



TLDs and Domain Names

Fahd A. Batayneh | pkSIG 2015 | October 7, 2015

Topics of Discussion

- Top Level Domains (TLDs)
- ccTLDs
- gTLDs
- New gTLDs
- DNS Root Servers
- Registry/Registrar/Registrant
- Dispute Resolution
- Business Opportunities in the Domain Name Industry

What is the DNS?



<http://www.youtube.com/watch?v=72snZctFFtA>

Top Level Domains (TLDs)

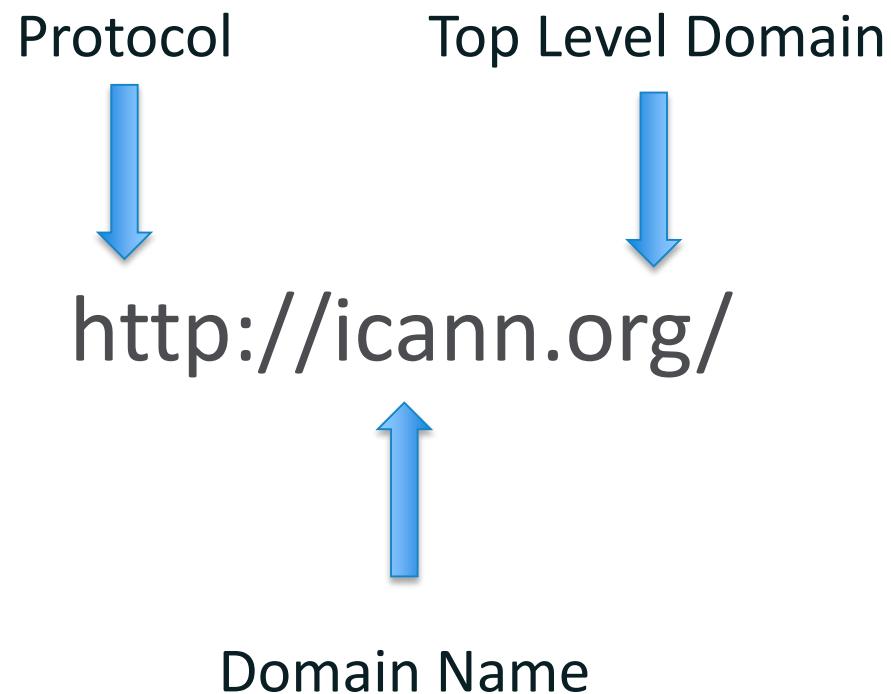
- Exists at the highest level of the DNS hierarchy
- Is the entry installed into the root-zone
- Consists of various groups:
 - ccTLDs (.pk, .kw, .jo, .eg, .tn)
 - gTLDs
 - Sponsored (.asia, .cat)
 - Unsponsored (.com, .net, .org)
 - New (.apps, .shop, .موقع, .شبكة)
 - IDN TLDs (.تونس, .تونسي, .تونسي, .تونسي)

What is a Domain Name?

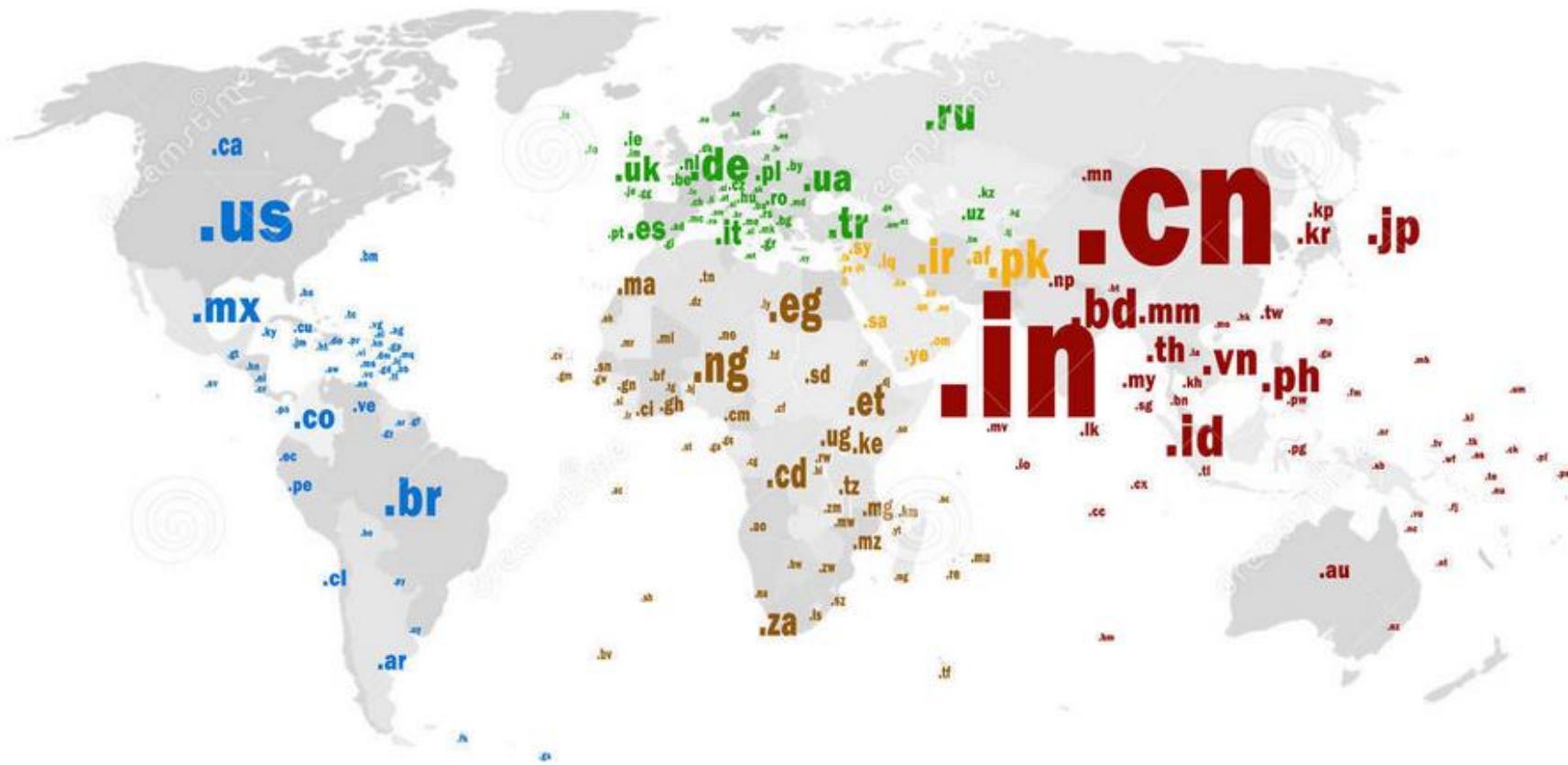


<http://www.youtube.com/watch?v=2ZUQ2Szu-JI>

Parts of a Domain Name?



Country Code TLDs (ccTLD)



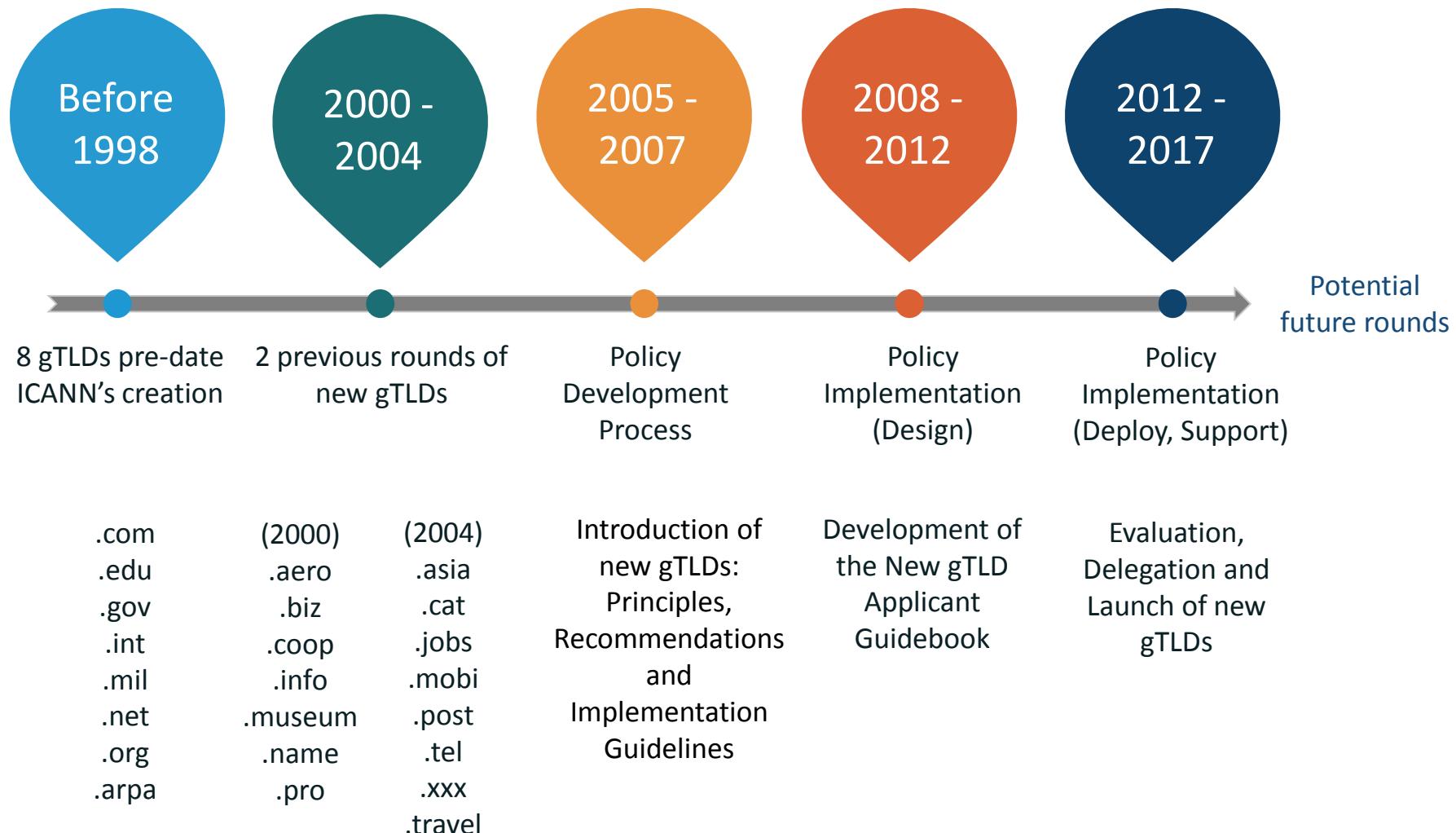
Source at <http://dreamstime.com/>



Where do ccTLDs get Their Code From?

- The ISO-3166 provides 2-letter codes and 3-letter codes of countries or a grouping of countries
- 249 assigned 2-letter codes
 - According to the UN, the world has 196 countries
- More at <https://www.iso.org/obp/ui/#search>

History of TLD



The 2012 New gTLDs Program

- Considered the 3rd round of New TLDs
- More than 5 years of community collaboration to come out with the “TLD Applicant Guidebook”
- A one-time application fee of USD 185,000, and an annual fee of USD 25,000
- Allowed the application of non-Latin names
- 1,930 applications received for 1,420 unique strings
 - 116 non-Latin strings

... cont. (The 2012 New gTLDs Program)

- A rigorous application and evaluation process
 - To ensure that TLDs remain operational and that registrants do not lose their domain names due to the instability of a TLD
- The New gTLD microsite at <http://newgtlds.icann.org/>
- More statistics on the program can be found at <http://newgtlds.icann.org/en/program-status/statistics>

Understanding New gTLDs



New gTLDs Application Breakdown by Region

1930

total number of applications received

911

North America

675

Europe

24

South America

17

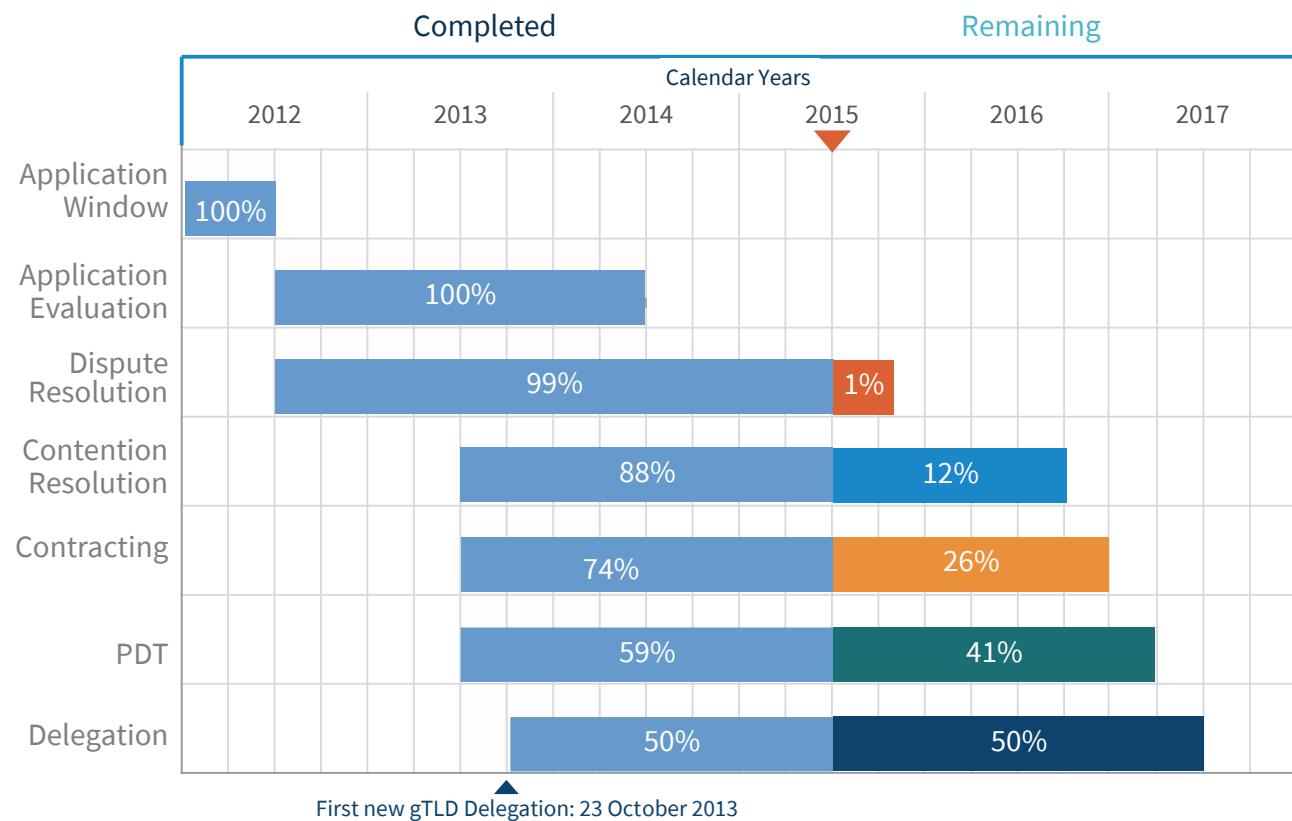
Africa

303

Asia Pacific

Where do Things Stand At?

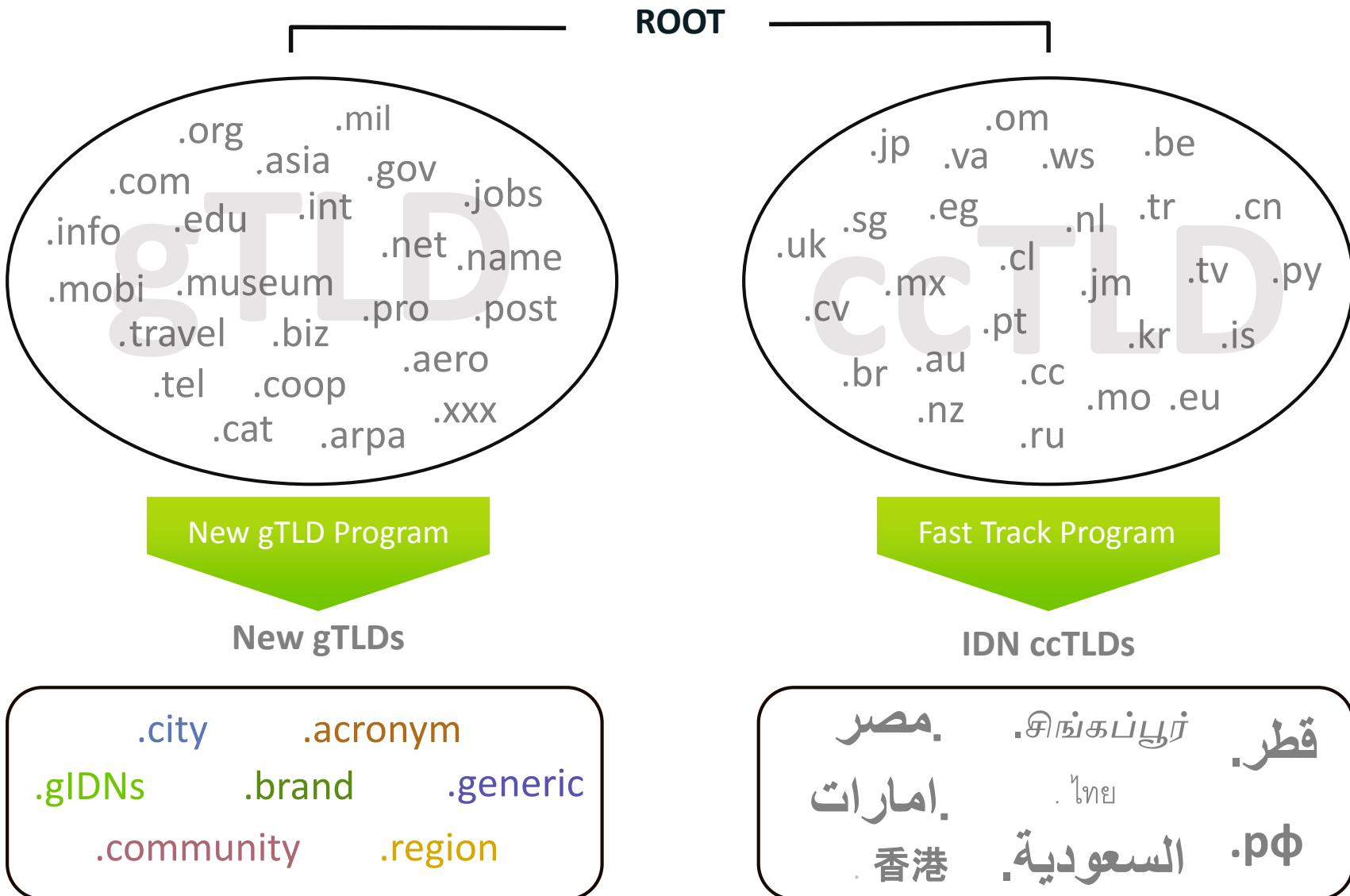
- 1930 applications
- 1300+ potential TLDs delegated by 2017
- 750 new gTLDs delegated (as of 18 September 2015)
- More than 50% of the IDN gTLDs have been delegated (103 applications)



The New Landscape



How does the Name Space Look Nowadays?



Root Servers

- Publish the root zone file to other DNS servers and clients on the Internet
- The root zone file describes where the authoritative servers for the DNS TLDs are located
- The root name server operators publish the root zone file as received from the IANA

Root Server Operators

1. A - VeriSign Global Registry Services
2. B - University of Southern California - Information Sciences Institute
3. C - Cogent Communications
4. D - University of Maryland
5. E - NASA Ames Research Center
6. F - Internet Systems Consortium, Inc.
7. G - U.S. DOD Network Information Center
8. H - U.S. Army Research Lab 13 root-servers
9. I - Autonomica/NORDUnet 12 Operators
10. J - VeriSign Global Registry Services
11. K - RIPE NCC
12. L – ICANN
13. M - WIDE Project

Root Servers Around the World (~480 instances)



Source at <http://root-servers.org/>

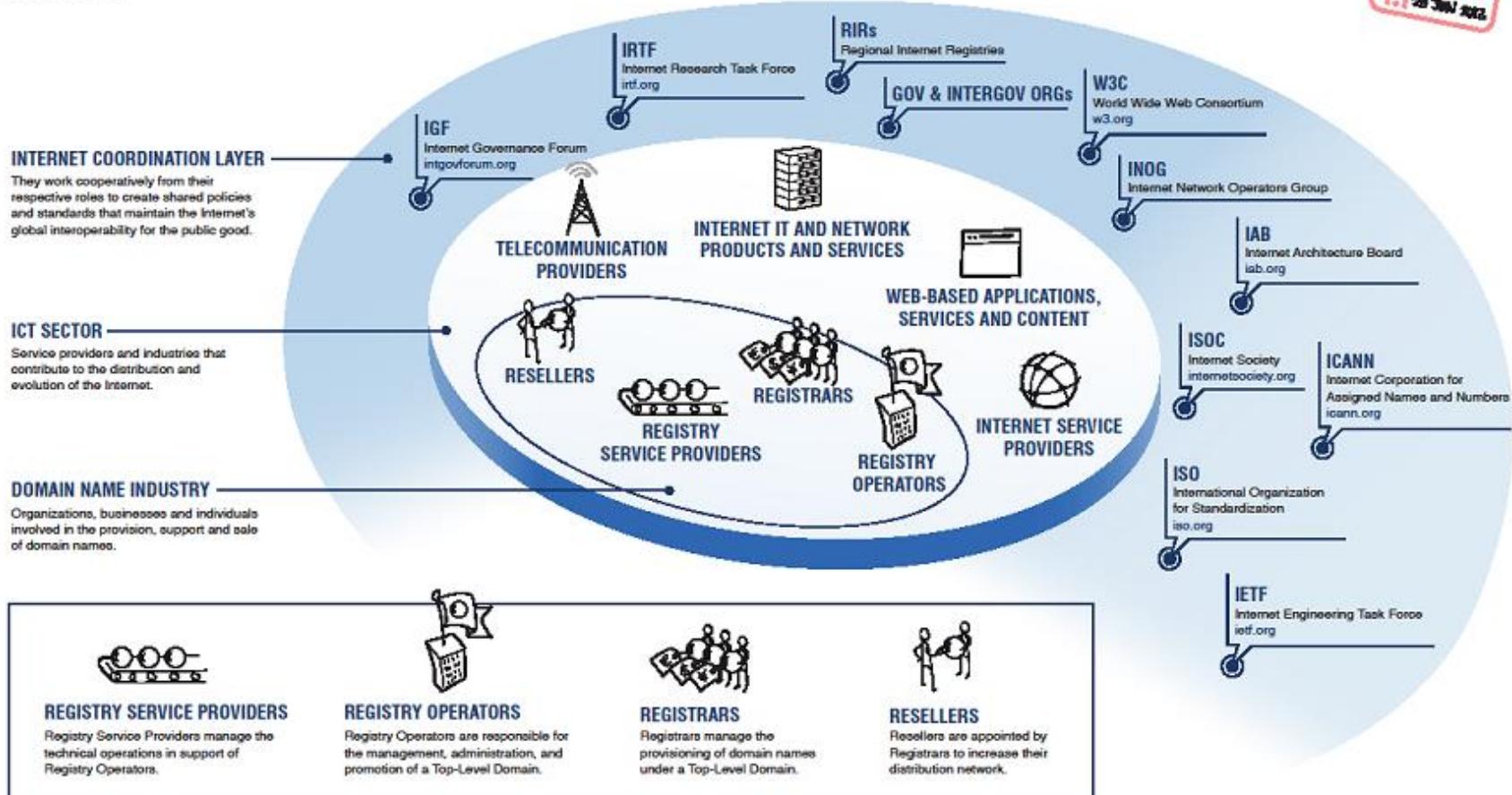
Registry/Registrar/Registerant

- **Registry** – The entity that operates a TLD
- **Registrar** – A middle-agent that sells domain names to registrants on behalf of a TLD Registry
 - ICANN Accredited Registrar vs. Resellers
 - ccTLDs vs. gTLDs
- **Registrant** – Entity registering a domain name

The DNS Industry Ecosystem

THE DOMAIN NAME INDUSTRY ECOSYSTEM

VERSION
1.1
28 JUL 2012



This graphic is a living document, designed to provide a high level view of the relationship between the different parties of the Domain Name Industry. It is for illustrative purposes only and is not intended to be a definitive guide. Some of the names of the documents may vary. Please provide feedback at www.xplanations.com/domainnameindustry.

2013 | Creative Commons Attribution-ShareAlike 3.0

Dispute Resolution

- Cases could arise where a registrant registers a domain name that could infringe the naming rights of others. This is also called “Cybersquatting”
- When ICANN was formed in 1998, one of the first things it worked on was the “Uniform Domain Name Dispute Resolution Process (UDRP)”
 - 5 accredited providers for gTLDs to date
 - ccTLDs can customize this to fit their needs and local laws
 - An expensive and sometimes long process
- All 5 providers listed at <http://bit.ly/1j6Rehz>

New gTLDs and Dispute Resolution

- With the introduction of 100s of New gTLDs, cybersquatting is on the high
- Several protection and DRP mechanisms devised
 - Trademark Clearinghouse (TMCH)
 - Uniform Rapid Suspension (URS)
 - Post Delegation Dispute Resolution Procedure (PDDRP)
- UDRP would not be the best path for clear cut cases that need urgent attention
 - TMCH is good for initial protection, while URS comes in for a faster and cheaper path for resolving disputes

Trademark Clearinghouse (TMCH)



<http://www.youtube.com/watch?v=ZXP2dieeZes>

Business Opportunities in the DNS Industry

- Registry Services
 - 4 ccTLDs in the region have state-of-the-art registries. Many others are moving to a shared system
 - Having a Registry/Registrar model is the way forward
- New gTLDs
 - Providing more options in the TLD space away from the traditional .com/.net/.org and the ccTLDs, and providing multilingual TLDs
- Accredited Registrars
 - More than 1000 registrars
 - More at <http://www.icann.org/en/resources/registrars/accreditation>

... continue (Business Opportunities)

- Dispute Resolution Services
 - Five providers in total the last of which joined in 2013
- Online Intellectual Property Protection Services
 - Domain names are digital assets online
- Domain Name Aftermarket
 - **insurance.com** sold for USD 35.6 million in 2010
 - **360.com** sold for USD 17 million in 2015
- Development of Domain Name Tools and Software

Domain Name Ecosystem in Pakistan

- ccTLD is .pk, and is run by PKNIC
- IDN ccTLD applied for is بیاکستان in Urdu by the MoITT
- 1 ICANN Accredited Registrar
- 5 root-server instances (2 in Karachi | 2 in Islamabad | 1 in Lahore)
- Membership from MoITT and PTA in the GAC
- ISOC Islamabad Chapter is an ALS
- Many Pakistani stakeholders attend ICANN meetings; mainly through the ICANN fellowship program

Questions?!



ICANN

Fahd Batayneh

Manager, Stakeholder Engagement, Middle East

E: fahd.batayneh@icann.org

W: <http://icann.org/>



twitter.com/icann

twitter.com/fahdbatayneh



gplus.to/icann



facebook.com/icannorg



weibo.com/ICANNorg



linkedin.com/company/icann



flickr.com/photos/icann



youtube.com/user/icannnews



slideshare.net/icannpresentations



Internationalized Domain Names (IDNs)

Sarmad Hussain | pkSIG 2015 | October 7, 2015

Internationalized Domain Names (IDNs)



<http://www.youtube.com/watch?v=wnauGpYh96c>

Overview of Presentation

○ IDNs at Top Level

- IDN TLD Program
 - Label Generation Ruleset (LGR)
 - LGR Toolset
 - IDN Variant Implementation
- IDN ccTLD Fast Track Process Implementation



○ IDNs at Second Level for gTLDs

- IDN Implementation Guidelines
- Reference LGR

○ Community Outreach and Involvement

ASCII Domain Name Label

www.cafe.com

Third Level
Domain

Second Level
Domain

Top Level
Domain (TLD)

Top Level Domains (TLDs)

- ◎ Country Code TLDs (ccTLDs)
 - ◎ .sg, .cn, .kh, .la, .mm, .th, .ca, ...
 - ◎ Two letter [a..z] codes, reserved for countries and territories by ISO 3166 standard
- ◎ Generic TLDs (gTLDs)
 - ◎ .com, .org, .net, .edu, ... - organizations
 - ◎ New gTLDs – 1930 applications in 2012

Domain Stakeholders

- ◎ ICANN
- ◎ Registry
- ◎ Registrar
- ◎ Reseller
- ◎ Registrant
- ◎ End-User

ASCII Domain Name Label

www.cafe.com

Third Level
Domain

Second Level
Domain

Top Level
Domain (TLD)

Forming ASCII Labels

Use LDH

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

Label length = 63

Other constraints (e.g. on hyphen)

Forming ASCII Labels

Use only Letters

- Letters [a-z]

Label length = 63

Internationalized Domain Name (IDN) Labels

ตัวอย่าง-ไทย

IDN Second
Level
Domain

IDN Top
Level
Domain

বাংলা

Бел

الجزائر

հայ

中國

ભારત

한국

ଓঁকো

Syntax of IDN Labels

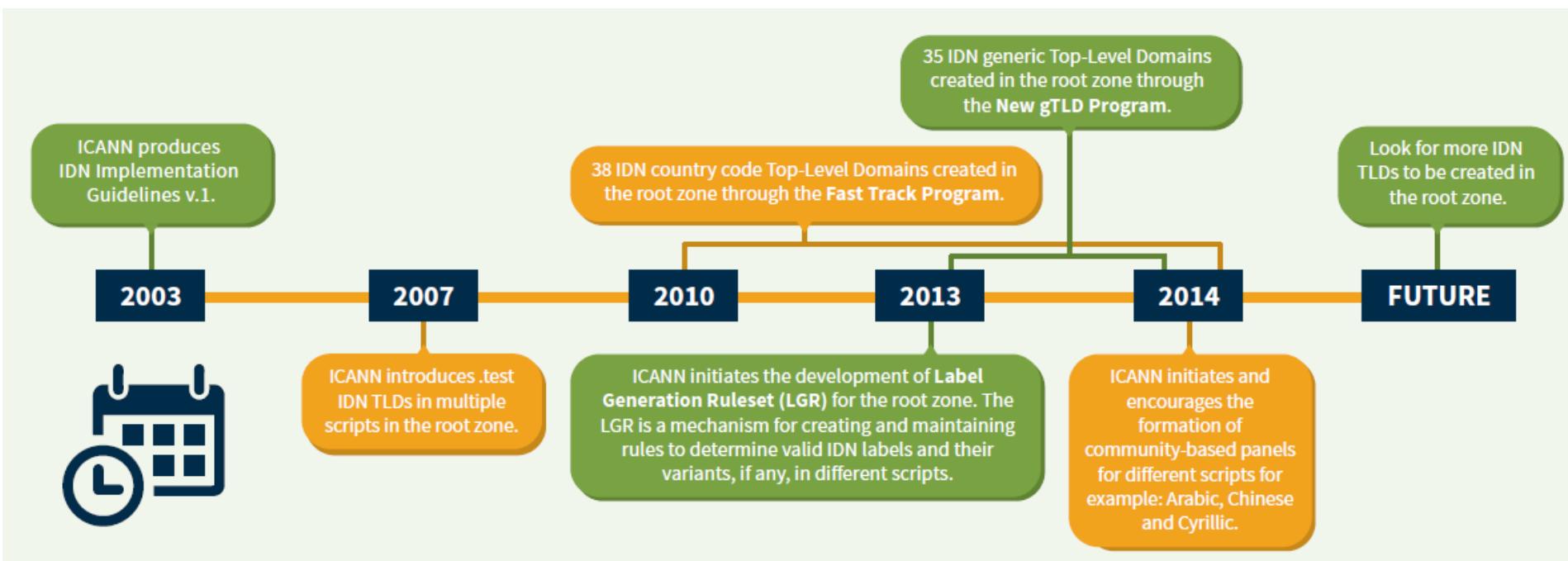
Valid U-Label: Unicode code points as constrained by IDNA2008

Valid A-Label - “xn--” followed by punycode of U-Label of length 59

Syntax of IDN Labels

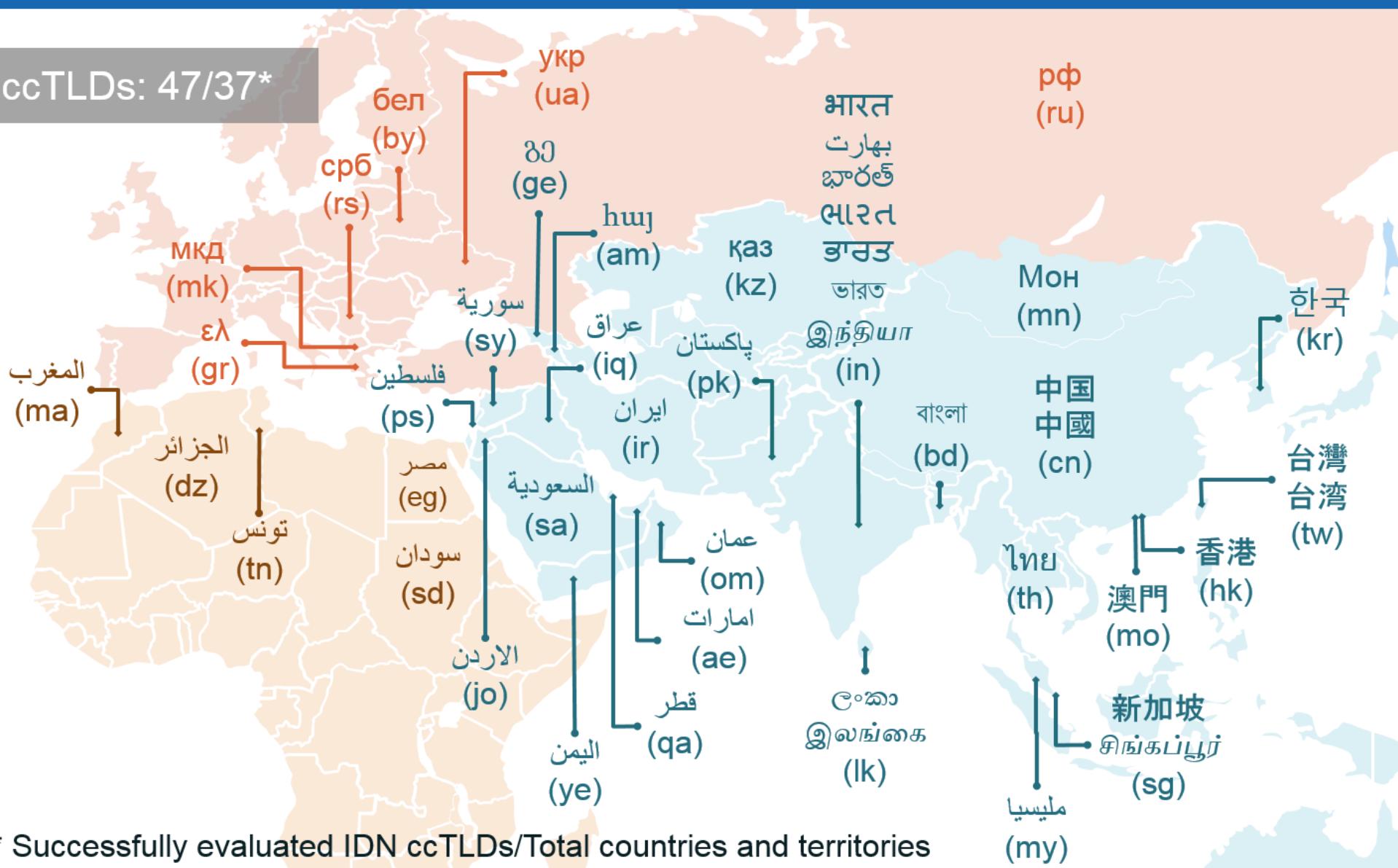
Valid U-Label, further constrained by the “letter” principle for TLDs
Valid A-Label

Timeline of ICANN's Progress on IDN TLDs



IDN Country Code Top-Level Domains

ccTLDs: 47/37*

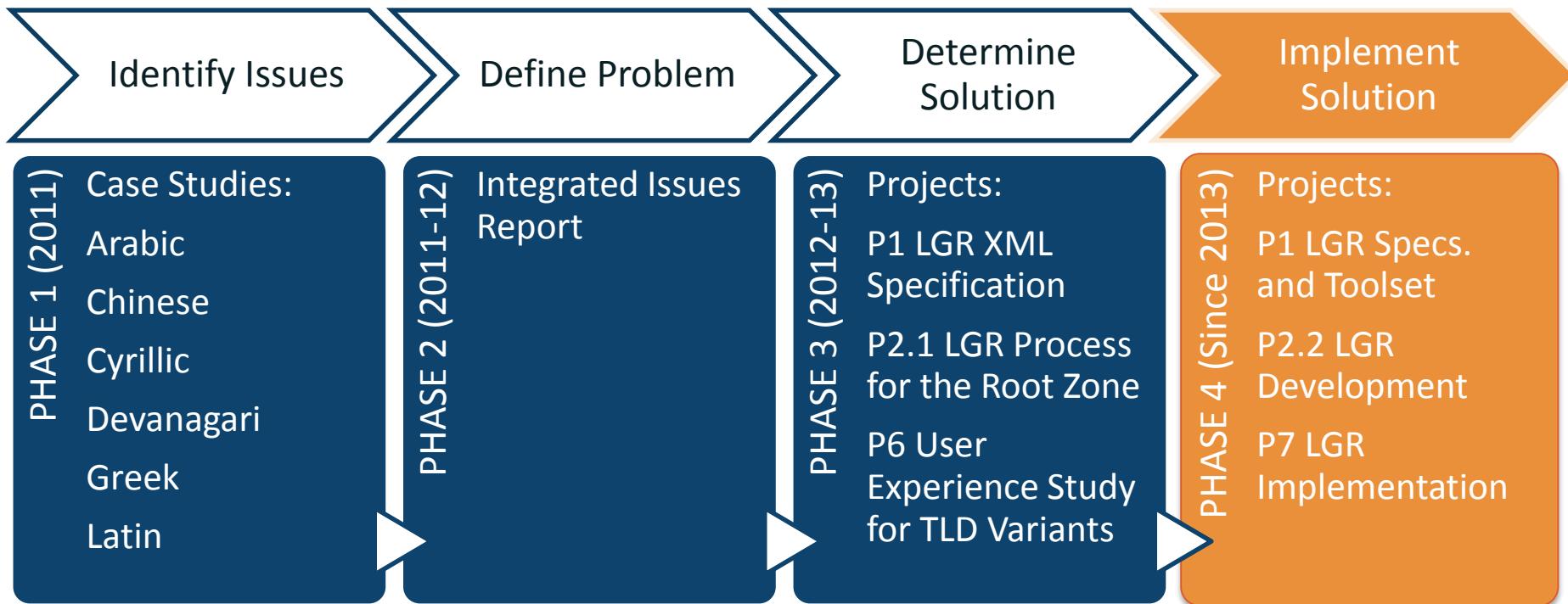


Successfully evaluated IDN ccTLDs/Total countries and territories

The background of the slide features a complex, abstract network graph. It consists of numerous small, semi-transparent circular nodes of varying sizes scattered across the frame, connected by a dense web of thin, light-colored lines that represent connections or links between the nodes. This visual metaphor represents the interconnected nature of the Internet and the global reach of the IDN TLD Program.

IDN TLD Program

IDN TLD Program



Community agreed to define a Label Generation Rules (LGR)

Reports and documentation of all completed projects available at:
<https://www.icann.org/resources/pages/reports-2013-04-03-en>

Label Generation Rules for the Root Zone

- For the Root Zone, single “table” containing data for all scripts
 - As it is a shared resource, must be conservative
 - Must be stable and secure
 - Must be based on inclusion based analysis
- For each script or writing system:
 - Which code points are valid for use?
 - Are any of these code points variants of each other?
 - Are there any additional constraints on the labels?

LGR for the Root Zone

Unicode

	000	001	002	003	004	005	006	007
0	NUL	DLE	SP	0	@	P	`	p
0000	0010	0020	0030	0040	0050	0060	0070	
1	SOH	DC1	!	1	A	Q	a	q
0001	0011	0021	0031	0041	0051	0061	0071	
2	STX	DC2	"	2	B	R	b	r
0002	0012	0022	0032	0042	0052	0062	0072	
3	ETX	DC3	#	3	C	S	c	s
0003	0013	0023	0033	0043	0053	0063	0073	
4	EOT	DC4	\$	4	D	T	d	t
0004	0014	0024	0034	0044	0054	0064	0074	
5	ENQ	NAK	%	5	E	U	e	u
0005	0015	0025	0035	0045	0055	0065	0075	
6	ACK	SYN	&	6	F	V	f	v
0006	0016	0026	0036	0046	0056	0066	0076	
7	BEL	ETB	'	7	G	W	g	w
0007	0017	0027	0037	0047	0057	0067	0077	
8	BS	CAN	(8	H	X	h	x
0008	0018	0028	0038	0048	0058	0068	0078	
9	HT	EM)	9	I	Y	i	y
0009	0019	0029	0039	0049	0059	0069	0079	
A	LF	SUB	*	:	J	Z	j	z
000A	001A	002A	003A	004A	005A	006A	007A	
B	VT	ESC	+	;	K	[k	{
000B	001B	002B	003B	004B	005B	006B	007B	
C	FF	FS	,	<	L	\	l	l
000C	001C	002C	003C	004C	005C	006C	007C	
D	CR	GS	-	=	M]	m	}
000D	001D	002D	003D	004D	005D	006D	007D	
E	SO	RS	.	>	N	^	n	~
000E	001E	002E	003E	004E	005E	006E	007E	
F	SI	US	/	?	O	-	o	DEL
000F	001F	002F	003F	004F	005F	006F	007F	

• • •

	OE8	OE9	OEA	OEB	OEC	OED	OEE	OEF
0				߻	߳	ߴ	ߵ	
00B0				߻	߳	ߴ	ߵ	
1	߱			߲	߳	ߴ	ߵ	
00B1				߲	߳	ߴ	ߵ	
2			߳	߳	߳	߳	߳	
00B2			߳	߳	߳	߳	߳	
3				߳	߳	߳	߳	
00B3				߳	߳	߳	߳	
4	߳	߳	߳	߳	߳	߳	߳	
00B4			߳	߳	߳	߳	߳	
5				߳	߳	߳	߳	
00B5				߳	߳	߳	߳	
6			߳	߳	߳	߳	߳	
00B6			߳	߳	߳	߳	߳	
7				߳	߳	߳	߳	
00B7				߳	߳	߳	߳	
8				߳	߳	߳	߳	
00B8				߳	߳	߳	߳	
9				߳	߳	߳	߳	
00B9				߳	߳	߳	߳	
A			߳	߳	߳	߳	߳	
00BA			߳	߳	߳	߳	߳	
B				߳	߳	߳	߳	
00BB				߳	߳	߳	߳	
C			߳	߳	߳	߳	߳	
00BC			߳	߳	߳	߳	߳	
D				߳	߳	߳	߳	
00BD				߳	߳	߳	߳	
E				߳	߳	߳	߳	
00BE				߳	߳	߳	߳	
F				߳	߳	߳	߳	
00BF				߳	߳	߳	߳	

• • •

LGR for the Root Zone

Unicode

IDNA2008 – by IETF

LGR for the Root Zone

Unicode

IDNA2008

Maximal Starting Repertoire – by Integration Panel of ICANN

LGR for the Root Zone

LGR Proposal – by **Generation Panel** of Script Community

Unicode

IDNA2008

Maximal Starting Repertoire (MSR)

X	X	X
X		
X	X	
X		
X		X X X
X		X
X		X

060 061 062 063 064 065 066 067 068 069 06A 06B 06C 06D 06E 06F																075 076 077			08A 08B 08C 08D 08E 08F					
۰ 0600	۱ 0610	۲ 0620	۳ 0630	۴ 0640	۵ 0650	۶ 0660	۷ 0670	۸ 0680	۹ 0690	۰ 06A0	۱ 06B0	۲ 06C0	۳ 06D0	۴ 06E0	۵ 06F0	۰ 0750	۱ 0760	۲ 0770	۰ 08A0	۱ 08B0	۲ 08C0	۳ 08D0	۴ 08E0	۵ 08F0
۶ 0601	۷ 0611	۸ 0621	۹ 0631	۰ 0641	۱ 0651	۲ 0661	۳ 0671	۴ 0681	۵ 0691	۶ 06A1	۷ 06B1	۸ 06C1	۹ 06D1	۰ 06E1	۱ 06F1	۶ 0751	۷ 0761	۸ 0771	۶ 08A2	۷ 08B2	۸ 08C2	۹ 08D2	۰ 08E2	۱ 08F1
۲ 0602	۳ 0612	۴ 0622	۵ 0632	۶ 0642	۷ 0652	۸ 0662	۹ 0672	۰ 0682	۱ 0692	۲ 06A2	۳ 06B2	۴ 06C2	۵ 06D2	۶ 06E2	۷ 06F2	۲ 0752	۳ 0762	۴ 0772	۲ 08A3	۳ 08B3	۴ 08C3	۵ 08D3	۶ 08E3	۷ 08F2
۳ 0603	۴ 0613	۵ 0623	۶ 0633	۷ 0643	۸ 0653	۹ 0663	۰ 0673	۱ 0683	۲ 0693	۳ 06A3	۴ 06B3	۵ 06C3	۶ 06D3	۷ 06E3	۸ 06F3	۳ 0753	۴ 0763	۵ 0773	۳ 08A4	۴ 08B4	۵ 08C4	۶ 08D4	۷ 08E4	۸ 08F4
۴ 0604	۵ 0614	۶ 0624	۷ 0634	۸ 0644	۹ 0654	۰ 0664	۱ 0674	۲ 0684	۳ 0694	۴ 06A4	۵ 06B4	۶ 06C4	۷ 06D4	۸ 06E4	۹ 06F4	۴ 0754	۵ 0764	۶ 0774	۴ 08A5	۵ 08B5	۶ 08C5	۷ 08D5	۸ 08E5	۹ 08F5
۵ 0605	۶ 0615	۷ 0625	۸ 0635	۹ 0645	۰ 0655	۱ 0665	۲ 0675	۳ 0685	۴ 0695	۵ 06A5	۶ 06B5	۷ 06C5	۸ 06D5	۹ 06E5	۰ 06F5	۵ 0755	۶ 0765	۷ 0775	۵ 08A6	۶ 08B6	۷ 08C6	۸ 08D6	۹ 08E6	۰ 08F6
۶ 0606	۷ 0616	۸ 0626	۹ 0636	۰ 0646	۱ 0656	۲ 0666	۳ 0676	۴ 0686	۵ 0696	۶ 06A6	۷ 06B6	۸ 06C6	۹ 06D6	۰ 06E6	۱ 06F6	۶ 0756	۷ 0766	۸ 0776	۶ 08A7	۷ 08B7	۸ 08C7	۹ 08D7	۰ 08E7	۱ 08F7
۷ 0607	۸ 0617	۹ 0627	۰ 0637	۱ 0647	۲ 0657	۳ 0667	۴ 0677	۵ 0687	۶ 0697	۷ 06A7	۸ 06B7	۹ 06C7	۰ 06D7	۱ 06E7	۲ 06F7	۷ 0757	۸ 0767	۹ 0777	۷ 08A8	۸ 08B8	۹ 08C8	۰ 08D8	۱ 08E8	۲ 08F8
۸ 0608	۹ 0618	۰ 0628	۱ 0638	۲ 0648	۳ 0658	۴ 0668	۵ 0678	۶ 0688	۷ 0698	۸ 06A8	۹ 06B8	۰ 06C8	۱ 06D8	۲ 06E8	۳ 06F8	۸ 0758	۹ 0768	۰ 0778	۸ 08A9	۹ 08B9	۰ 08C9	۱ 08D9	۲ 08E9	۳ 08F9
۹ 0609	۰ 0619	۱ 0629	۲ 0639	۳ 0649	۴ 0659	۵ 0669	۶ 0679	۷ 0689	۸ 0699	۹ 06A9	۰ 06B9	۱ 06C9	۲ 06D9	۳ 06E9	۴ 06F9	۹ 0759	۰ 0769	۱ 0779	۹ 08AA	۰ 08BA	۱ 08CA	۲ 08DA	۳ 08EA	۴ 08FA
A 060A	B 061A	C 062A	D 063A	E 064A	F 065A	G 066A	H 067A	I 068A	J 069A	K 06AA	L 06BA	M 06CA	N 06DA	O 06EA	P 06FA	A 075A	B 076A	C 077A	A 08AB	B 08AC	C 08AD	D 08AE	E 08AF	F 08BF
F 060B	G 061B	H 062B	I 063B	J 064B	K 065B	L 066B	M 067B	N 068B	O 069B	P 06AB	۰ ۰6B8	۱ ۰6C8	۲ ۰6D8	۳ ۰6E8	۴ ۰6F8	F 075B	G 076B	H 077B	F 08AB	G 08AC	H 08AD	I 08AE	J 08AF	K 08BF
C 060C	D 061C	E 062C	F 063C	G 064C	H 065C	I 066C	J 067C	K 068C	L 069C	M 06AC	N 06BC	O 06CC	P 06DC	۰ ۰6EC	۱ ۰6FC	C 075C	D 076C	E 077C	C 08AC	D 08AD	E 08AE	F 08AF	G 08BF	H 08CF
D 060D	E 061D	F 062D	G 063D	H 064D	I 065D	J 066D	K 067D	L 068D	M 069D	N 06AD	O 06BD	P 06CD	۰ ۰6DD	۱ ۰6ED	۲ ۰6FD	D 075D	E 076D	F 077D	D 08AD	E 08AE	F 08AF	G 08BF	H 08CF	I 08DF
E 060E	F 061E	G 062E	H 063E	I 064E	J 065E	K 066E	L 067E	M 068E	N 069E	O 06AE	P 06BE	۰ ۰6CE	۱ ۰6DE	۲ ۰6EE	۳ ۰6FE	E 075E	F 076E	G 077E	E 08AD	F 08AC	G 08AD	H 08AF	I 08BF	J 08CF
F 060F	G 061F	H 062F	I 063F	J 064F	K 065F	L 066F	M 067F	N 068F	O 069F	P 06AF	۰ ۰6BF	۱ ۰6CF	۲ ۰6DF	۳ ۰6EF	۴ ۰6FF	F 075F	G 076F	H 077F	F 08AD	G 08AC	H 08AD	I 08AF	J 08BF	K 08CF

Label Generation Rules (LGR)

۵۳۲	۰۶۴۳	ک	۰۶۵۲	۰۶۶۲	۰۶۷۲	۰۶۸۲	۰۶۹۲	۰۶۸۲	۰۶۸۲	۰۶۸۲	۰۶۸۲
۵۳۳	۰۶۵۳	۰۶۴۳	۰۶۶۳	۰۶۷۳	۰۶۸۳	۰۶۹۳	۰۶۹۳	۰۶۸۳	۰۶۸۳	۰۶۸۳	۰۶۸۳
۵۳۴	۰۶۴۴	ل	۰۶۵۴	۰۶۶۴	۰۶۷۴	۰۶۸۴	۰۶۹۴	۰۶۹۴	۰۶۸۴	۰۶۸۴	۰۶۸۴
۵۳۵	۰۶۴۵	م	۰۶۵۵	۰۶۶۵	۰۶۷۵	۰۶۸۵	۰۶۹۵	۰۶۹۵	۰۶۸۵	۰۶۸۵	۰۶۸۵
۵۳۶	۰۶۴۶	ن	۰۶۵۶	۰۶۶۶	۰۶۷۶	۰۶۸۶	۰۶۹۶	۰۶۹۶	۰۶۸۶	۰۶۸۶	۰۶۸۶
۵۳۷	۰۶۴۷	ه	۰۶۵۷	۰۶۶۷	۰۶۷۷	۰۶۸۷	۰۶۹۷	۰۶۹۷	۰۶۸۷	۰۶۸۷	۰۶۸۷
۵۳۸	۰۶۴۸	و	۰۶۵۸	۰۶۶۸	۰۶۷۸	۰۶۸۸	۰۶۹۸	۰۶۹۸	۰۶۸۸	۰۶۸۸	۰۶۸۸
۵۳۹	۰۶۴۹	ی	۰۶۵۹	۰۶۶۹	۰۶۷۹	۰۶۸۹	۰۶۹۹	۰۶۹۹	۰۶۸۹	۰۶۸۹	۰۶۸۹
۵۳A	۰۶۴A	ی	۰۶۵A	۰۶۶A	۰۶۷A	۰۶۸A	۰۶۹A	۰۶۹A	۰۶۸A	۰۶۸A	۰۶۸A

- Valid code points
- Variants code points

پاکستان

پاکستان

- Label constraints

- Cannot mix ک and ڪ in a label

کلکته ✓

کلکته ✓

کلکته ✗

کلکته ✗

LGR Specification

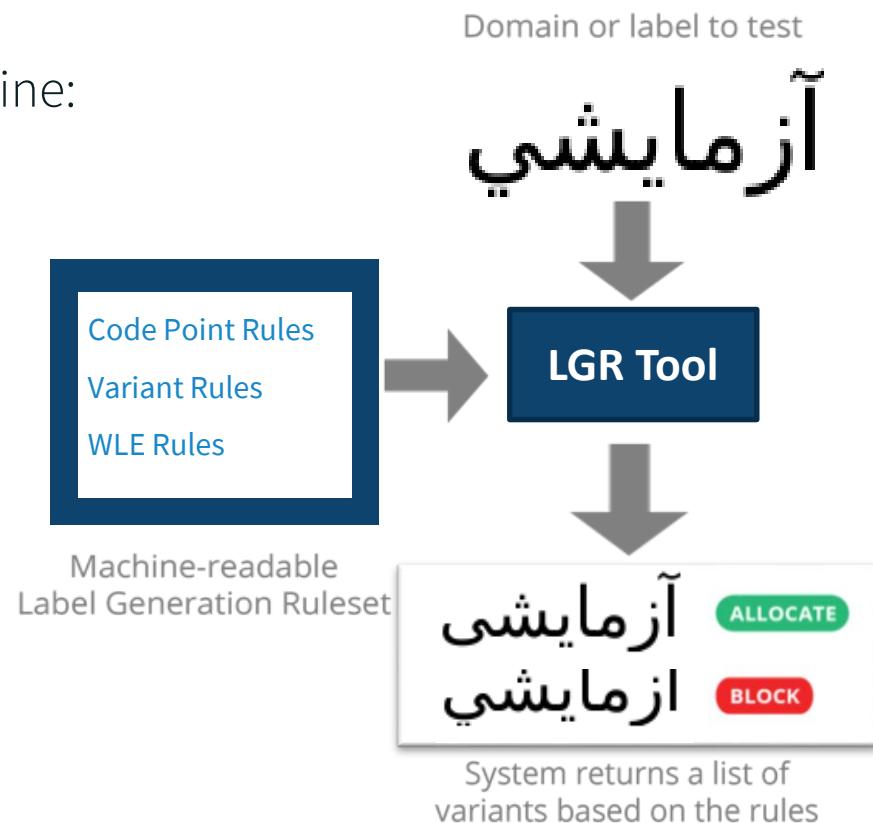
- ◎ LGR machine-readable specifications at
<https://datatracker.ietf.org/doc/draft-davies-idntables>
- ◎ Example: excerpt from MSR-2 XML file

```
...
<range first-cp="0780" last-cp="07B0" tag="sc:Thaa" ref="3"/>
<char cp="07B1" tag="sc:Thaa" ref="5"/>
<char cp="08A0" tag="sc:Arab" ref="12"/>
<range first-cp="08A2" last-cp="08AC" tag="sc:Arab" ref="12"/>
<range first-cp="08E4" last-cp="08EF" tag="sc:Arab" ref="12"/>
<range first-cp="08F4" last-cp="08FE" tag="sc:Arab" ref="12"/>
<range first-cp="0901" last-cp="0903" tag="sc:Deva" ref="0"/>
<char cp="0904" tag="sc:Deva" ref="6"/>
<range first-cp="0905" last-cp="0939" tag="sc:Deva" ref="0"/>
<range first-cp="093A" last-cp="093B" tag="sc:Deva" ref="11"/>
<char cp="093C" tag="sc:Deva" ref="0"/>
<range first-cp="093E" last-cp="094D" tag="sc:Deva" ref="0"/>
<char cp="094F" tag="sc:Deva" ref="11"/>
<range first-cp="0956" last-cp="0957" tag="sc:Deva" ref="11"/>
<char cp="0972" tag="sc:Deva" ref="9"/>
<range first-cp="0973" last-cp="0977" tag="sc:Deva" ref="11"/>
<range first-cp="0979" last-cp="097A" tag="sc:Deva" ref="10"/>
<range first-cp="097B" last-cp="097C" tag="sc:Deva" ref="8"/>
<range first-cp="097E" last-cp="097F" tag="sc:Deva" ref="8"/>
<range first-cp="0981" last-cp="0983" tag="sc:Beng" ref="0"/>
...
...
```

LGR Specification and Toolset

- ◎ LGR machine-readable specifications at:
<https://datatracker.ietf.org/doc/draft-davies-idntables>
 - LAGER WG at IETF
- ◎ Open source LGR Toolset tentative timeline:
 - Create LGR - available
 - Use LGR – 12/15
 - Manage LGRs – 3/16

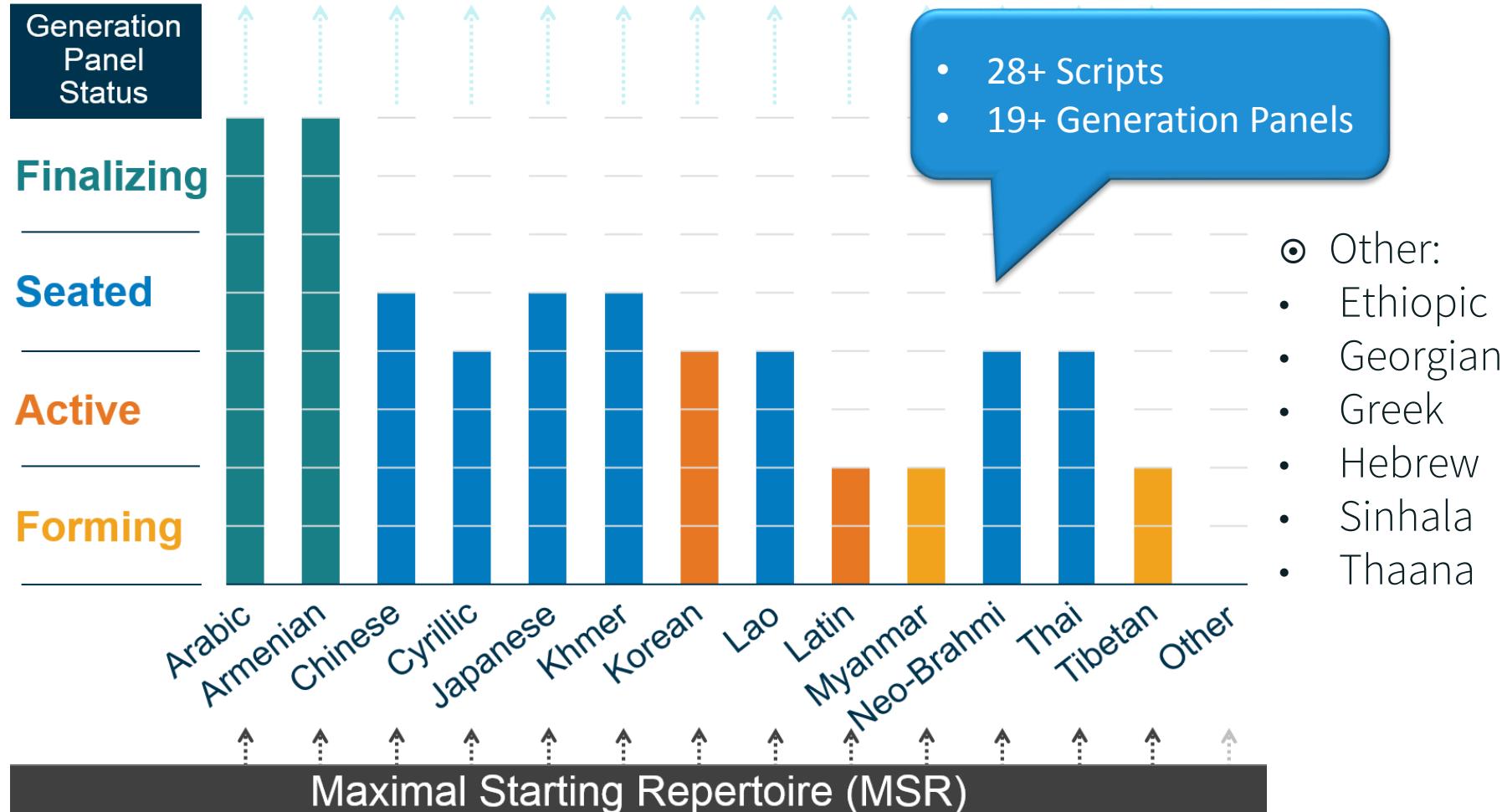
```
...
<char cp="06CC" >
    <var cp="0649" type="blocked" />
    <var cp="064A" type="allocatable" />
</char>
...
```



Status of LGR Development

ICANN 54

Label Generation Rules (LGR)



XML representation of Arabic script LGR

- <https://www.icann.org/en/system/files/files/proposed-arabic-lgr-22aug15-en.xml>
- Example of Arabic Script:

Code point definition:

```
<char cp="0647" tag="sc:Arab" ref="0 100">
  <var cp="0629" type="blocked" />
  <var cp="06BE" type="blocked" />
  <var cp="06C0" type="blocked" />
  <var cp="06C1" type="allocatable" />
  <var cp="06C2" type="blocked" />
  <var cp="06C3" type="blocked" />
  <var cp="06D5" type="blocked" />
</char>
```

WLE rule:

```
<rule name="no-mix-alef-maksura-farsi-yeh">
  <choice>
    <rule>
      <char cp="0649" />
      <any count="0+" />
      <char cp="06CC" />
    </rule>
    <rule>
      <char cp="06CC" />
      <any count="0+" />
      <char cp="0649" />
    </rule>
  </choice>
</rule>
```

Code Points

	060	061	062	063	064	065	066	067	068	069	06A	06B	06C	06D	06E	06F		075	076	077		08A	08B	08C	08D	08E	08F
0	ء	ଁ	ڏ	ڙ	-	ڦ	ڻ	ڻ	ڻ	ڻ	ڳ	ڻ	ڻ	ڻ	ڻ	ڻ	ڻ	ڦ	ڦ	ڦ		گ	ڳ	ڻ	ڻ	ڻ	ڻ
1	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
2	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
3	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
4	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
5	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
6	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
7	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
8	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
9	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
A	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
B	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
C	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
D	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
E	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ
F	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ	ڦ		ڦ	ڦ	ڦ	ڦ	ڦ	ڦ

Example of Variant Code Points

Conventional Arabic Orthography	Western (African) Orthography
ف	ب
ق	ف
ن	ل
ڭ	ڭ
ى	ڭ

Example of Variant Code Points

Unicode Code Point	Isolated Form	Initial Form	Final Form	Medial Form	Applicable Principle No.
0649	ى	ب	بى	بب	1
06CC	ى	پ	بى	پپ	1
064A	ي	پ	بى	پپ	1
06D0	ي	پ	بى	پپ	4
067B	پ	پ	بې	پپ	4
06CD	ى		بى		3
06D2	ئ		بې		2 (06CC), 5 (064A)
0626	ئ	ئب	بې	ئب	3

Example of Whole Label Evaluation Rules

S.No.	Code points cannot co-occur within a label	Notes
1	0643 and 06A9	ک and ڪ
2	0643 and 06AA	ک and ڪ
3	06CC and 0649	ڻ and ڻ
4	06C1 and 0647	ڻ and ڻ
5	0647 and 06D5	ڻ and ڻ
6	0647 and 06BE	ڻ and ڻ
7	0629 and 06C3	ڙ and ڙ
8	06D1 and 06BD	ڦ and ڦ
9	067E and 06BD	ڦ and ڦ
10	0641 and 06A2	ڦ and ڦ
11	0642 and 06A7	ڦ and ڦ

Summary of the Arabic LGR Proposal

Number of code points: 128.

Variants:

Total number of variants: 192 (this is more than the code points as the variants are directional)

Number of variants for type 'allocatable': 26.

Number of variants for type 'blocked': 166.

WLE Rules:

Number of rules defined: 17.

LGR at work !

For the IDN ccTLD of Iran: ایران

Code point sequence: 0627 06CC 0631 0627 0646

Variants Generated : 400

allocatable=4 (including 1 original)

blocked=396

Questions?!



ICANN

Sarmad Hussain

Senior Manager, IDN Program

E: sarmad.hussain@icann.org

W: <http://icann.org/>



twitter.com/icann



gplus.to/icann



facebook.com/icannorg



weibo.com/ICANNorg



linkedin.com/company/icann



flickr.com/photos/icann



youtube.com/user/icannnews



slideshare.net/icannpresentations