming packets. The list is called the **Target Set IP List**, and an outward-facing IP address is called a Target Set Address (**TSA**). A packet arriving at the IPFE and destined for the TSA is forwarded to an address in the Target Set IP List.

There can be as many as 16 IP addresses in the target set IP list and thus the IPFE may distribute traffic among as many as 32 physical or virtual application servers. Each server in the target set IP list can have a **Weighting** indicating that the IPFE should apportion more or fewer connections to that server. The load balancing algorithm for apportioning connections is also configurable through a number of settings. The TSA, target set IP list, weighting, and load balancing algorithm settings are together called a **Target Set**. There can be as many as 32 independent target sets configured on one IPFE.

The IPFE neither interprets nor modifies anything in the TCP or SCTP payload. The IPFE also does not maintain TCP or SCTP state, but keeps sufficient state to route all packets for a particular session to the same application server.

Return traffic from the application server to the client (both TCP and SCTP) does not pass through the IPFE, but routes directly to the gateway.

Switch MAC Address Cache and Ping Feature

In a certain deployments where all traffic passes through the IPFE, no Ethernet packets go directly to the DA-MP from the gateway (or remote peer, for the case that a remote peer is on the local network segment). Rather, all Ethernet packets come to the DA-MP by way of the IPFE. Any intermediate Switch would be unaware that the Ethernet jack ("switch port") of the gateway (or peer) is a viable path for packets emitted by the DA-MP. In this case, the Switch would broadcast that packet to all Ethernet switch ports as a last resort. This creates network flooding.

For this situation, even if the switch had knowledge of the aforementioned switch port, this information expires after five minutes on typical switch configurations.

The solution to this problem is to keep the switch tables up-to-date with periodic pings to remote peers or gateways. An ICMP or ARP ping every two minutes, from the DA-MPs, is sufficient.

To run the ping on a particular DA-MP, login as root and run

```
/usr/TKLC/dsr/bin/pingAllLivePeers -v
```

Use pingAllLivePeers -h for options. These commands can be used for diagnostics. Note that background operation logs to /var/log/messages and /var/log/cron.

High Availability

The IPFE supports active-standby or active-active high availability (HA) when paired with a second IPFE instance. The mated pair of IPFEs expose typically one or two TSAs per configured IP version.

Each TSA can operate in an active-standby mode, where all traffic to a given TSA goes to the active (for that TSA) IPFE, if it is available. If the active IPFE fails or if its mate is explicitly selected as Active, traffic to the TSA goes to the mate IPFE. For active-active HA, the addresses must be configured in pairs, where one IPFE is active for one address in a pair, and the mate is active for the other.

Note that the IPFE supports more than two TSAs, and in fact when both IPv4 and IPv6 are supported, the IPFE is usually configured with at least four TSAs. An IPFE and its mate are numbered 1 and 2, whereas an IPFE pair is numbered A and B. The four IPFEs are numbered A1, A2, B1, and B2.

For multi-homed SCTP connections, the **Target Set** is represented by both a public address and an alternate address. Each application server in the Target Set must also be configured for multihomed SCTP.

IPFE Associations

The IPFE stores an Association record about each connection. The Association contains the information necessary to identify packets belonging to a connection and to identify the application server that the IPFE has selected for the connection. The IPFE forwards all packets associated with a particular connection to the selected application server.

The specific packet-identifying information is the source IP address and the source port number. For each target set, packets matching both by source address and source port are routed to the same target application server. SCTP verification tags are also used as identifiers since, with the SCTP protocol, the source IP address can change.

All association information is replicated between mated IPFEs, but not between IPFE pairs.

Association information is isolated to a target set so that the target sets behave independently.

Because returning packets bypass the IPFE, the IPFE has limited knowledge of the state of the connection. The IPFE cannot determine if a connection has reconnected from the same source port, nor whether the connection has been terminated. The IPFE attempts to use the available state information to make the best possible judgments about when an association is stale. A stale connection is removed and subsequent packets originating from the same IP address and from the same source port are treated as a new connection: the load balancing algorithm is freshly applied.

An association is considered stale if:

- No packets have been received for the duration of the **Delete Age** setting in the **Target Set** configuration.
- The transactions of the `Connect-CER-CEA-Disconnect` form are the only transactions to have taken place for a period of time of Delete Age.
- The IPFE is able to track the TCP sequence numbers and determined if an authentic FIN and subsequence SYN are in evidence that a TCP connection has disconnected and reconnected. This tracking works for certain idealized TCP connections only.
- The IPFE is able to track the SCTP verification tag and determined if an authentic SHUTDOWN and subsequence INIT are in evidence that a SCTP connection has disconnected and reconnected. This tracking works for certain idealized SCTP connections only.

Load Balancing

If a packet is not matched by any association, the IPFE creates a new association by choosing an application server from the target set IP list. The choice is based on the load balance algorithm setting. The IPFE is designed to keep the connection on the same MP across reconnection in a period, if possible. This enables the upper layer transaction to be complete after reconnection and minimizes the impact to other MPs for a bouncing case. If the original application server is not available, reconnecting connections is distributed to other application servers available. However, after the unavailable application server recovers, the connections are not redistributed back for continuity purpose, so ongoing traffic is not disturbed.

Regardless of the algorithm, the IPFE raises a minor alarm of `Out of Balance: High` or `Out of Balance: Low` on an application server whenever it is receiving a statistically high or low amount of traffic in comparison to others within the same target set.

If an application server determines that it has reached fully loaded capacity, then it notifies the IPFE not to send it further new connections. This is called Stasis.

Application servers may go in and out of Stasis automatically according to the current traffic.

There are two load balance algorithms available:

- Hash: load balancing achieves by sending the new connection to a server based on hashing the originating port and IP address. Hash load balancing removes an application server from consideration for new connections whenever it is incurring an `Out of Balance: High` alarm. In this way reconnecting connections are always directed to application servers that are moderately loaded. This feature is independent of Stasis notifications.
- Least load : chooses the server with the least load as reported by the application server. If the loads of two or more of the least-loaded servers are within a configurable percentage of each other, they are considered equally loaded, and the IPFE distributes connections to them in a round-robin fashion. By using the load data reported from the application server, IPFE can better manage the actual traffic load on each MP, although this may introduce some latency between the distribution and the actual situation. In some extreme conditions, such as huge burst of traffic, the latency might cause uneven distribution of a newly assigned connection. However, the application reported data can benefit for cases like different server capabilities and other traffic assignment (not TSA traffic), which is more common than corner cases for latency issue.

IPv4 and IPv6 support

A Target Set can be created as either IPv4 or IPv6. However a target set cannot support mixed address types. This means that SCTP multi-homed endpoints can contain address types of either IPv4 or IPv6 but not both.

Throttling

In the case of signaling storms, the IPFE provides a configurable parameter which limits the IPFE's throughput rate and prevents the maxing out of its CPU. Throttling causes the IPFE to drop packets in order to keep the load from overwhelming the IPFE. The packet/second rate limit implementation creates an even dropping of packets that would cause client TCP/SCTP stacks to withhold their rates to just below the threshold, as happens when there is an overloaded router in the path. Throttling is on per-local-port bases, for example, each local port (such as 3868) is apportioned the configured amount.

Failure and Recovery Scenarios

An IPFE that has a mate and at least two target set addresses can handle different failure and recovery scenarios.

> **Note:**
>
> The following failover scenarios describe what happens with the IPFE-A1 and IPFE-A2 pair. A failover involving the IPFE-B1 and IPFE-B2 pair is handled exactly the same way.

This section discusses how the following IPFE setup can gracefully handle the failure and recovery of various components in the system:

- Two IPFEs, IPFE-A1 and IPFE-A2, each responsible for one target set address. IPFE-A1 is public for TSA1, and IPFE-A2 is public for TSA2.
- Two target sets, each with three application servers and the target set addresses TSA1 and TSA2.
  - TSA1 has application servers Server1, Server2, and Server3
  - TSA2 has application servers Server4, Server5, and Server6
- Two clients, each configured with TSA1 and TSA2.

These failure and recovery scenarios apply to a single component outage.

IPFE Failure and Recovery

If IPFE-A1 fails, the system handles it in the following manner:

- IPFE-A1's mate, IPFE-A2, detects the failure.
- IPFE-A2 takes over IPFE-A1's TSA, TSA1.
- There are no changes to the application servers in TSA1. TSA1 continues to comprise Server1, Server2, and Server3
- Traffic for TSA1 continues to go to TSA1, which is now managed by IPFE-A2
- IPFE-A2 continues to route TSA1 traffic to Server1, Server2, and Server3 - no different than they were before the failure.
- IPFE-A2 also continues to route traffic for TSA2 to Server4, Server5, and Server6.
- No disruption of service occurs.
- New connection requests for TSA1 is routed to Server1, Server2, or Server3.
- New connection requests for TSA2 is routed to Server4, Server5, or Server6.

When IPFE-A1 recovers, the following happens:

- IPFE-A2 detects that IPFE-A1 has recovered and relinquishes control of TSA1.
- IPFE-A1 assumes control of TSA1.
- Traffic that went to TSA1 continues to go to TSA1.
- The clients are unaware that a recovery has occurred.
- New connection requests for TSA1 continue to be routed to Server1, Server2, or Server3.
- New connection requests for TSA2 continue to be routed to Server4, Server5, or Server6.

Application Server Failure and Recovery

When an application server, for example, Server1, fails, then the following occurs:

- The connections from the client fail.
- Other connections through TSA1 to Server2 and Server3 remain established.
- When a connection breaks due to an application server failure, and the peer tries to reestablish the connection on the same TSA, then the connection is delegated to another available application servers in the same TSA.

- IPFE-A1 routes new connection requests to the remaining application servers (Server2 and Server3). If all application servers in a target set fail, and IPFE-A1 receives a request for a new connection to TSA1, then it alternatively notifies the client that the request cannot be fulfilled using either a TCP RST packet (for TCP connections) or a configurable ICMP message.

When Server1 recovers:

- IPFE-A1 detects Server1's availability.
- IPFE-A1 routes new connection requests to Server1.
- Some imbalance across application servers in TSA1 exists after recovery. IPFE-A1 monitors for imbalances in traffic and distributes new connections to reduce the imbalance.

Switch MAC Address Cache and Ping Feature

In certain deployments where all traffic passes through the IPFE, no Ethernet packets go directly to the DA-MP from the gateway (or remote peer, for the case that a remote peer is on the local network segment). Rather, all Ethernet packets come to the DA-MP by way of the IPFE. Any intermediate Switch would be unaware that the Ethernet jack (switch port) of the gateway (or peer) is a viable path for packets emitted by the DA-MP. In this case, the Switch would broadcast that packet to all Ethernet switch ports as a last resort. This creates network flooding.

For this situation, even if the switch had knowledge of the aforementioned switch port, this information expires after five minutes on typical switch configurations.

The solution to this problem is to keep the switch tables updated with periodic pings to remote peers or gateways. An ICMP or ARP ping every two minutes, from the DA-MPs, is sufficient.

To run the ping on a particular DA-MP, login as root and type:

```
/usr/TKLC/dsr/bin/pingAllLivePeers -v
```

For options, type:

```
pingAllLivePeers -h
```

These commands can be used for diagnostics.

> **Note:**
>
> The background operation logs to /var/log/messages and /var/log/cron

Enclosure Failure and Recovery

In the enclosure failure scenario we assume that the IPFE is co-located with the application servers in its target set. In this case, IPFE-A1 is in an enclosure with Server1, Server2, and Server3.

When the enclosure containing IPFE-A1, Server1, Server2, and Server3 fails:

- All connections to all servers in the enclosure fail.
- IPFE-A2 detects that IPFE-A1 is down and starts servicing TSA1.
- Clients with existing connections to TSA1 detect that TSA1 is unavailable and send traffic to TSA2.
- Depending on configuration, IPFE-A2 optionally sends a TCP RST (for TCP connections) or a configured ICMP message in response to client connection requests to TSA1.

When the enclosure recovers:

- IPFE-A2 detects that IPFE-A1 has recovered and relinquishes control of TSA1.
- IPFE-A1 takes over control of TSA1.
- Since TSA1 did not have any existing connections during the failure, no special handling of existing connections is required.
- Over a time, clients are expected to route new connections to TSA1, resulting in connections to recovered servers in the associated target set.
- In the interim, there is a substantial imbalance between the two IPFEs as well as between the servers in the two TSAs. The IPFEs monitor the traffic for imbalances and distribute new connections to reduce the imbalance.

External Connectivity Failure and Recovery

If external connectivity to the IPFE, say IPFE-A1, fails:

- Connections to IPFE-A1 and TSA1 fail.
- IPFE-A2 does not take over TSA1 since it sees IPFE-A1 as available. That is, internal connections still work.
- Clients with failed connections to TSA1 must send traffic to TSA2.
- Clients attempting to create new connections to TSA1 fail.
- IPFE-A2 and TSA2 carry all the traffic for all the clients.

When external connectivity is restored:

- There are no existing connections for TSA1 to handle.
- IPFE-A1 still retains control over TSA1.
- Clients route new connections to TSA1 over time.
- In the interim, there is a substantial imbalance between the two IPFEs as well as between the servers in the two TSAs. The IPFEs monitors the traffic for imbalances and distribute new connections to reduce the imbalance.

Bulk Import and Export

The *Diameter Common User's Guide* describes the use and operation of Bulk Import and Export functions:

- **Help**, and then **Diameter Common**, and then **Bulk Import**
- **Help**, and then **Diameter Common**, and then **Bulk Export**

The Bulk Import and Export functions can be used to export Diameter, IPFE, and Application configuration data in CSV files to a location outside the system, and to import the files (usually edited) into the system where the Import function is executed.

Bulk Import

The Bulk Import operations use configuration data in ASCII Comma-Separated Values (CSV) files (.csv), to insert new data into, update existing data in, or delete existing data from the configuration data in the system.

> **Note:**
>
> Some configuration data can be imported only with the Update operation, and other data can be imported with Insert and Delete operations but not Update. Refer to the *Diameter Common User's Guide* or the **Diameter Common**, and then **Import** Help for valid Import operations.

Import CSV files can be created by using a Bulk Export operation, or can be manually created using a text editor.

> **Note:**
>
> The format of each Import CSV file record must be compatible with the configuration data in the release used to import the file. Across different release versions, column counts may not be compatible, and the import fails.

Files that are created using the Bulk Export operation can be exported either to the local Status & Manage File Management Directory (**Status & Manage**, and then **Files** page), or to the local Export Server Directory.

CSV files that are in the local File Management area can be used for Bulk Import operations on the local system.

Files can be created manually using a text editor; the files must be uploaded to the File Management area of the local system before they can be used for Import operations on the local system.

Multiple Import operations can be performed:

- Insert new configuration data records that do not currently exist in the system
- Update existing configuration data in the system
- Delete existing configuration data from the system

Each Import operation creates a log file. If errors occur, a Failures CSV file is created that appears in the File Management area. Failures files can be downloaded, edited to correct the errors, and imported to successfully process the records that failed. Failures files that are unchanged for more than 14 days and log files that are older than 14 days are automatically deleted from the File Management area.

Bulk Export

The Bulk Export operation creates ASCII Comma-Separated Values (CSV) files (.csv) containing Diameter, IPFE, and Application configuration data. Exported configuration data

can be edited and used with the Bulk Import operations to change the configuration data in the local system without the use of GUI pages. The exported files can be transferred to and used to configure another system.

Each exported CSV file contains one or more records for the configuration data that was selected for the Export operation. The selected configuration data can be exported once immediately, or exports can be scheduled to periodically occur automatically at configured times.

Configuration data can be exported in one Export operation:

- All exportable configuration data in the system
- All exportable configuration data from the selected Application, IPFE, or Diameter (each component's data is in a separate file)
- Exportable configuration data from a selected configuration component for the selected Application, IPFE, or Diameter

Exported files can be written to the File Management Directory in the local File Management area (**Status & Manage**, and then **Files** page), or to the Export Server Directory for transfer to a configured remote Export server.

CSV files that are in the local File Management area can be used for Bulk Import operations on the local system.

If the export has any failures or is unsuccessful, the results of the export operation are logged to a log file with the same name as the exported file but with a .log extension. Successful export operations are not logged.

3
IPFE Configuration Options

The **IPFE**, and then **Configuration**, and then **Options** page allows you to manage IPFE configuration.

Configuration Options Elements

Table 3-1 describe the fields on the **IPFE**, and then **Configuration**, and then **Options** page. An asterisk before the value field means the configuration is mandatory.

Table 3-1 IPFE Configuration Elements

Element	Description	Data Input Notes
	Inter-IPFE Synchronization	
IPFE-A1 IP Address	This address must reside on the internal management interface (IMI) network. This address is used for replicating association data between IPFEs and is not exposed to application clients. If left blank, the IPFE does not replicate association data. Although optional, this configuration is required for a fully functioning installation.	Format: IPv4 or IPv6 address, or left blank Default: blank
IPFE-A2 IP Address	This address must reside on the internal management interface (IMI) network. This address is used for replicating association data between IPFEs and is not exposed to application clients. If left blank, the IPFE does not replicate association data. Although optional, this configuration is required for a fully functioning installation.	Format: IPv4 or IPv6 address, or left blank Default: blank
IPFE-B1 IP Address	This address must reside on the internal management interface (IMI) network. This address is used for replicating association data between IPFEs and is not exposed to application clients. If left blank, the IPFE does not replicate association data. Although optional, this configuration is required for a fully functioning installation.	Format: IP address, or left blank Default: blank

Table 3-1 (Cont.) IPFE Configuration Elements

Element	Description	Data Input Notes
IPFE-B2 IP Address	This address must reside on the internal management interface (IMI) network. This address is used for replicating association data between IPFEs and is not exposed to application clients. If left blank, the IPFE does not replicate association data. Although optional, this configuration is required for a fully-functioning installation.	Format: IP address, or left blank Default: blank
* State Sync TCP Port	This port establishes and maintains a connection to its mate. If the connection is lost, it attempts to re-establish the connection based on the configuration of the state sync reconnect interval.	Format: text box Range: 1-65535 Default: 19041
* State Sync Reconnect Interval	Reconnect interval, in seconds, for syncing kernel state between IPFEs.	Format: text box Range: 1-255 seconds Default: 1
* Gratuitous ARP Type	Specify type of gratuitous ARP broadcast to send.	Format: ARP Request, ARP Reply, Send both types Default: ARP Request
	Traffic Forwarding	
* Application Traffic TCP Reject Option	How to reject TCP connections when no application servers are available. When no application servers are available, the IPFE must reject the TCP traffic that it receives. The IPFE can either drop packets or it can communicate to the application clients with TCP or ICMP messages. Select the option that can be best handled by the application client.	Format: list Range: • TCP Reset • Drop Packet • ICMP Host Unreachable • ICMP Port Unreachable • ICMP Administratively Prohibited Default: TCP Reset
* Application Traffic SCTP Reject Option	How to reject SCTP connections when no application servers are available. When no application servers are available, the IPFE must reject the SCTP traffic that it receives. The IPFE can either drop packets or it can communicate to the application clients with ICMP messages. Select the option that can be best handled by the application client.	Format: list Range: • Drop Packet • ICMP Host Unreachable • ICMP Port Unreachable • ICMP Administratively Prohibited Default: ICMP Host Unreachable
	Packet Counting	

Table 3-1 (Cont.) IPFE Configuration Elements

Element	Description	Data Input Notes
* Imbalance Detection Throughput Minimum	This value applies only to the hash algorithm selection. This is the value below which no throughput analysis is performed regarding the distribution of connections. This setting should not be changed from its default unless the IPFE is being tested with a very low load. This setting ensures the IPFE does not mark application servers as imbalanced when it is distributing very few messages between them.	Format: text box Range: 1-2147483647 Default: 20000
* Least Load Threshold	This value can be set to a packets-per-second rate below which the Least Load algorithm reverts to round robin.	Format: text box Range: 1-2147483647 Default: 1
* Cluster Rebalancing and Accounting	Support for cluster rebalancing and packet accounting in measurements. When this is disabled, all accumulation of packet and byte measurements cease. Overload detection also stops. The disabled state is useful only for troubleshooting, which should be done by My Oracle Support . Contact My Oracle Support before disabling measurements and overload detection.	Format: list Range: • Enabled • Disabled Default: Enabled
	Application Server Monitoring	
* Monitoring Port	TCP port to try periodic connections or monitoring of application servers. The IPFE opens a TCP connection to the application server's IP address and this port. The application server must listen on this port and should send heartbeats.	Format: text box Range: 1-65535 Default: 9675

Table 3-1 (Cont.) IPFE Configuration Elements

Element	Description	Data Input Notes
* Monitoring Connection Timeout	How long to wait for a connection to complete when polling the application servers for aliveness in seconds. If the IPFE detects that an application server has missed a configurable number of heartbeats - that is, more than that number of seconds have elapsed since the most recent heartbeat was received - then it considers the application server to be down. The IPFE removes a down application server from the traffic balancing pool and attempts to reconnect to the server.	Format: text box Range: 1 - 255 Default: 3
Monitoring Connection Try Interval	Interval in seconds of periodically connecting to application servers to test for aliveness. While an application server is down, the IPFE periodically attempts to reconnect to it based on this configuration.	Format: text box Range: 1 - 255 Default: 10

Table 3-1 (Cont.) IPFE Configuration Elements

Element	Description	Data Input Notes
Monitoring Protocol	Application liveness monitoring method. If any Target Set has load balancing of **Least Load**, then this setting cannot be changed from **Heartbeat** due to the need for load information in the monitoring packets. The monitoring protocol allows the IPFE to determine the liveness of the application servers. The IPFE determines this either by listening for heartbeat messages from the application servers. When the protocol is set to Heartbeat, the IPFE connects to the monitoring port, sustains the connection, and receives heartbeat packets from the application server. In this case, the failure to receive a heartbeat packet within the period **Back-end Connection Timeout** indicates the server is dead. A dead server is removed from the traffic balancing pool. The IPFE attempts connections on the monitoring port until the server responds. When the server responds, the IPFE adds it back to the pool.	Format: list Range: • Heartbeat • None Default: Heartbeat
	Throttling and DoS Protection	
Global Packet Rate Limit	Combined packet rate limit for a single IPFE at which overload throttling is applied.	Format: text box Range: 10000 - 10000000 Default: 500000

Configuring the IPFE

The **IPFE**, and then **Configuration**, and then **Options** page set up data replication between IPFEs, specify port ranges for TCP traffic, and set application server monitoring parameters.

1. Click **IPFE**, and then **Configuration**, and then **Options**.

   The fields are described in Table 3-1.

2. Under the Inter-IPFE Synchronization section complete the following entries:

   a. Enter the IP addresses for IPFE-A1, IPFE-A2, IPFE-B1, and IPFE-B2 in the corresponding **IPFE-Xn IP Address** field.

      These are internal addresses used by the IPFEs to replicate association data. These addresses should reside on the IMI network.

b. Select the **State Sync TCP Port** used for syncing kernel state between IPFE.

c. Set the **State Sync Reconnect Interval** for syncing kernel state between IPFE.

d. Select the type of **Gratuitous ARP Type** broadcast to send.

3. Under the Traffic Forwarding section complete the following entries:

   a. If not application servers are available, select how to reject TCP connections in the **Application Traffic TCP Reject Option** list.

   b. If not application servers are available, select how to reject SCTP connections in the **Application Traffic SCTP Reject Option** list.

4. Under the Packet Counting section complete the following entries:

   a. Set a value for **Imbalance Detection Throughput Minimum**.

      The default setting is 20000 and should not be changed from its default unless the IPFE is being tested with a very low load.

   b. Set a value for **Least load Threshold**.

      This value can be set to a packets-per-second rate below which the Least Load algorithm reverts to round robin.

   c. If you want the support for cluster re-balancing and packet accounting in measurements, select **Enabled** from the list in the **Cluster Rebalancing and Accounting** field.

      When this is disabled, all accumulation of packet and byte measurements cease.

5. Under the Application Server Monitoring section complete the following entries:

   a. Set a value for **Monitoring Port**

      The application server must listen on this port and should send heartbeats.

   b. Set the wait time for a connection to complete when polling the application servers for aliveness in the **Monitoring Connection Timeout** field

   c. Set the interval for periodically connecting to application servers to test for aliveness in the **Monitoring Connection Try Interval** field.

   d. If you want to monitor the liveness of application servers, select **Heartbeat** in the **Monitoring Protocol** field.

      If any Target Set has load balancing of Least Load, then this setting cannot be changed from **Heartbeat** due to the need for load information in the monitoring packets.

6. Under the Throttling and DoS protection section complete the following entry:

   - Set the overload throttling in the **Global Packet Rate Limit** field.

7. Click **Apply** or **Cancel**

For the IPFE to be fully functional, you must assign application servers to a Target Set and associate the Target Set with the IPFE. See Add a Target Set to add a new Target Set.

4
IPFE Target Sets Configuration

The **IPFE**, and then **Configuration**, and then **Target Sets** page allows you to assign a list of application server IP addresses to a target set and associate the target set with an IPFE pair.

Target Sets Configuration Elements

A Target Set associated with an IPFE maps a single externally available IP address to a set of IP addresses for application servers.

In general, it is inadvisable to reduce delete age value to less than the default. However, a TSA that has connections with longer SCTP heartbeat interval may require this value to be increased from default.

The Table 4-1 describes the fields on the **IPFE**, and then **Configuration**, and then **Target Sets** page.

The Table 4-2 describes the fields on the view, insert, and edit pages.

Table 4-1 Target Sets configuration elements (View pages)

Field	Description	Data Input Notes
Target Set Number	Unique ID identifying the target set.	Format: Numeric Range: 1-32
Target Set Address	Public IP address to present to the outside world.	Format: IPv4 or IPv6 address The target set address must be on the XSI network
Target Set IP List	List of IP addresses of the associated application servers.	Format: IPv4 or IPv6 address IP address type must match that of the target set Address. The IP addresses in target set IP list must be on the XSI network.
Weighting	Weighting value is used to apportion load between application servers within the target set.	Format: Numeric Range: 0-65535 Default: 100
Supported Protocols	The protocols supported by this target set.	Format: Options Range: TCP only, SCTP only, Both TCP and SCTP Default: Both TCP and SCTP

Table 4-1 (Cont.) Target Sets configuration elements (View pages)

Field	Description	Data Input Notes
Preferred Active	The IPFE that primarily handles traffic for this target set. Disabled means that the target set is defined, but not currently in use by an IPFE.	Format: Options Range: IPFE-A1, IPFE-A2, IPFE-B1, IPFE-B2 Default: IPFE-A1 If an option is not activate, you need configure the IPFE address under **IPFE**, and then **Configure**, and then **Options**.
Preferred Standby	The mate of the Preferred Active IPFE. If the Preferred Active IPFE is unavailable, the Preferred Standby server takes over.	If the preferred standby IPFE has been configured, it is set when you select the preferred active IPFE.

Table 4-2 Target Sets configuration elements (Insert and Edit pages)

Field	Description	Data Input Notes
	Target Set	
* TS Number	Unique ID identifying the TSA.	Format: List Range:1-32 Default: 1
Protocols	A target set can support SCTP, TCP, or both.	Format: Options Range: TCP only, SCTP only, Both TCP and SCTP Default: Both TCP and SCTP
Disable	Select to disable this target set, but preserve it in this configuration.	Format: Checkbox Range: Disable
* Delete Age	Connections are dropped if idle for this time (seconds). When setting this value please take into account that TCP connections can sometimes be idle for long periods of time depending on the application protocol.	Format: Text box, numeric Range: 10 - 3110400 Default: 600

Table 4-2 (Cont.) Target Sets configuration elements (Insert and Edit pages)

Field	Description	Data Input Notes
Load Balance Algorithm	Algorithm used to determine where new connections should go. Hash: load balancing by sending the new connection to a server based on hashing the originating port and IP address. Least Load: load balancing by choosing the server with the least load as reported by the application server. (Requires Monitoring Protocol to be set to Heartbeat.) The load of an application server is calculated using the load equation: $L(m,c) = (F_m * m/m_{total} + F_c * c/_{ctotal}) * W_{high}/w$ where m and m_{total} are the currently reserved and total capacity of ingress MPS (messages per second), respectively; c and c_{total} are the number of current connections and total connection capacity, respectively; w and w_{high} are the application server weighing and the highest weighting in the Target IP List, respectively. The value c includes, as an added component, the rate of new connections, in order to smooth the distribution of a sudden flood of new connections.	Format: Options Range: Hash, Least Load Default: Least Load
	Least Load Parameters	
* MPS Factor	Factor F_m in load equation. The total $F_m + F_c$ is normalized to 100 on commit of this form.	Format: Text field; numeric Range: 0 - 100 Default: 50
* Connection Count Factor	Factor F_c in load equation. The total $F_m + F_c$ is normalized to 100 on commit of this form.	Format: Text field; numeric Range: 0 - 100 Default: 50

Table 4-2 (Cont.) Target Sets configuration elements (Insert and Edit pages)

Field	Description	Data Input Notes
* Allowed Deviation	Percentage within which two application servers' L(m,c) results are considered to be equal, which is used to smooth out load distribution. If the difference in load between the lowest and next least-loaded application server is greater than or equal to this value, then the IPFE applies the Least Load algorithm and assigns new connections to the least loaded application server. If the difference in load between the lowest and next least-loaded application server is less than this value, then the IPFE distributes the connection in a weighted round-robin fashion between the application servers that are within the Allowed Deviation range.	Format: Text field; numeric Range: 0 - 50 Default: 5
Peer Node Aware Least Load	Enable peer node group awareness when directing connections. When enabled, the IPFE distribute connections from the same peer node group across servers in the target set to provide server redundancy for that group of peers. The IPFE keeps a group count of the connections from a peer node group to each server in the target set. Servers with a group count difference that is equal to or greater than `D` from the lowest group count is generally not considered, such as, if `D` is 1, the effect is to send the connection to the server with the lowest group count.	Format: Checkbox Default: Enable
Peer Node Group Distribution Threshold	The value `D` in Peer Node Aware	Format: Text field; numeric Range: 1 - 10 Default: 1
	Public IP Address	
* Address	Public IP address presented to the outside world. Do not edit if in use by a local node.	Format: IPv4 or IPv6 address

Table 4-2 (Cont.) Target Sets configuration elements (Insert and Edit pages)

Field	Description	Data Input Notes
Active IPFE	IPFE that primarily handles traffic for this TSA. If the active IPFE fails, then its mate takes over. IPFE-A1 and IPFE-A2 are mates. IPFE-B1 and IPFE-B2 are mates. If these options are disabled, IPFE Addresses under IPFE>Configuration>Options need to be configured.	Format: Options
	Alternate Public IP Address	
Alternate Address	Optional alternate public IP address presented to the outside world. For SCTP, this address serves as an non-primary protocol-linked failover address. For TCP, this address can serve as an independent address. If this field is populated, then the column alternate IP address target set IP List must be populated. Do not edit if in use by a local nodes.	Format: IPv4 or IPv6 address
Active IPFE for Alternate address IPFE	The IPFE that primarily handles traffic for this TSA's alternate address. If the active IPFE fails, then its mate takes over. IPFE-A1 and IPFE-A1 are mates. IPFE-B1 and IPFE-B2 are mates. The setting for this field should complement the setting of Active IPFE in order to provide and alternative path for SCTP dual-homed traffic. This allows SCTP connections with a very short heartbeat interval to transmit on the alternate path if the heartbeat timeout is short than the IPFE switchover delay.	Format: Options
	Target Set IP List	
IP Address	Public IPv4 or IPv6 address for the application server.	Format: List
Alternate IP Address	Alternate IP address for the application server.	Format: List

Table 4-2 (Cont.) Target Sets configuration elements (Insert and Edit pages)

Field	Description	Data Input Notes
Description	Free-form description for the application server.	Format: Text
* Weighting	Weighting value used to apportion load between application servers within the target set. The following formula determines the selection of an application server: Application server's % chance of selection = (Application server weight / Sum of all weights in the target set) * 100. If all application servers have an equal weight, they have an equal chance of being selected. If application servers have unequal capacities, give a higher weight to the servers with the greater capacity.	Format: Text field; numeric Range: 0 - 65535

Add a Target Set

Before you can add a Target Set, you must configure at least one IPFE in **IPFE**, and then **Configuration**, and then **Options**.

Use this task to add a target set to the IPFE configuration. Define the list of application server IP addresses for the target set and associate the target set with an IPFE.

Target Sets associated with an IPFE may be completely overlapping, but may not be partially overlapping. A warning appears if overlapping target sets are associated with an IPFE.

Partially overlapping target set example:
Target Set 1: Application Server 1, Application Server 2
Target Set 2: Application Server 2, Application Server 3

Completely overlapping target set example:
Target Set 1: Application Server 1, Application Server 2
Target Set 2: Application Server 1, Application Server 2

1. Select **IPFE**, and then **Configuration**, and then **Target Sets**.

   The fields are describe in Table 4-2.

2. Click either **Insert IPv4** or **Insert IPv6**.

   If no IPFE has been configured, an error message is displayed.

3. Under the Target Set section complete the following entries:

   a. Select the **TS Number** for the target set.

   b. Select the **Protocols** this target set supports.

c. If you want to configure the target set, but not enable its use, select **Disable**.

d. Set the **Delete Age** timer. The timer must be greater than Diameter Watchdog timer.

e. Select **Hash** or **Least Load** in the **Load Balance Algorithm** field.

4. Under the Least Load Parameters section complete the following entries:

   a. Set the **MPS Factor**.

   b. Set the **Connection Count Factor**.

   c. Set the deviation percentage of **MPS Factor** and **Connection Count Factor** in the **Allowed Deviation** field.

   d. If you want to **Enable** peer node group awareness when directing connections, check the box in the **Peer Node Aware Least Load** field.

      When enabled, the IPFE distributes connections from the same peer node group across servers in the target set to provide server redundancy for that group of peers. The IPFE keeps a group count of the connections from a peer node group to each server in the target set. Servers with a group count difference that is equal to or greater than `D` from the lowest group count are generally not considered, such as, if `D` is 1, the effect is to send the connection to the server with the lowest group count.

   e. Set the `D` value as describe in step 4d (**Peer Node Aware Least Load**) in the **Peer Node Group Distribution Threshold field**.

5. Under the Public IP Address section complete the following entries:

   a. Provide an IP **Address** to represent this target set to the outside world.

      The IP address format is either IPv4 or IPv6 depending on which button you selected in step 2. This IP address must reside on the XSI network.

   b. Select the **Active IPFE** that primarily handles traffic for this TSA.

      If an IPFE is unavailable for selection, that IPFE has not been configured.

      If the Active IPFE fails, then its mate takes over.
      If configured, the partner of the active IPFE is the standby IPFE

6. Under the Alternate Public IP Address section complete the following entries:

   a. Provide an optional **Alternate Address** that is a public IPv4 IP address to represent this target set to the outside world.

      For SCTP this address serves as an non-primary protocol-linked failover address.

      For TCP, this address can serve as an independent address.

      If this field is populated, then the column alternate IP address under target set IP list must be populated. Do not edit if in use by a local node.

   b. Select the **Active IPFE for alternate address** that handles traffic for this TSA's alternate address.

      If the Active IPFE fails, then its mate takes over.

7. Under the Target Set IP List section complete the following entries:

   a. Select an IP address in the **IP Address** field.

      This IP address must reside on the XSI network.

   b. Optionally, select an alternate IP address in the **Alternate Address** field.

   c. Enter a textual description for the application server in the **Description** field.

d. Provide a weighting value in the **Weighting** field.

   The weighting value is used to control the traffic distribution among the application servers.

e. Click **Add** to add another IP address to the list.

   You may add up to 16 IP addresses per target set.

8. Click **OK**, **Apply**, or **Cancel**.

After application servers have been added to a target set, the IPFE distributes traffic across them.

Edit a Target Set

Use this task to edit a Target Set.

When the IPFE Configuration Target Sets [Edit] page opens, the fields are initially populated with the current values for the selected target set.

Target Sets associated with an IPFE may be completely overlapping, but may not be partially overlapping. A warning appears if overlapping target sets are associated with an IPFE.

Partially overlapping target set example:
Target Set 1: Application Server 1, Application Server 2
Target Set 2: Application Server 2, Application Server 3

Completely overlapping target set example:
Target Set 1: Application Server 1, Application Server 2
Target Set 2: Application Server 1, Application Server 2

1. Click **IPFE**, and then **Configuration**, and then **Target Sets**.
2. Select the target set you want to edit and click **Edit**.
3. Update the relevant fields.

   For more information about each field please see Table 4-2.
   An IP address can be removed from the **Target Set IP List** by clicking the X at the end of the **Weighting** field. The target set IP cannot be modified.
4. Click **OK**, **Apply**, or **Cancel**.

Delete a Target Set

Use this task to delete a Target Set.

1. Select **IPFE**, and then **Configuration**, and then **Target Sets**.
2. Select the target set you want to delete and click **Delete**.
3. Click **OK** or **Cancel** on the confirmation screen.