

ICANN

VIRTUAL COMMUNITY FORUM

70

Email Address Internationalization

Champika Wijayatunga
Technical Engagement Manager – APAC

22 March 2021

Overview

www.cafe-123.com



②

Forming ASCII Labels

Use LDH

- **Letters** [a-z]
- **Digits** [0-9]
- **Hyphen** [H]

Label length = 63

Other constraints (e.g. on hyphen)

①

Forming ASCII Labels

Use only Letters

- **Letters** [a-z]

Label length = 63

Domain Name Mnemonics in ASCII

Using LDH

- Letters [a-z]
- Digits [0-9]
- Hyphen (H)

2

	0	1	2	3	4	5	6	7
0	NUL	DLE	space	0	@	P	`	p
1	SOH	DC1 XON	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3 XOFF	#	3	C	S	c	s
4	EOT	DC4	\$	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(8	H	X	h	x
9	HT	EM)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
B	VT	ESC	+	;	K	[k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M]	m	}
E	SO	RS	.	>	N	^	n	~
F	SI	US	/	?	O	_	o	del

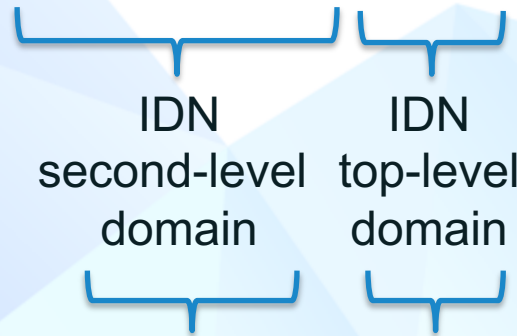
Using Letters only

- Letters [a-z]
- ~~Digits [0-9]~~
- ~~Hyphen (-)~~

①

	0	1	2	3	4	5	6	7
0	NUL	DLE	space	0	@	P	`	p
1	SOH	DC1 XON	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3 XOFF	#	3	C	S	c	s
4	EOT	DC4	\$	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(8	H	X	h	x
9	HT	EM)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
B	VT	ESC	+	;	K	[k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M]	m	}
E	SO	RS	.	>	N	^	n	~
F	SI	US	/	?	O	_	o	del

ตัวอย่าง.ไทย



② Syntax of IDN Labels

Valid U-Label:

Unicode code points as constrained by the “LDH” scheme within IDNA 2008

① Syntax of IDN Labels

Valid U-label,

further constrained by the “letter” principle for TLDs

- Encoding glyphs into codepoints covering the different scripts used for languages
 - Codepoints are shown in hex using the U+XXXX notation
- Codepoints are typically carried using the UTF-8 (Unicode Transformation Format, 8 bit) format
 - Variable number of bytes for a single codepoint.
 - ASCII is used as is
 - Gold standard for carrying Unicode codepoints, in web, protocols, etc.

	060	061	062	063	064	065	066	067	068	069
0	 0600	و	ي	ذ	-	ِ	٠	ُ	پ	ڈ
1	 0601	و	ء	ر	ف	ّ	١	أ	خ	ڑ
2	 0602	و	آ	ز	ق	ُ	٢	أ	خ	ز
3	 0603	و	أ	س	ك	ّ	٣	إ	ج	ر
4	 0604	و	ؤ	ش	ل	ُ	٤	ج	ر	ر
5	 0605	و	إ	ص	م	ِ	٥	أ	خ	ر
6	 0606	و	ئ	ض	ن	ِ	٦	و	ج	د

- ⊙ Multiple ways to use a glyph:
 - è = U+00E8
 - e + ` = è = U+0065 U+02CB
- ⊙ Normalization is a process to ensure that whatever the user types, the end representation will be the same.
 - For the two entries above, Normalization Form C (NFC) will generate U+00E8 for both
- ⊙ Note: case folding is not stable (i.e. upper to lower to upper does not always result with the same value)

- ◉ How to use Unicode for domain names?
 - IDNA2008 standard by IETF (RFCs 5890-5894)
 - Determines which code points can be used – sets a baseline
 - Specifies NFC form for labels represented in UTF8 format and specifies additional constraints and conversion “on-the-wire”
 - Further recommendations by community-based work and published by ICANN org to address security and stability considerations
 - Top-level domains: [Root Zone Label Generation Rules](#)
 - Second and other levels: [Reference Second Level Label Generation Rules](#)

- What is EAI
 - Having UTF8 support for:
 - Mailbox name (before the @ sign)
 - Domain name (after the @ sign)

- What is not EAI
 - Having UTF8 support in:
 - Subject line
 - Address comments
 - Message body
 - MIME provides all these in conventional mail
 - Use of any character set other than UTF-8

Categories of Domain Names and Email Addresses

- Domain Names

- **Newer** top-level domain names:
- **Longer** top-level domain names:
- **Internationalized** domain names

example.**sky**

example.**abudhabi**

普遍接受-测试.世界

- Internationalized email addresses (EAI)

- ASCII@IDN
- UTF8@ASCII
- UTF8@IDN
- UTF@IDN; right to left scripts

marc@**société**.org

ईमेल@example.com

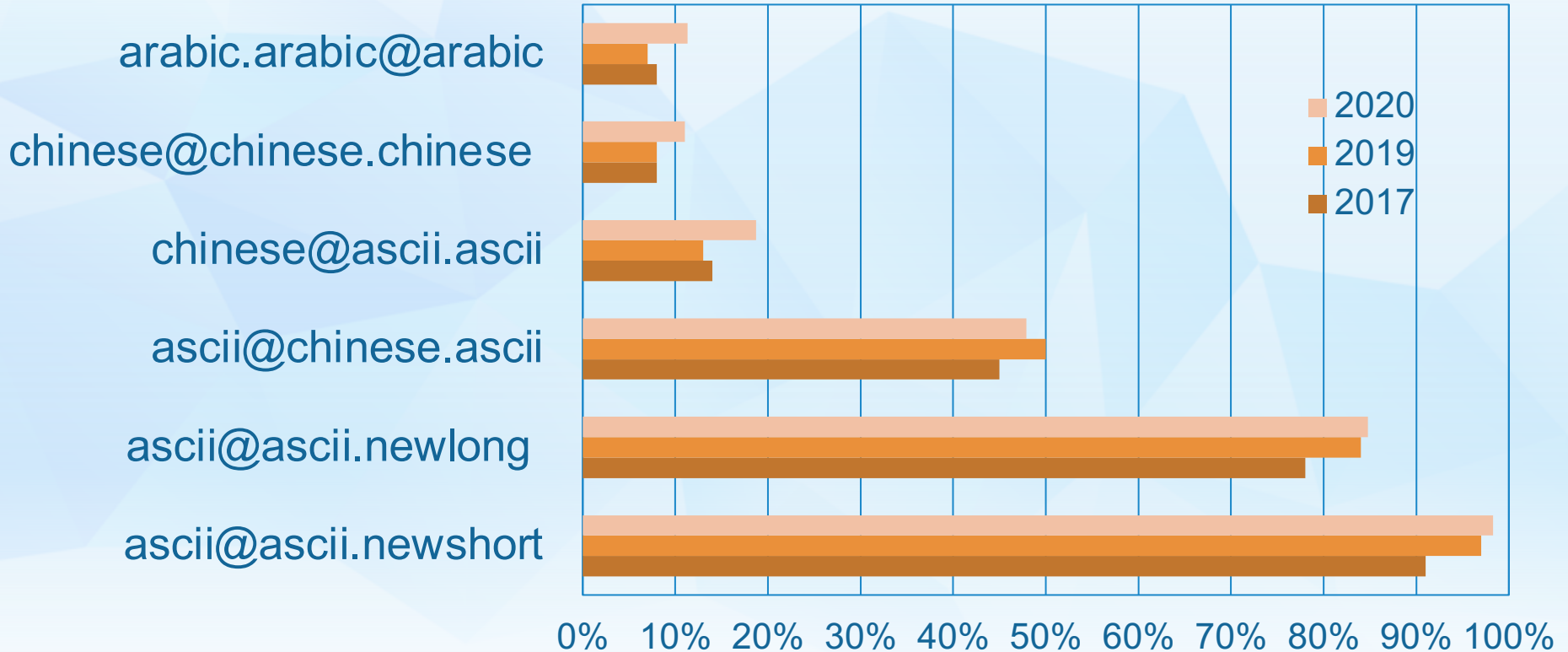
测试@普遍接受-测试.世界

ای-میل@مثال.موقع

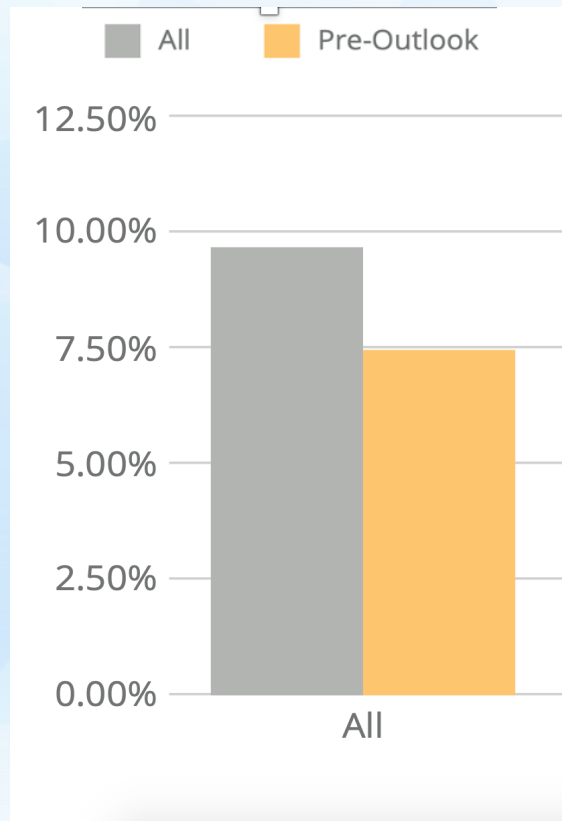
Acceptance of Email Addresses in Websites Globally



For details, see [UASG027](#)



Estimated Support of EAI in Email Systems Under All TLDs



Only 9.7% of the domains sampled were EAI ready in 2019

based on mail servers found through MX records in zones of All TLDs

For details on methodology, see [UASG021D: EAI Readiness in TLDs](#)

- ⦿ **No EAI support** - only ASCII email addresses supported by the tools and services

- ⦿ **Level 1** - can exchange email with EAI addresses
 - Receive email from an EAI address
 - Send email to an EAI address
 - Cannot create mailbox and domain name in UTF8

- ⦿ **Level 2** - Level 1 + can create EAI addresses
 - Receive email from an EAI address
 - Send email to an EAI address
 - Create mailbox and domain name in UTF8

EAI Support by Email Tools and Services

Name	MUA	MSA	MTA	MDA	MSP	Web mail
Coremail	Few	All L2	Most L2	Few	All L2	Most L2
MS Outlook.cpm	Most L1	Most L1	Most L1	None	None	Most L1
Yandex Mail	Few	None	None	Few	Part	Few
Roundcube	Most L2					
Apple Mail	Few					
Apple iOS Mail 14.x	Most L2					
Mozilla Thunderbird	Few					
MS Outlook	Most L1					
MS Exchange Server		All L1	All L1	Few		
Exim		Most L2	All L2			
Postfix		All L2	All L2			
Sendmail		Not tested	Not tested			
Fetchmail				Not tested		
Courier		All L2	All L2	All L2		
Gmail	All L1	All L1	All L1	Few		
XgenPlus		Not tested	Not tested	Not tested	All L2	Not tested

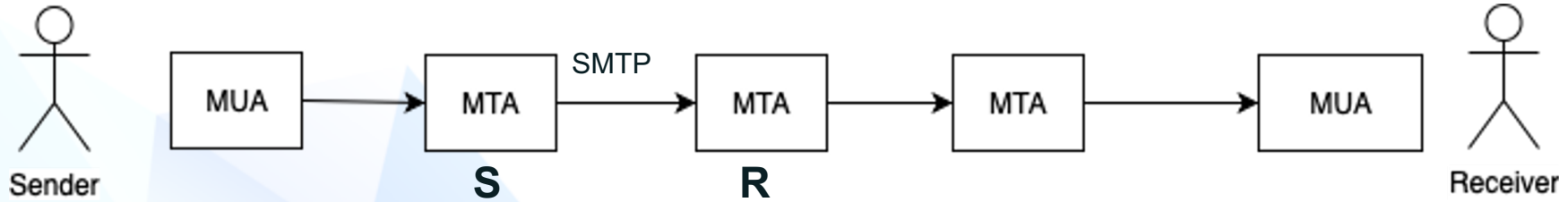
See detailed testing results in [UASG030: Evaluation of EAI Support in Email Software and Services Report](#)

Configuring for EAI

- ◎ SMTP
 - Is augmented to support EAI
 - Has a signaling flag (SMTPUTF8) to specify support of EAI
 - All SMTP servers in the path must support EAI to successfully deliver the email

- ◎ POP/IMAP
 - Are augmented to properly support EAI
 - Have a signaling flag to specify support of EAI
 - Could “half support” EAI by providing a downgraded email version to the non-EAI conforming email software clients

SMTPUTF8 Example

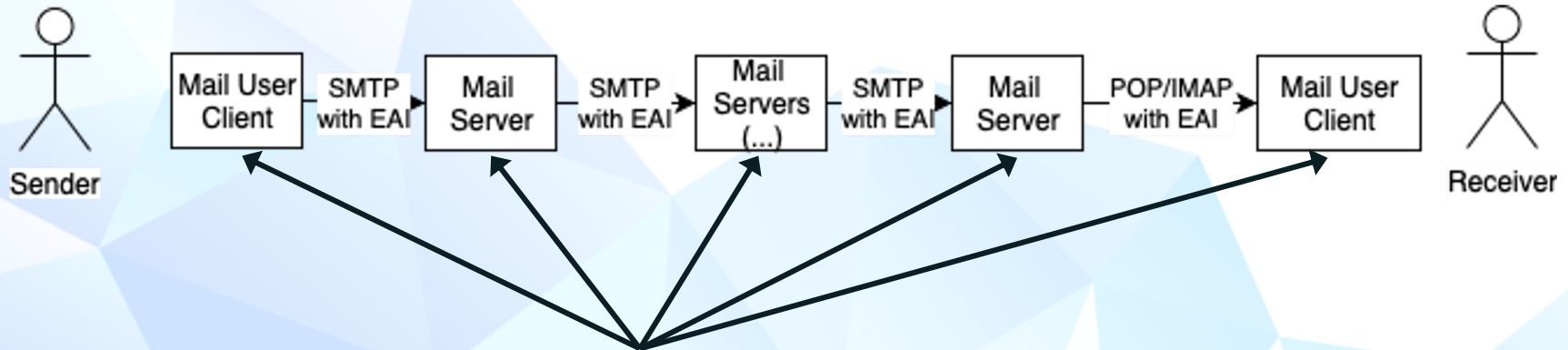


Server S forwarding an email to server R

```
S: <connect>
R: 220 receive.net ESMTP
S: EHLO sender.org
R: 250-8BITMIME
R: 250-SMTPUTF8 ← Specific SMTPUTF8 Signaling (EAI support)
R: 250 PIPELINING
S: MAIL FROM:<猫王@普遍接受-测试.世界> SMTPUTF8
R: 250 Sender accepted
S:RCPT TO:<ray@receive.net>
R:250 Recipient accepted
S:DATA
R:354 Send your message
S:From: 猫王 <猫王@普遍接受-测试.世界>
S:To: ray@receive.net
S:Subject: 我们要吃午饭吗?
S:
S:How about lunch at 12:30?
S:.
R:250 Message accepted 389dck343fg34
S:QUIT
R:221 Sayonara
```

- ⦿ POP:
 - UTF8 command
- ⦿ IMAP
 - ENABLE UTF8 = ACCEPT command

Protocol Changes, Delivery Path Considerations



To send and receive an email with EAI:

- All email parties involved in the delivery path have to be updated for EAI support
- If a single SMTP server in the path does not support EAI, then the email is not delivered

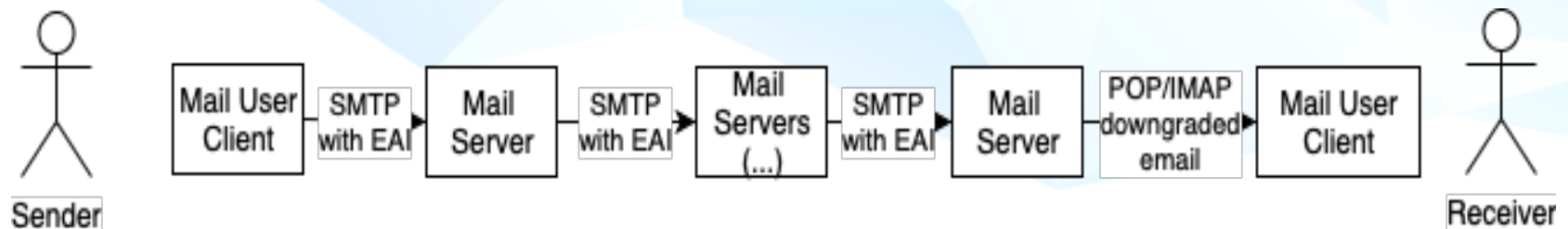
Protocol Changes, Delivery Path Considerations

- What happens when one email (SMTP) server in the path does not support EAI?
 - The last server trying to send to the next hop:
 - Sends back to the sender user a report of unable to deliver.
 - Drops the email.
 - Similar to reports that a sender receives when an email address does not exist.



Protocol Changes, Delivery Path Considerations

- ◉ What happens when the receiver client software (IMAP/POP) does not support EAI?
 - ◉ The IMAP/POP server can be “nice”:
 - By providing a downgraded version of the email. (RFC6857)
 - Changing the EAI to non-EAI version of the local part.
 - ◉ If IMAP/POP server can not be “nice”, then should send a report back to the sender...
 - But that is not always possible as the “mail server” may just be an IMAP/POP server, not SMTP.
- ◉ While it may look interesting, downgrading may cause many issues for the users and the sysadmin to debug issues. Try to avoid using the downgrading mechanism if you can.



- ◎ Case folding
 - In ASCII, email users expect the equivalence of lowercase and uppercase. For example, PETER@example.com and peter@example.com will be delivered to the same mailbox
 - Typically for EAI, such case folding functionality is not automatically implemented in most EAI-ready software

- ◎ SPAM
 - EAI emails may be considered as spam by spam-filtering software even when proper SPF/DKIM records are enabled

- ◎ Software/Services
 - Not every server/client software and services support EAI

Considerations for Mailbox Names Using EAI

Considerations for Mailbox Names Using EAI

- Script Mixing
 - Allow limited script mixing only when clear user need based on local practice
 - Consider security and confusion due to script mixing for mailbox and domain name
- Preventing Invalid and Unstably-Rendered Strings
 - Check if [Reference IDN tables](#) meet desired mailbox string and update as needed
 - Use a string validation tool (e.g. [LGR Tool](#)) to validate the mailbox strings
- Right-to-Left (RTL) Script Consideration
 - Avoid script mixing with right-to-left scripts to avoid confusability and security issues

Considerations for Mailbox Names Using EAI

- ◎ Aliases and Display Names Consideration
 - Consider alias-creation option for the user interface during the mailbox name selection process. ASCII alias can be allowed with an EAI mailbox name
 - Optionally allow the user to add additional aliases at a later time
- ◎ Signs and Symbols
 - Avoid using signs and symbols, especially that do not exist on keyboard/input devices
 - If required for your market, the dot (.), underscore (_), hyphen (-) and plus sign (+) are commonly used
 - Review any additional signs (if needed) and ensure it does not cause a security issue

Get Involved!

- ◉ For more information, email info@uasg.tech or UAProgram@icann.org
- ◉ Access all UA documents and presentations at website: <https://uasg.tech>
- ◉ Access details of ongoing work from wiki pages: <https://community.icann.org/display/TUA>
- ◉ Register to participate or listen in the UA discussion list at: <https://uasg.tech/subscribe>
- ◉ Register to participate in UA working groups [here](#)

- See <https://uasg.tech> for a complete list of reports.
 - Universal Acceptance Quick Guide: [UASG005](#)
 - Introduction to Universal Acceptance: [UASG007](#)
 - Quick Guide to EAI: [UASG014](#)
 - EAI – A Technical Overview: [UASG012](#)
 - EAI – Evaluation of Major Email Software and Services: [UASG021B](#)
 - Universal Acceptance Readiness Framework: [UASG026](#)
 - Considerations for Naming Internationalized Email Mailboxes: [UASG028](#)
 - Evaluation of EAI Support in Email Software and Services Report: [UASG030](#)

Engage with ICANN – Thank You and Questions



One World, One Internet

Visit us at icann.org

Email: champika.wijayatunga@icann.org



[@icann](https://twitter.com/icann)



facebook.com/icannorg



youtube.com/icannnews



flickr.com/icann



linkedin/company/icann



slideshare/icannpresentations



soundcloud/icann