

ICS 33.020

L 07

pt  
tsn  
通  
标  
网  
.net.cn

**YD**

# 中华人民共和国通信行业标准

YD/T 2030-2009

---

## 互联网中文电子邮件地址框架 总体技术要求

General technical specification for Chinese internet email address  
(IETF RFC 4952, Overview and Framework for Internationalized Email, NEQ)

2009-12-11 发布

2010-01-01 实施

---

中华人民共和国工业和信息化部 发布

电话：82054513 <http://www.ptsn.net.cn>



## 目 次

前 言	II
引 言	III
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
4 协议概述	4
5 协议要求	4

电话：82054513 <http://www.ptsnet.cn>

## 前 言

本标准是互联网中文电子邮件地址系列标准之一。该系列标准预计的名称和结构如下：

1. 互联网中文电子邮件地址框架总体技术要求；
2. 简单邮件传输协议（SMTP）扩展支持互联网中文电子邮件地址技术要求；
3. 互联网中文电子邮件地址格式的邮件头技术要求；
4. 互联网中文电子邮件地址与普通邮件系统兼容技术要求；
5. 在POP中支持互联网中文电子邮件地址技术要求；
6. 在IMAP中支持互联网中文电子邮件地址技术要求；
7. 互联网中文电子邮件地址客户端技术要求。

本标准对应IETF RFC 4952《国际化电子邮件地址框架》（英文版），与IETF RFC 4952的一致性程度为非等效。本标准在技术内容上与IETF RFC 4952保持一致，其中第3章“术语和定义”是将IETF RFC 4952中的术语和定义修改归集，并增加了一些适合本标准的特定的术语和定义；第4章和第5章则依据IETF RFC 4952和IETF国际化电子邮件地址工作组的相关核心草案的基本思路，将其中有关国际化电子邮件地址的规定都转换成只针对中文电子邮件地址的规定，但是其技术思路未作修改。

本标准由中国通信标准化协会提出并归口。

本标准起草单位：中国互联网络信息中心（CNNIC）、工业和信息化部电信研究院、中兴通讯股份有限公司、清华大学、中国移动通信集团公司。

本标准主要起草人：李晓东、姚健康、曹蓟光、李明栋、崔 勇、段晓东。

## 引 言

互联网是一个基于开放互联协议的网络，电子邮件发送是互联网上的基础服务，可以用来传送互联网信息，电子邮件服务使互联网信息的传递更加快捷。传统的电子邮件地址都是ASCII格式的，随着中文域名的推广使用，人们越来越迫切需要中文电子邮件地址。实现对中文电子邮件地址的支持首先应实现对国际化电子邮件地址的支持。

随着互联网的发展，中文用户的数量不断增加，对于中文域名的使用需求也在增加。但是中文域名的最大应用中文电子邮件地址由于缺乏相关标准，没有得到很好的应用和推广。中文电子邮件地址和传统的英文电子邮件地址有较大差别，比如：中文电子邮件地址的字符集比传统的英文域名的字符集大很多。为了规范中文电子邮件地址的使用，让中文用户能够方便的通过中文电子邮件地址来使用互联网的各种应用服务，尽快制定中文电子邮件地址标准，进而推动中文电子邮件地址的使用是十分重要的。

# 互联网中文电子邮件地址框架总体技术要求

## 1 范围

本标准规定了在互联网体系上使用中文电子邮件地址的框架体系总体技术要求，从服务器端和客户端提出相应的技术规范。

本标准适用于各级电子邮件地址注册管理机构、电子邮件地址服务提供商以及软件厂商开发支持中文电子邮件地址的应用或者服务等。

## 2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

ISO/IEC 10646	信息技术 通用多八位编码字符集（UCS）
IETF RFC 822	互联网信息格式旧版
IETF RFC 1652	SMTP扩展支持8比特MIME传输
IETF RFC 1939	邮局协议第三版本
IETF RFC 2822	互联网信息格式新版
IETF RFC 3454	国际化字符串预处理
IETF RFC 3492	一种适用于国际化域名应用的对UNICODE的编码方法：Punycode
IETF RFC 3501	互联网信息交互协议第四版本
IETF RFC 5335	国际化电子邮件地址格式的邮件头技术要求
IETF RFC 5336	SMTP扩展支持国际化电子邮件地址技术要求
IETF RFC 5504	国际化电子邮件地址与普通邮件系统兼容技术要求

## 3 术语、定义和缩略语

### 3.1 术语和定义

下列术语和定义适用于本标准。

#### 3.1.1

##### 电子邮件

在计算机网络上，用户终端之间往来的信函。

#### 3.1.2

##### 最终投递MTA

一种SMTP服务器，它可以控制邮件地址本地部分的格式并且允许检查和解释地址的本地部分，它从网络接受信息投递给邮箱或者其他本地过程而不是中转（relay）。从网络的观点来看，任何本地的投

递安排，比如保存在一个信息仓库，转交给特殊的信息投递程序或代理以及获取信息机制全都是在最终投递MTA之后，因此并不是SMTP传输或者最终投递MTA过程的一部分。

### 3.1.3

#### Unicode字符

Unicode根据其位置或码位来识别字符，给每个字符提供的一个惟一的数字。比如说，U+12AB指的是在Unicode 3.2表中位于12AB处的字符。本标准的Unicode字符应符合ISO/IEC 10646的规定。Unicode字符集包含ASCII字符集。

### 3.1.4

#### UTF-8

Unicode 是分配整数给字符的编码表，UTF-8是将Unicode中的一串字符表示为一串字节的方法。在UTF-8中，字符采用1~6个8比特字节的序列进行编码。在一个8比特字节的一个序列中，字节的高位为0，其他的7位用于字符值编码。 $n$  ( $n>1$ ) 个8比特字节的一个序列中，初始的8比特字节中高 $n$ 位为1，接着一位为0，此字节余下的位包含被编码字符值的位。接着的所有8比特字节的最高位为1，再接着下一位为0，余下每个字节6位包含被编码字符的位。

### 3.1.5

#### ASCII地址

如果一个地址中的每一个字符都属于ASCII字符集，且符合IETF RFC 821中规定的格式，那么这个地址是“全部ASCII”的地址或者是一个“ASCII地址”。

### 3.1.6

#### UTF8SMTP地址

如果电子邮件地址中至少有一个字符满足属于Unicode字符集但是不属于ASCII字符集，那么这个地址被称为“UTF8SMTP地址”。

### 3.1.7

#### 中文域名 Chinese domain name

含有中文域名字段的域名。

### 3.1.8

#### ASCII用户

一个ASCII用户只使用ASCII地址，不使用也不能使用UTF8SMTP地址。

### 3.1.9

#### i18mail 用户

一个“i18mail用户”指拥有一个或多个UTF8SMTP地址，这些用户可能也拥有ASCII地址。如果用户拥有多于一个账户和相应的电子邮件地址，或者同一个电子邮件地址有多于一个别名，他可以通过某种方法选择在发出的邮件中使用哪个地址。在这种情况下，是不可能通过地址来辨别发件人或者收件人是否是一个i18mail用户。

### 3.1.10

#### 信息 message

一个信息是从一个用户（发送者）利用特定的电子邮件地址发送到另一个或者多个接收电子邮件地址(经常仅仅称作用户或接收用户)。

### 3.1.11

#### 追踪头trace field

邮件头部信息中为了追踪邮件所经过的路径而专门设的主机字段。

### 3.1.12

#### Punycode

一种编码转换规则。运用这种规则应可实现Unicode字符串和ASCII字符串的相互转换。详见IETF RFC 3492。

### 3.1.13

#### 中文电子邮件系统

支持本标准的邮件系统。

### 3.1.14

#### ASCII电子邮件系统

只支持IETF RFC 822和IETF RFC 2822所规定的ASCII格式形式地址的邮件系统。

### 3.1.15

#### 电子邮件地址本地部分

电子邮件地址“@”的左半部分。

### 3.1.16

#### 电子邮件地址域名部分

电子邮件地址“@”的右半部分。

### 3.1.17

#### POP协议

IETF RFC 1939规定的邮局协议。

### 3.1.18

#### IMAP协议

IETF RFC 3501规定的互联网信息交互协议。

### 3.1.19

#### 字符串预处理 Stringprep

按照IETF RFC 3454的规定，对字符串进行处理的过程。

## 3.2 缩略语

下列缩略语适用于本标准。

ACE	ASCII Compatible Encoding	ASCII兼容编码
CDN	Chinese Domain Name	中文域名
DNS	Domain Name System	域名系统
Email	Electronic Mail	电子邮件
IMAP	Interactive Mail Access Protocol	交互式邮件接入协议

MDA	Message Delivery Agent	邮件递送代理
MS	Message Store	信息存储器
MSA	Message Submission Agent	邮件提交代理
MTA	Mail Transfer Agent	邮件传递代理
MUA	Mail User Agent	邮件用户客户端
POP	Post Office Protocol	邮局协议
SMTP	Simple Mail Transfer Protocol	简单邮件传输协议

#### 4 协议概述

IETF RFC 3490已经允许国际化域名和中文域名CDN的使用。目前，国内还没有完全中文化的互联网名字体系。域名只是各种需要中文化的名字和标识符中的一种。在很多环境中，仅仅是中文域名并不能很好地方使用户使用中文化的互联网，需要更多的中文标识符来支持。广大中文域名用户在使用中文域名的时候都迫切需要与中文域名相关的应用，与中文域名最相关的一个应用就是中文电子邮件地址的使用。

中文电子邮件地址更能符合中文互联网用户的上网。为了支持中文电子邮件地址，需要对原有的邮件系统进行扩展支持中文电子邮件地址。中文电子邮件地址的本地部分和域名部分可以先通过字符串预处理的方式来判定该电子邮件地址是否合适作为合法的电子邮件地址。

电子邮件地址中文化不是简单的把SMTP信封做些改变，或修改“From, To和Cc”字段，或进行特殊的编码来显示本地的字符。为了对收到的电子邮件地址更有用，处理的中文电子邮件地址必须和它们产生时的环境保持一致。因此必须建立一个中文化电子邮件通信环境以便使用中文的用户能够很好的进行交流，需要允许在邮件的信封和信头里都能够使用UTF-8格式的字符，而这要求SMTP扩展支持UTF-8编码以允许中文电子邮件地址的发送和接收。

#### 5 协议要求

##### 5.1 总体要求

本标准要求更新现有的SMTP协议和电子邮件地址的格式，以便允许中文电子邮件地址的显示和传输。下面从服务器端和客户端来具体介绍协议的实现。

##### 5.2 SMTP 扩展支持中文电子邮件地址

SMTP协议要求进行扩展来支持中文电子邮件地址。这个扩展的关键字是“UTF8SMTP”，作为一个SMTP扩展，“UTF8SMTP”定义为：

——允许在电子邮件地址中使用UTF-8字符串，包括电子邮件地址本地部分和域名部分。

——允许在电子邮件地址中有选择的使用UTF-8格式的中文字符串。

——要求服务器声明 8BITMIME扩展[IETF RFC 1652]以及客户端支持8比特传输，这样头部信息可以不用通过特殊的内容传输编码（content-transfer-encoding）就能够传输。

——提供必要的信息来支持向下兼容机制。

支持中文电子邮件地址的中文电子邮件系统应遵循以下原则：



a) 中文电子邮件地址可能会进入不同的系统或子系统，这些系统可能会对中文电子邮件地址进行字符转换或进行编码转换。如果电子邮件地址的本地部分含有中文，域名部分不宜使用punycode编码来显示给用户，以保持编码和格式的一致性。

b) 一个SMTP中继可以有以下选择：

1) 或者明确的识别格式，可以通过ESMTP的“UTF8SMTP”声明，来明确标示支持中文电子邮件地址；

2) 可以选择和使用ASCII地址，对信息进行处理以便与现有的ASCII电子邮件系统兼容；

3) 拒绝发送邮件，然后给发送者返回一个未投递通知信息，这样发送者可以采取其他方法来发送邮件，如果因为下一跳的系统不支持“UTF8SMTP”扩展，而且没有足够的信息可以利用实现降级，那么必须拒绝或者产生并且发送一个未投递信息给发送者。

c) 目前没有一种可行的方法来正确识别UTF-8字符，允许多种编码的字符容易引起混乱，也不利于世界范围内的邮件的互通互联，本标准规定在电子邮件地址及其头部禁止使用非UTF-8编码的字符。在SMTP服务器做DNS域名记录查询时，应遵循IETF RFC 3490中规定的格式，以punycode编码的ACE格式向DNS服务器提交数据。

### 5.3 邮件头和信封支持中文电子邮件地址

传统的邮件信息格式只允许ASCII字符出现在邮件头里，本标准要求中文电子邮件地址系统必须允许非ASCII字符出现在邮件头里，这些字符是以UTF-8编码格式的UNICODE字符，通过UTF-8编码传输电子邮件头部域。

允许在邮件头里使用非ASCII字符，会影响SMTP客户端、SMTP服务器、邮件用户客户端和网关等各种解析和处理邮件信息的进程。在IETF RFC 5336里规定了用“UTF8SMTP”扩展来阻止非ASCII字符的传输，来避免在传输过程中带有邮件头的信息被错误解析。使用“UTF8SMTP”扩展并不能阻止非ASCII邮件头信息传递给邮件存储器，如果这些存储器没有更新支持中文电子邮件系统，可能不会正确解析这些邮件信息，因此这些存储系统如POP、IMAP等也必须更新支持中文电子邮件地址系统。本节的目的是允许非ASCII字符在邮件头里传输，并不规定如何将这些信息传递给非中文电子邮件系统。IETF RFC 5504规定了在邮件传输过程中遇到不支持中文电子邮件地址系统的SMTP服务器时候的具体处理办法。

在邮件头里将主要做如下变化：

a) 允许邮件头里出现UTF-8编码的UNICODE字符；

b) 在MIME头里增加message/global类型；

c) 邮件头的语法格式进行扩展支持UTF-8编码；

d) 追踪头（trace field）的格式语法更新。

### 5.4 兼容现有的ASCII电子邮件系统

由于中文电子邮件系统的出现，互联网上必然也存在ASCII电子邮件系统。对于任何SMTP扩展机制，都有可能一个SMTP客户端要求一些属性而服务器并不支持要求的属性。如果一个信封地址或邮件头信息包含非ASCII字符，这封邮件就不能投递给不支持UTF-8扩展的SMTP服务器。对于中文电子邮件地址的投递，需要在传输过程中每一个邮件服务器都支持“UTF8SMTP”扩展。当其中的个别服务器不支持这个扩展时候，在IETF RFC 5336中规定了4种方法来处理。如果客户端使用的电子邮件地址中带有服务器端可以使用的alt-address参数，本标准规定可以依照本节描述的方法进行向后兼容处理。对于发送服务

器的选择基本上处于发送者客户端的控制之下，并且中间可能的中转服务器的选择也可能在最终投递MTA服务器的管理之下。为了中文电子邮件能够顺利的投递，发送服务器和最终投递的服务器的管理者应该尽可能的使传输过程中所涉及到的邮件服务器都支持中文电子邮件地址系统。

IETF RFC 5504规定了具体的向后兼容支持中文电子邮件地址系统的详细要求。向后兼容机制可以在行使SMTP客户端功能的MUA、MSA和MTA端实现，也可以在MDA、POP和IMAP等存储邮件端实现。向后兼容机制应尽量保护原来的信息不受破坏。

向后兼容机制主要包含下列4个部分：

- a) 新的头字段定义；
- b) SMTP向后兼容；
- c) 邮件头字段向后兼容；
- d) MIME头字段向后兼容机制。

## 5.5 POP 扩展支持中文电子邮件系统

### 5.5.1 LANG 能力

POP3允许大多数的响应需要返回给用户可读的文本，但是POP3协议规定这些文本必须用ASCII字符。LANG功能和命令允许POP3客户端与服务器协商应该使用什么语言来传递这些文本。

为了简化解析，所有的POP3服务器都应允许中文字符。

### 5.5.2 UTF8 能力

这个功能向POP3增加"UTF8"命令，这个命令切换会话进程从ASCII到UTF8模式。邮件投递系统可以存储UTF8格式的信息或只存储ASCII格式的字符或二者都可以。UTF8模式对邮件投递系统中的ASCII格式的信息没有影响。在UTF8模式下，UTF8和ASCII信息都被不经转化的直接发送给客户端，如果不是UTF8模式，那么邮件投递系统中的UTF8信息必须使用IETF RFC 5504中的方法做向下兼容的处理。

## 5.6 IMAP 扩展支持中文电子邮件系统

现有的IMAP基础协议禁止在基础字串中或带引号的字串中使用8位的字符。支持中文电子邮件地址的IMAP应扩展支持“UTF-8”能力从而支持8位字符，应允许IETF RFC 5335规定的邮件头格式，同时也要确立一种机制来支持UTF-8格式的邮箱名字和登录用户名及密码。IMAP客户端使用ENABLE命令来通知服务器可以使用与UTF-8相关的机制。

## 5.7 邮件客户端扩展支持中文电子邮件系统

邮件客户端具有和邮件提交代理MSA相互作用的界面来发送邮件，和邮件存储交互来收取邮件。收取邮件的接口直接进入文件系统或进入POP或IMAP服务器。客户端也提供了用户界面，允许终端用户读取、显示、撰写邮件。

支持中文电子邮件地址的客户端应该有能够支持UTF-8格式的字符的能力。对于一个支持UTF8SMTP的MUA，会遇到多种可能性，基于电子邮件信封和正文是否包含非ASCII字符以及MSA 是否支持UTF8SMTP扩展等问题。

如果MSA不支持UTF8SMTP扩展的话，MUA不应该发送带有UTF8SMTP头部的信息，可以采用IETF RFC 5336中提供的4种方法来处理，这4种方法是：

- 用ASCII重写头部；
- 拒绝信息；

- 寻找一个能够到达目的的替代路由（如果客户端与MSA通过接口相连MTA）；
- 向下兼容信息。

在邮件的接收过程中，有可能原始的邮件是一个UTF8SMTP邮件而在传输过程中向下兼容为ASCII。一个邮件头部如果有一个或者多个头部带有前缀"Downgraded-"的字段，可以确认这封邮件在传输过程中进行了向下兼容处理。

---