

Multipurpose Internet Mail Extensions



David Jonatan Correa

Camila García

Ángeles Tella Arena

Tecnicatura Universitaria en Programación Informática

Universidad Nacional de Quilmes

Abril de 2014

Introducción

Multipurpose Internet Mail Extensions -extensiones multipropósito de correo de internet- son una serie de convenciones dirigidas al intercambio en Internet de todo tipo de archivos. Fue creado para expandir las capacidades limitadas del correo electrónico y en particular para permitir la inserción de documentos (como imágenes, sonido y texto) en un mensaje expresando el tipo del contenido y tipo de código mediante encabezados. Una parte importante del MIME está dedicada a mejorar las posibilidades de transferencia de texto en distintos idiomas y alfabetos.

Las extensiones de MIME soportan:

- Texto en conjuntos de caracteres distintos de ASCII (American Standard Code for Information Interchange);
- Adjuntos que no son de tipo texto;
- Cuerpos de mensajes con múltiples partes (multi-part);
- Información de encabezados con conjuntos de caracteres distintos de ASCII;
- Capacidad de enviar múltiples adjuntos en un solo mensaje;
- Longitud ilimitada del mensaje;
- Uso de texto enriquecido (diseños, fuentes, colores, etc.).



Los tipos de contenido definidos por el estándar MIME tienen gran importancia también fuera del contexto de los mensajes electrónicos. Por ejemplo, HTTP requiere que los datos sean transmitidos en un contexto de mensajes tipo e-mail aunque los datos pueden no ser un e-mail propiamente dicho.

En la actualidad ningún programa de correo electrónico o navegador de Internet puede considerarse completo si no acepta MIME en sus diferentes facetas (texto y formatos de archivo).

Historia

En los primeros días de ARPANET, el correo electrónico consistía exclusivamente en mensajes de texto escritos en inglés y expresados en ASCII, regulado por SMTP (Simple Mail Transfer Protocol). El RFC¹ 822 especificaba los encabezados, pero dejaba el contenido en manos del usuario. Hoy día, este método ya no es adecuado. Los problemas incluyen envío y recepción de:

- Mensajes en idiomas con acentos (por ejemplo, español, francés y alemán).
- Mensajes en alfabetos no latinos (por ejemplo, hebreo y ruso).
- Mensajes en idiomas sin alfabetos (por ejemplo, chino y japonés).
- Mensajes que no contienen texto (por ejemplo, audio y video).

ENCABEZADO	SIGNIFICADO
Date:	fecha y hora de envío de mensajes
Reply-to:	dirección de correo electrónico a la que deben enviarse las contestaciones
Message-Id:	número único para referencia posterior a este mensaje
In-Reply-To:	identificador del mensaje al que éste responde
References:	otros identificadores de mensaje pertinentes
Keywords:	claves seleccionadas por el usuario
Subject:	resumen corto del mensaje para desplegar una línea

Algunos campos usados en el encabezado de mensaje RFC 822.



¹ Request for Comments son una serie de publicaciones que describen diversos aspectos del funcionamiento de Internet y otras redes, como protocolos, procedimientos, o comentarios e ideas sobre estos.

Se propuso una solución en el RFC 1341 y se actualizó en los RFCs 2045-2049. Esta solución, llamada **MIME**, continúa usando el formato RFC 822, pero agrega una estructura al cuerpo del mensaje y define reglas de codificación para los mensajes no ASCII. Al no desviarse del 822, los mensajes MIME pueden enviarse usando los programas y protocolos de correo electrónico existentes.

Fue desarrollada en 1991 por la IETF (Grupo de Trabajo en Ingeniería de Internet, Internet Engineering Task Force en inglés) y desde 1994 todas las extensiones MIME están especificadas de forma detallada en diversos documentos oficiales disponibles en Internet.

RFC principales:

- RFC 2045: MIME Part One: Format of Internet Message Bodies
- RFC 2046: MIME Part Two: Media Types
- RFC 2047: MIME Part Three: Message Header Extensions for Non-ASCII Text
- RFC 2048: MIME Part Four: Registration Procedures
- RFC 2049: MIME Part Five: Conformance Criteria and Examples

RFC secundarios:

- RFC 1524: The formal description of mailcap files. Mailcap files describe how to handle media types.
- RFC 2015: MIME Security with Pretty Good Privacy (PGP).
- RFC 2110: MIME E-mail Encapsulation of Aggregate Documents, such as HTML (MHTML).
- RFC 2111: Content-ID and Message-ID Uniform Resource Locators.
- RFC 2112: The MIME Multipart/Related Content-type.
- RFC 2183: Defines the syntax and semantics of the "Content-Disposition" header to convey presentational information.
- RFC 2184: MIME Parameter Value and Encoded Word Extensions: Character Sets, Languages, and Continuations

Prácticamente todos los mensajes de correo electrónico escritos son transmitidos en formato MIME a través de SMTP (Simple Mail Transfer Protocol), y están tan cercanamente asociados con éstos que usualmente se les llama mensaje SMTP/MIME.

Sin MIME, SMTP sólo admite el envío de caracteres de texto y además usando sólo 7 bits. Esto hace que pueda enviarse un conjunto muy limitado de caracteres de texto (128), y no pueden enviarse caracteres con acentos y eñes, entre otros.

Si queremos enviar un correo cuyo asunto es "Camión" y el texto es "Acentos áéíóó y otros Ñ ñ ü", el texto con caracteres de 7 bits que se enviaría al servidor SMTP sería:

```
From: yo@yo.com
To: chuidiang@gmail.com
Message-ID: <12097592.01175854969750.JavaMail.Administrador@trasto2>
Subject: =?Cp1252?Q?Cami=F3n?=
MIME-Version: 1.0
Content-Type: text/plain; charset=Cp1252
Content-Transfer-Encoding: quoted-printable
```

```
Acentos =E1=E9=ED=F3=F3 y otros =D1 =F1 =FC
```

En las tres primeras líneas hay datos sobre el correo, indicados en la forma en la forma que SMTP requiere.

La siguiente línea es el subject, que se ha convertido a caracteres normales, quitando el tilde de camión y reemplazandolo por =F3. Como el subject va delante de la cabecera MIME, lleva su conversión por defecto, decidida por el cliente de correo.

En el mensaje se han reemplazado todos los tildes y caracteres extraños por caracteres normales. El punto de la última línea es como el protocolo SMTP indica que es el final de mensaje.

Encabezados

MIME define cinco nuevos encabezados de mensaje:

ENCABEZADO	SIGNIFICADO
MIME-version:	Identifica la versión de MIME.
Content-transfer-encoding:	Como se envuelve el mensaje para su transmisión.
Content-type:	Naturaleza del mensaje.
Content-description: (opcional)	Cadena de Texto que describe el mensaje.
Content-Id: (opcional)	Identificador único.

MIME-version: Indica al agente de usuario receptor del mensaje con qué versión de MIME está tratando. Se considera que cualquier mensaje que no contenga un encabezado MIME-version: es un mensaje de texto normal en inglés, y se procesa como tal.

MIME-Version: 1.0

Content-Transfer-Encoding: indica la manera en que está envuelto el cuerpo para su transmisión, sino podría tener problemas con la mayoría de los caracteres distintos de las letras, números y signos de puntuación.

Content-Transfer-Encoding: 7bit

TIPO	DESCRIPCIÓN
7bit	El cuerpo contiene los caracteres ASCII de 7-bits con una longitud máxima de 1000 caracteres.
8bit	Puede haber caracteres de 8-bits que no sean ASCII pero la longitud máxima del cuerpo es limitada a 1000 caracteres.
Binary	Caracteres binarios de 8-bits sin limitación de 1000 caracteres en el cuerpo.
Quoted-printable	Esto es útil cuando la información esta compuesta, en gran parte, por caracteres cuyos valores de bytes están fuera del rango ASCII. Los caracteres con valor decimal equivalentes del 33 al 61 en ASCII, son representados en ASCII. Otros son representados como una hexa-representación de dos dígitos precedida por el signo '='. Los caracteres que no son de texto son reemplazados por secuencias hexa de seis dígitos.
Base 64	Bloques de 6-bit de información entrante es codificada en bloques de salida de 8-bit

Tipos para Content-Transfer-Encoding.

Content-Type: Especifica la naturaleza del cuerpo del mensaje, definiéndolo por tipo/subtipo. Hay casos en los que el content-type es de tipo multipart, en los cuales una o más clases de datos son combinados. El cuerpo del mensaje debe contener varias 'partes', encapsuladas por boundaries (límites). Todas deben comenzar con un boundary, y sólo la última parte tiene uno de cierre.

Content-Type: video/mpeg

ó

Content-Type: multipart/mixed;
boundary=ejemploBoundary

Esta parte es ignorada. Por lo general, se utiliza para enviar mensajes a usuarios de clientes viejos que no soporten MIME.

--ejemploBoundary

Este es texto plano ASCII implícito (no se especifica el tipo).

Esta parte del cuerpo no utiliza boundary de cierre.

--ejemploBoundary

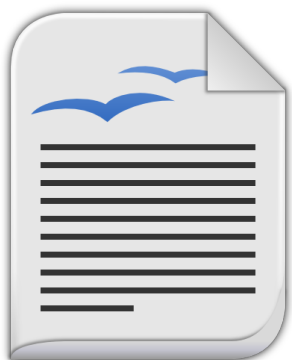
Content-type: text/plain; charset=us-ascii

Este es texto plano ASCII explícito.

Esta parte del cuerpo sí utiliza boundary de cierre, ya que es la última.

--ejemploBoundary--

Esta parte también es ignorada.



TIPO	SUBTIPO	DESCRIPCIÓN
Text	<ul style="list-style-type: none"> • Plain • Enriched 	<ul style="list-style-type: none"> - Texto sin formato - Texto con comando de formatos sencillos.
Image	<ul style="list-style-type: none"> • Gif • Jpeg 	<ul style="list-style-type: none"> - Imagen fija en formato GIF - Imagen fija en formato JPEG
Audio	<ul style="list-style-type: none"> • Basic 	<ul style="list-style-type: none"> - Sonido
Video	<ul style="list-style-type: none"> • Mpeg 	<ul style="list-style-type: none"> - Pelicula en formato MPEG
Application	<ul style="list-style-type: none"> • Octet-stream • Postscript 	<ul style="list-style-type: none"> - Secuencia de bytes no interpretada - Documento imprimible en PostScript
Message	<ul style="list-style-type: none"> • Rfc822 • Partial • Extern 	<ul style="list-style-type: none"> - Mensaje MIME RFC 822 - Mensaje dividido para su transmisión - El mensaje mismo debe obtenerse de la red
Multipart	<ul style="list-style-type: none"> • Mixed • Alternative • Parallel • Digest 	<ul style="list-style-type: none"> - Partes independientes en el orden especificado - Mismo mensaje en diferentes formatos - Las partes deben verse en forma simultánea - Cada parte es un mensaje RFC 822 completo

Tipos y subtipos para Content-type.

Content-Description: Header opcional que explica brevemente lo que está en el mensaje, con el fin de que el destinatario sepa si vale la pena decodificar y leer el mensaje. Si la cadena dice: "Foto de mi hamster" y la persona que recibe el mensaje no es un gran fanático de los hamsters, el mensaje probablemente será descartado en lugar de decodificado para dar una foto a color de alta definición.
Content-Description: Foto de mi hamster en HD.

Content-Id: Header opcional que identifica unívocamente el cuerpo del mensaje; usa el mismo formato que el encabezado estándar Message-Id. Tiene el formato <xxxx@yyyy>, indicando a la izquierda un número de serie, la fecha cuando se generó el mensaje, o ambas cosas, y a la derecha indica donde se originó la parte MIME.

Se usa para identificarlo en varios contextos, particularmente para conseguir data referenciada por el mecanismo de message/external-body². En ese caso, es obligatorio que se especifique este header.

Content-ID: <my.random.id@some.domain>

Content-ID: <part1.02080004.04000407@sample.com>

² message/external-body especifica que un objeto puede contener una referencia o puntero a algún tipo de data en vez de contener la data en sí.

La vida como un mail MIME

Cuando componemos un mensaje en un programa de Emails que permite MIME, este realiza las siguientes acciones:

- Si el mensaje esta en ASCII solamente, no realiza nada y le dice al cliente del email receptor que solo espere texto plano.
- Si el mensaje contiene uno o más archivos adjuntos y el cuerpo posee un formato HTML, cada parte es mirada y tratada por separado.

Primero, se determina el formato de la información. Es necesario decirle al cliente del email receptor que hacer con la información, y asegurarse de que sea codificado propiamente para no perder nada en la transferencia. Luego, la información es codificada si está en un formato distinto a un texto plano en ASCII. Durante el proceso de codificación, la información es convertida al texto plano conveniente para los mensajes RFC 822.

Finalmente, la información codificada es insertada en el mensaje, y se informa al cliente del email receptor que tipo de información esperar tal como la existencia de archivos adjuntos, cómo están codificados y cuál era el formato del archivo original.

Del lado del receptor, el proceso es inverso. Primero, el cliente de mail lee la información que fue añadida por el cliente del mail enviado buscando saber si hay archivos adjuntos, cómo debe decodificarlos y cómo manejar los archivos resultantes. Luego, cada parte del mensaje es extraída y decodificada si es necesario.

Finalmente, el cliente de mail muestra las partes resultantes al usuario. El cuerpo de texto plano es mostrado en línea en el cliente de mail junto con la imagen adjunta. El programa adjunto con el mensaje es también mostrado con un icono de "adjunto" y el usuario puede decidir qué hacer con él, tanto como correrlo desde el mismo mail o guardarlo en el disco.



¿Por qué son importantes los tipos correctos de MIME?

Si el servidor de red o la aplicación informan un tipo incorrecto de MIME para el contenido, un navegador de red no tiene forma de saberlo de acuerdo a la especificación HTTP, más si tenemos en cuenta que el autor especificó el contenido para ser procesado y mostrado en una forma diferente que la impuesta por el tipo MIME informado.

Algunos navegadores intentan determinar el tipo adecuado de MIME en servidores mal configurados, suponiendo el tipo adecuado. El problema con este método es que el navegador continúa procesando el contenido aunque, por ejemplo, una imagen haya sido informada como texto plano.

¿Por qué los navegadores no deberían suponer tipos MIME?

Pérdida del control

Si el navegador ignora el tipo MIME reportado, los administradores de red y los autores dejarán de tener el control sobre cómo sus contenidos serán procesados.

Por ejemplo, un sitio orientado para desarrolladores de red puede enviar determinados ejemplos de documentos HTML como enteros text/html ó como text/plain para lograr documentos con enteros procesados y mostrados como HTML ó como código fuente. Si el navegador supone el tipo MIME, esta posibilidad dejará de estar disponible para el autor.

Seguridad

Algunos tipos de contenidos, tales como programas ejecutables, son inseguros. Por este motivo, esos tipos MIME son generalmente restringidos en términos de qué acciones tomará el navegador de red al recibirlos. Por ejemplo, un programa ejecutable no debería ser ejecutado en la computadora de un usuario, y en su lugar debería aparecer un cuadro de diálogo para preguntar si desea descargar el archivo.

La suposición de tipos MIME ha llevado a fallas de seguridad importantes, debido a autores maliciosos que reportaban el tipo MIME de un archivo peligroso como si fuera uno seguro, evitando así el cuadro de diálogo de descarga normal.

S/MIME

S/MIME (Secure / Multipurpose Internet Mail Extensions, del inglés, Extensiones de Correo de Internet de Propósitos Múltiples / Seguro) es un estándar para criptografía de clave pública y firmado de correo electrónico encapsulado en MIME.

S/MIME provee los siguientes servicios de seguridad criptográfica para aplicaciones de mensajería electrónica:

- autenticación, integridad y no repudio (mediante el uso de firma digital) y
- privacidad y seguridad de los datos (mediante el uso de cifrado)

La entidad MIME completa se cifra y se empaca en un objeto que luego se inserta en una entidad MIME application/pkcs7-mime.

```
Content-Type: application/pkcs7-mime;  
             smime-type= enveloped-data;  
             name=smime.p7m
```

La funcionalidad S/MIME está construida en la mayoría de los clientes de correo electrónico modernos y son capaces de interoperar entre ellos.

Antes de que S/MIME pueda usarse en alguna de las aplicaciones antes mencionadas, se debe obtener e instalar una clave/certificado individual tanto de la autoridad certificadora (AC) interna como de una AC pública. Una buena práctica es usar claves privadas separadas (y certificados asociados) para firma y para cifrado.

Este tipo de mail “seguro” tuvo su primera aparición en 1995, convirtiéndolo en una de las tantas especificaciones que había en ese entonces con respecto a la seguridad en los mensajes. Más tarde, en 1998, salió una segunda versión, la cual logró que S/MIME pasara de ser un estándar entre muchos, al estándar principal. Un año más tarde, salió una tercera versión con el objetivo de poder enriquecer las capacidades de la versión anterior, incorporando servicios adicionales como “Secure Receipts”, “Triple-Wrapping” y “Security Labels”. Esta es la versión que se utiliza hoy en día, pero vale tener en cuenta que algunos de los clientes de mail principales -Gmail y Hotmail incluidos- no utilizan esta adición segura de MIME.



Bibliografía

Redes de Computadoras - Andrew Tanenbaum, Pearson, Prentice Hall, 4ta edición

[Balabit.com - The MIME Protocol](#)

[ComputerNotes.com - MIME](#)

[RogueWave.com - Using the MIME Headers Effectively](#)

[W3.org - The Multipart Content-Type](#)

[About.com - How MIME works](#)

[Technet.Microsoft.com - Understanding S/MIME](#)