MRA Maintenance

Maintenance Mode on the Expressway

Maintenance mode on the Expressway has been enhanced so that you can bring an MRA system down in a managed way.

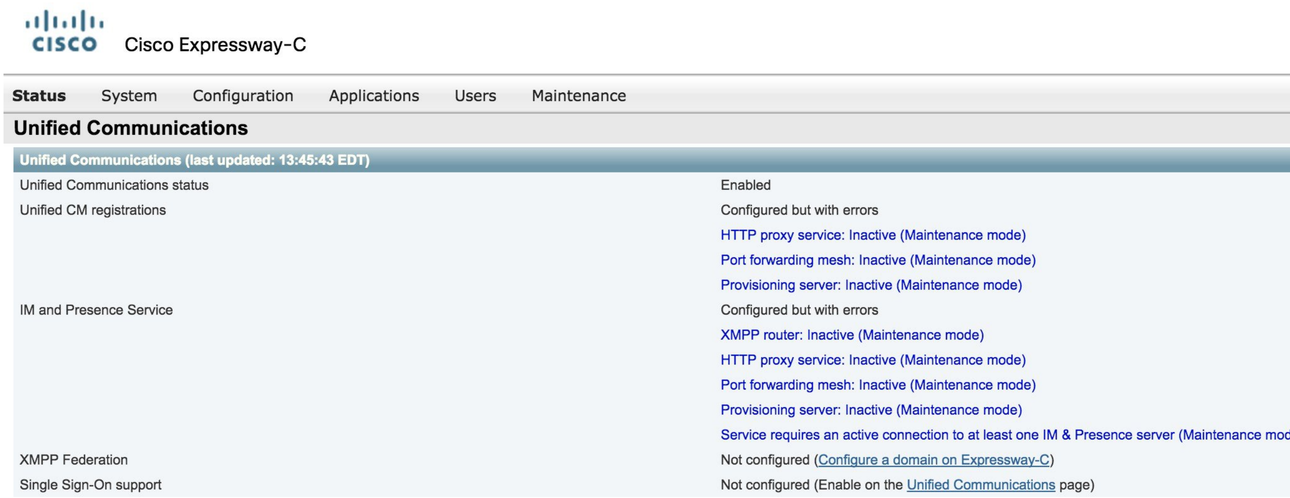
When you engage maintenance mode, the Expressway stops accepting new calls or proxy (MRA) traffic. Existing calls and chat sessions are not affected.

As users end their sessions normally, the system comes to a point when it is not processing any traffic of a certain type, and then it shuts that service down.

If users try to make new calls or start new chat sessions while the Expressway is in maintenance mode, the clients will receive a service unavailable response, and they might then choose to use another peer (if they are capable). This fail-over behavior depends on the client, but restarting the client should resolve any connection issues if there are active peers in the cluster.

The Unified Communications status pages also show (Maintenance Mode) in any places where MRA services are affected.

Figure 1: Maintenance Mode on Expressway-C



Unified Communications

Unified Communications (last updated: 13:45:43 EDT)	
Unified Communications status	Enabled
Unified CM registrations	Configured but with errors HTTP proxy service: Inactive (Maintenance mode) Port forwarding mesh: Inactive (Maintenance mode) Provisioning server: Inactive (Maintenance mode)
IM and Presence Service	Configured but with errors XMPP router: Inactive (Maintenance mode) HTTP proxy service: Inactive (Maintenance mode) Port forwarding mesh: Inactive (Maintenance mode) Provisioning server: Inactive (Maintenance mode) Service requires an active connection to at least one IM & Presence server (Maintenance moc
XMPP Federation	Not configured (Configure a domain on Expressway-C)
Single Sign-On support	Not configured (Enable on the Unified Communications page)

Limitation for CE endpoints

Maintenance mode is not supported over MRA for endpoints running CE software. The Expressway drops MRA calls from these endpoints when you enable maintenance mode.

MRA Registration Counts

From X12.6.1 onward, the **Status** > **Overview** page on Cisco Expressway-E lets you monitor up-to-date usage information for SIP devices that are registered over MRA. The **Overview** page contains the following fields:

MRA Registration:

- **Current**—The total number of devices that are currently registered over MRA.
- **Peak**—The peak count for MRA registrations since the last Expressway restart.

Authorization Rate Control

The Expressway can limit the number of times that any user's credentials can be used, in a given configurable period, to authorize the user for collaboration services. This feature is designed to thwart inadvertent or real denial of service attacks, which can originate from multiple client devices authorizing the same user, or from clients that reauthorize more often than necessary.

Each time a client supplies credentials to authorize the user, the Expressway checks whether this attempt would exceed the **Maximum authorizations per period** within the previous number of seconds specified by the **Rate control period**.

If the attempt would exceed the chosen maximum, then the Expressway rejects the attempt and issues the HTTP error 429 “Too Many Requests”.

The authorization rate control settings are configurable in the **Advanced** section of the **Configuration** > **Unified Communications** > **Configuration** page.

Credential Caching

Note These settings do not apply to clients that are using SSO (common identity) for authenticating via MRA.

The Expressway caches endpoint credentials which have been authenticated by Unified CM. This caching improves overall performance because the Expressway does not always have to submit endpoint credentials to Unified CM for authentication.

The caching settings are configurable in the **Advanced** section of the **Configuration** > **Unified Communications** > **Configuration** page.

Credentials refresh interval specifies the lifetime of the authentication token issued by the Expressway to a successfully authenticated client. A client that successfully authenticates should request a refresh before this token expires, or it will need to re-authenticate. The default is 480 minutes (8 hours).

Credentials cleanup interval specifies how long the Expressway waits between cache clearing operations. Only expired tokens are removed when the cache is cleared, so this setting is the longest possible time that an expired token can remain in the cache. The default is 720 minutes (12 hours).

Clustered Expressway Systems and Failover Considerations

You can configure a cluster of Expressway-Cs and a cluster of Expressway-Es to provide failover (redundancy) support as well as improved scalability.

Details about how to set up Expressway clusters are contained in Expressway Cluster Creation and Maintenance Deployment Guide and information about how to configure Jabber endpoints and DNS are contained in “Configure DNS for Cisco Jabber”.

Note that when discovering Unified CM and IM and Presence Service servers on Expressway-C, you must do this on the primary peer.

Expressway Automated Intrusion Protection

From X8.9 onwards, automated intrusion protection is enabled, by default, for the following categories:

- http-ce-auth
- http-ce-intrusion
- sshpfwd-auth
- sshpfwd-intrusion
- xmpp-intrusion

This change affects new systems. Upgraded systems keep their existing protection configuration.

On Expressway-C

The Expressway-C receives a lot of inbound traffic from Unified CM and from the Expressway-E when it is used for Mobile and Remote Access.

If you want to use automated protection on the Expressway-C, you should add exemptions for all hosts that use the automatically created neighbor zones and the Unified Communications secure traversal zone. The Expressway does not automatically create exemptions for discovered Unified CM or related nodes.

On Expressway-E

You should enable the Automated protection service (**System** > **System administration**) if it is not yet running.

To protect against malicious attempts to access the HTTP proxy, you can configure automated intrusion protection on the Expressway-E (**System** > **Protection** > **Automated detection** > **Configuration**).

We recommend that you enable the following categories on the Expressway-E:

- HTTP proxy authorization failure and HTTP proxy protocol violation. Do not enable the HTTP proxy resource access failure category.
- XMPP protocol violation

Note The Automated protection service uses Fail2ban software. It protects against brute force attacks that originate from a single source IP address.

Configure Exemptions

If you have Automated Intrusion Protection configured, use this procedure to configure exemptions for IP address ranges from one or more protection categories.

One example where you may need an exemption is if you have multiple MRA users connected behind a NAT using the same public IP address. This may trigger protection due to the incoming traffic from the single IP address.

Note This procedure assumes you have the Automated Intrusion Protection enabled on Expressway-E and disabled on Expressway-C, which is the recommended deployment.

Procedure

Step 1 On Expressway-E, go to **System** > **Protection** > **Automated detection** > **Exemptions**.

Step 2 Click on the **Address** that you want to configure or click **New** to configure a new address.

Step 3 Enter the **Address** and **Prefix Length** to define the IP address range that you want to exempt.

Step 4 Select from the categories to which you want to apply the exemption. For the example situation where you have multiple users behind a NAT, the following categories would apply:

- HTTP Proxy Authentication Failure

- HTTP Proxy Resource Access Failure
- SIP Authentication Failure

Step 5 Click **Add Address**.

Check the Unified Communications Services Status

You can check the status of the Unified Communications services on both Expressway-C and Expressway-E.

Procedure

Step 1 Go to **Status** > **Unified Communications**.

Step 2 Review the list and status of domains, zones and (Expressway-C only) Unified CM and IM and Presence Service servers.

The page displays any configuration errors along with links to the relevant configuration page that you access to address the issue.

Why You Need to Refresh the Discovered Nodes?

When the Expressway-C discovers a Unified Communications node, it establishes a connection to read the information required to create zones and search rules to proxy requests originating from outside of the network in towards that node. **This configuration information is static.** Expressway only reads it when you manually initiate discovery of a new node, or when you refresh the configuration of previously discovered nodes. If any related configuration has changed on a node after you discover it, the mismatch between the new configuration and what the Expressway-C knows of that node is likely to cause some kind of failure.

The information that the Expressway-C reads from the Unified Communications node is different for each node type/role. These are examples of UC configuration that you can expect to require a refresh from the Expressway. The list is not exhaustive. If you suspect that a configuration change on a node is affecting MRA services, you should refresh those nodes to eliminate one known source of potential problems.

- Changing cluster (such as adding or removing a node)
- Changing security parameters (such as enabling Mixed Mode)
- Changing connection sockets (such as SIP port configuration)
- Changing TFTP server configuration
- Upgrading node software

Devices cannot connect during the refresh

It takes some time to restore services after a server refresh and while the refresh is in progress, Jabber clients and other endpoints are unable to connect over MRA. It is not possible to provide accurate timings as they vary depending on the deployment. For straightforward deployments the refresh typically takes 5 to 10 seconds, but very complex configurations may take upwards of 45 seconds.

Refresh Servers on the Expressway-C

You must refresh the Cisco Unified Communications Manager and Cisco Unity Connection nodes defined on the Expressway-C. This fetches keys that the Expressway needs to decrypt the tokens.

Procedure

Step 1 For Unified CM, go to **Configuration** > **Unified Communications** > **Unified CM servers** and click **Refresh servers**.

Step 2 For Unity Connection, go to **Configuration** > **Unified Communications** > **Unity Connection servers** and click **Refresh servers**.