Japanese IDN and Search Navigation

April 14, 2010
at INET Asia 2010

Hiro Hotta
JPRS

Hello, this is me.

IDNs under .JP

* IDN : Internationalized Domain Name

Variation of JP domain names

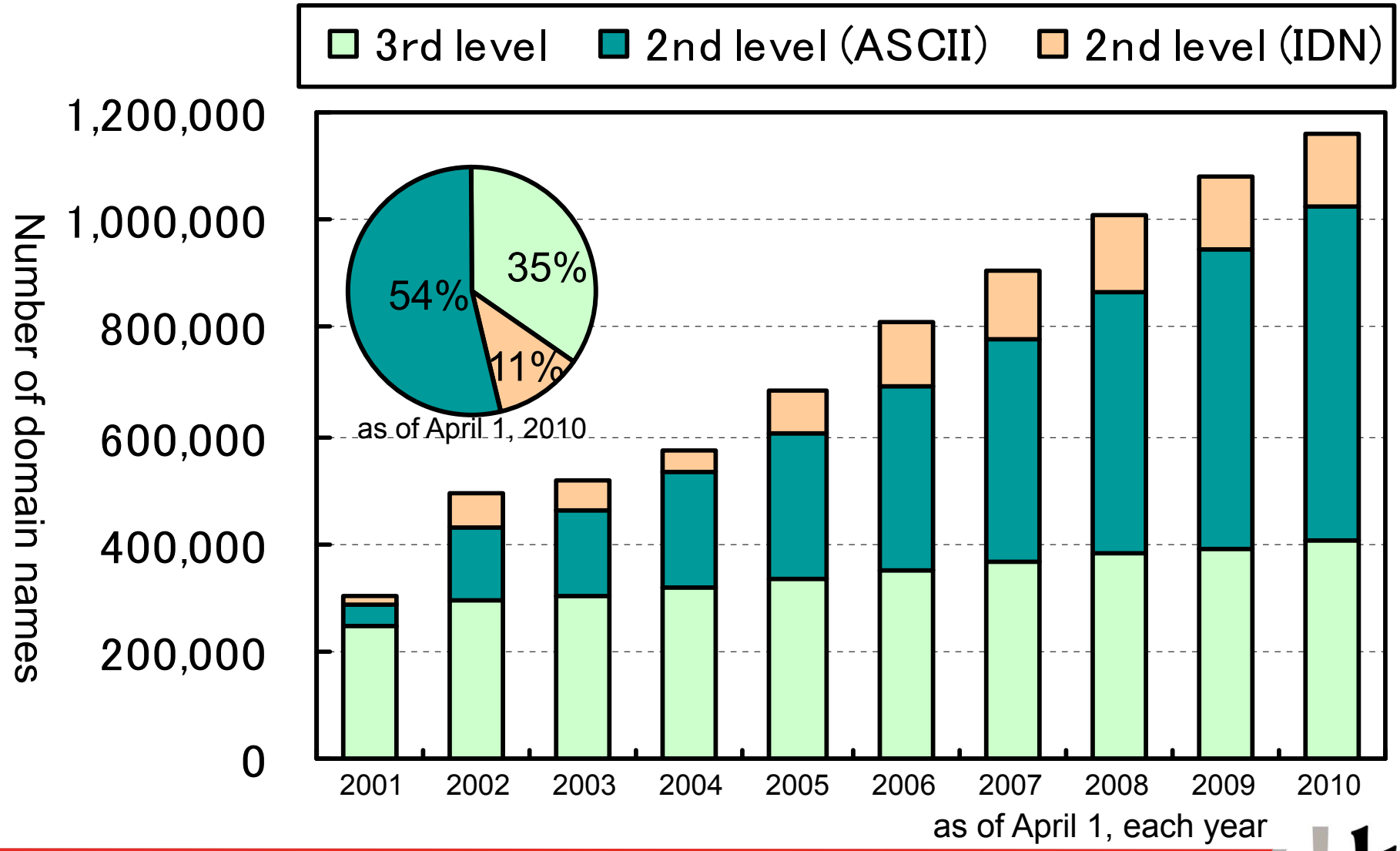
as of April 1, 2010

Third level（total: 401,587）		
****.AD.JP	JPNIC members and registrars	275
****.AC.JP	Educational institutions	3,540
****.CO.JP	Companies	336,724
****.GO.JP	Governmental organizations	791
****.OR.JP	Judicial persons other than companies	25,982
****.NE.JP	Network services	16,981
****.GR.JP	Groups	7,969
****.ED.JP	Schools	4,609
****.LG.JP	Local governments	1,867
geographic	Entities being in prefectures / core cities	2,849
Second level（total: 756,983）		
****.JP	Anyone (ASCII labels)	624,548
++++.JP	Anyone (IDN labels)	132,435

（****:ASCII, ++++:IDN）

total 1,158,570

Growth of .JP domain names



Japanese Domain Names under .JP

日本レジストリサービス.JP

- registration at the 2nd level (called General-use JP domain names) opened up in Feb. 2001
  - ASCII label
  - Japanese label
- Japanese domain names, consisting of Kanji, Hiragana, and Katakana characters as well as ASCII characters, can be registered
  - up to 15 characters
  - characters from 4 scripts can be mixed
- Japanese domain names are registered only as General-use JP domain names

Introduction of Japanese JP

- introduced on 22/Feb/2001
  - as {IDN}.jp
  - characters used in Japanese documents
    - Kanji, Hiragana, Katakana, ASCII LDH
- priority registration
  - 22/Feb/2001 - 23/Mar/2001
  - trademarks, registered names, university names, personal names in full
  - In case of competition, registrants were determined by random draw
- concurrent registration
  - 02/Apr/2001 - 23/Apr/2001
  - all applications which arrived in this period were regarded as arrived at the same time, not in the order received
  - In case of competition, registrants were determined by random draw
- first-come-first-served basis registration
  - 07/May/2001 -

Reserved Japanese Domain Names

- prefectures; large cities; prefectural capital cities
- single characters in Hiragana, Katakana, numbers written in Chinese characters, prolonged sound symbols, and others.
  - あ、イ、五、ー、...
- names of primary and secondary educational organizations (primary schools, junior high schools, etc.)
  - names ending with "小学校 primary school)," "中学校 (junior high school)" and "高等学校 (high school)."
- names of international inter-governmental organizations (such as the United Nations)
- names related to administrative, judicial, and legislative agencies
- Japanese common nouns
  - Ex) service, station, sightseeing, . . . (that may appear in yellow pages)
- names required for JP registry operations
  - ジェイピーニック、ドメイン名、日本語ドメイン名、...

Results of Priority Registration

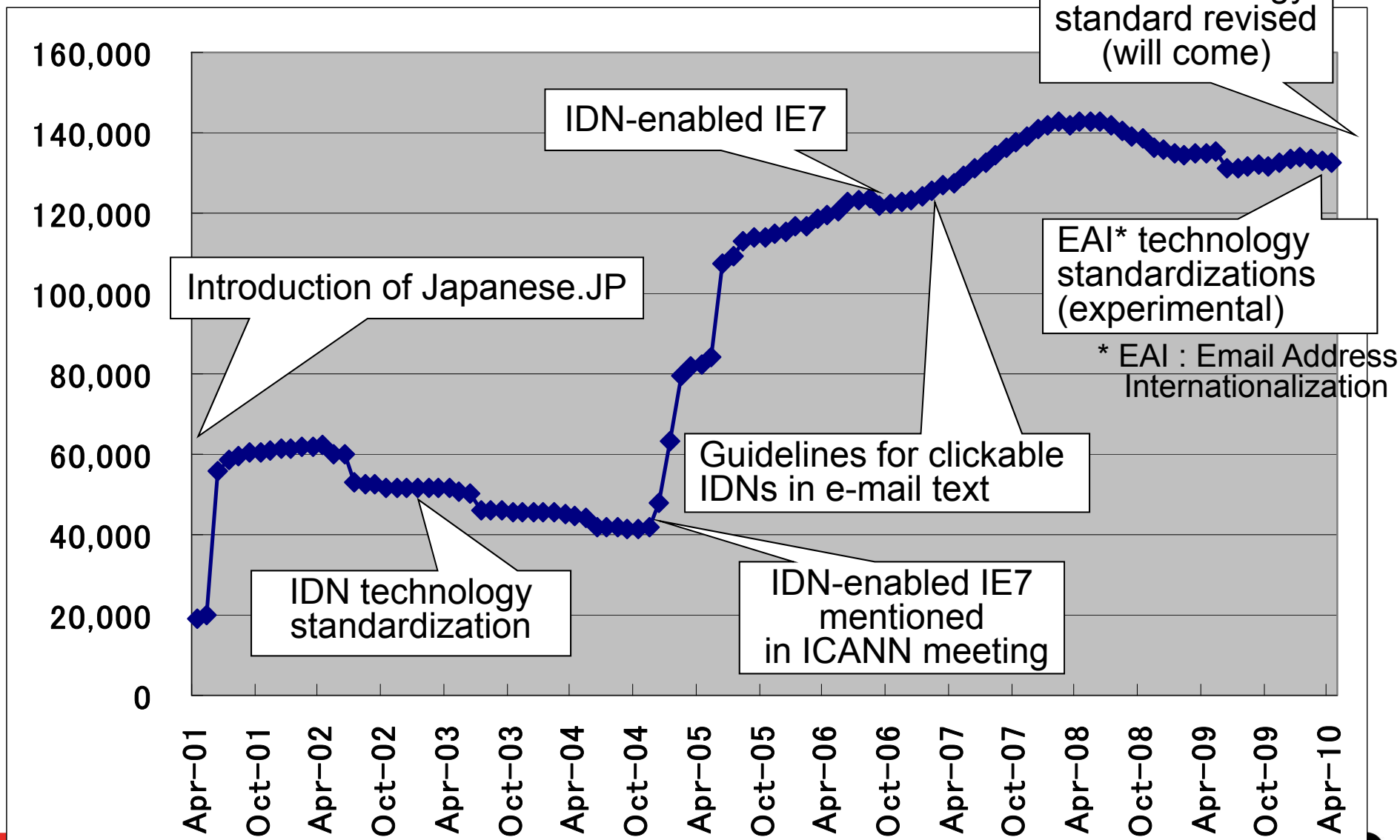
- Priority Registration Applications

Category	ASCII	Japanese	Total
Trademarks	8,300	11,900	20,100
Registered names	0	12,400	12,400
Personal names	200	600	800
academic	0	400	400
Total number of applications	8,500	25,400	33,800
Number of domain names registered	6,500	22,600	29,100

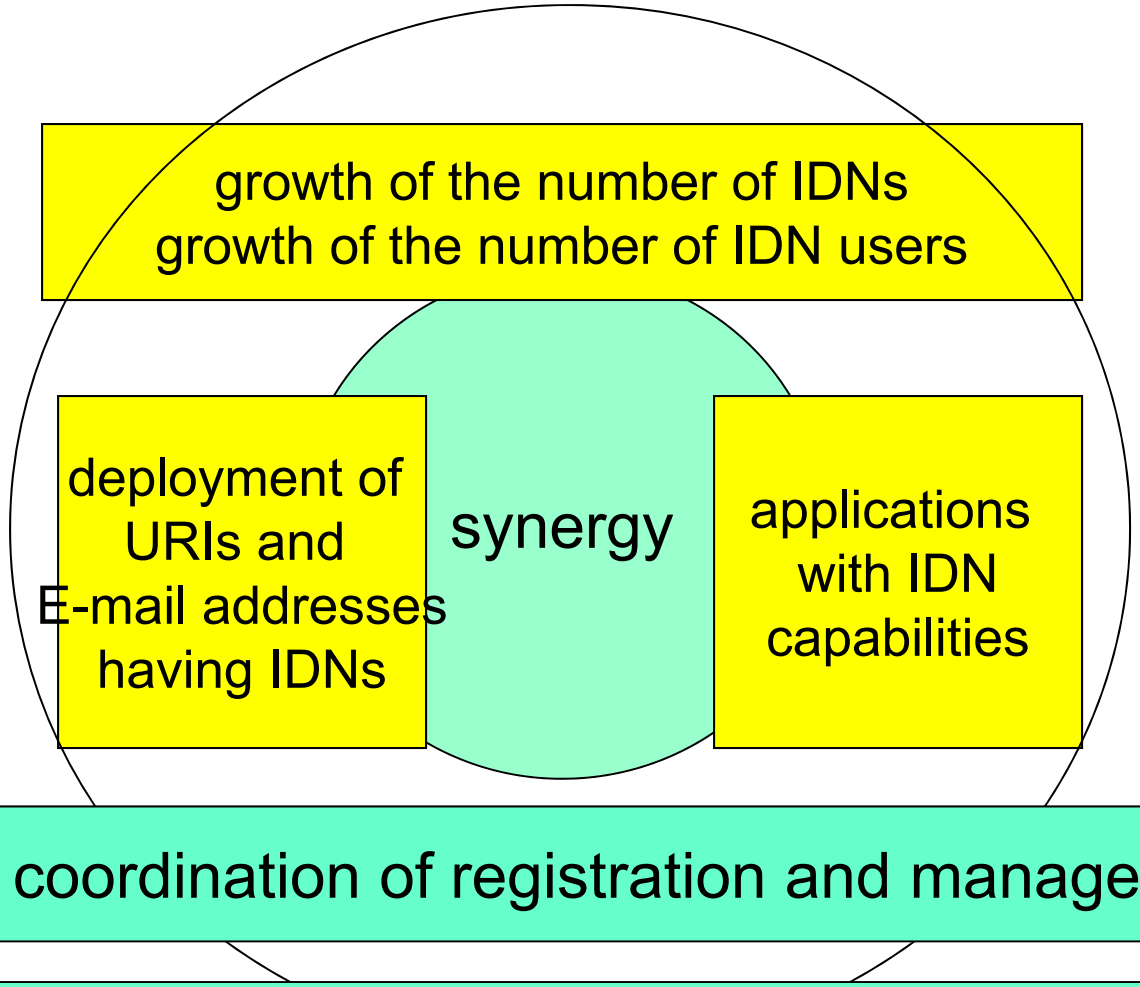
Results of Concurrent Registration

Category		ASCII	Japanese	Total
Number of applications	Multiple applications	32,500	41,700	74,200
	Single applications	22,600	23,400	46,000
	Total	55,100	65,200	120,300
Number of domain names	Multiple applications	4,600	5,200	9,800
	Single applications	22,600	23,400	46,000
	Total	27,200	28,600	55,800

Growth of {Japanese}.jp



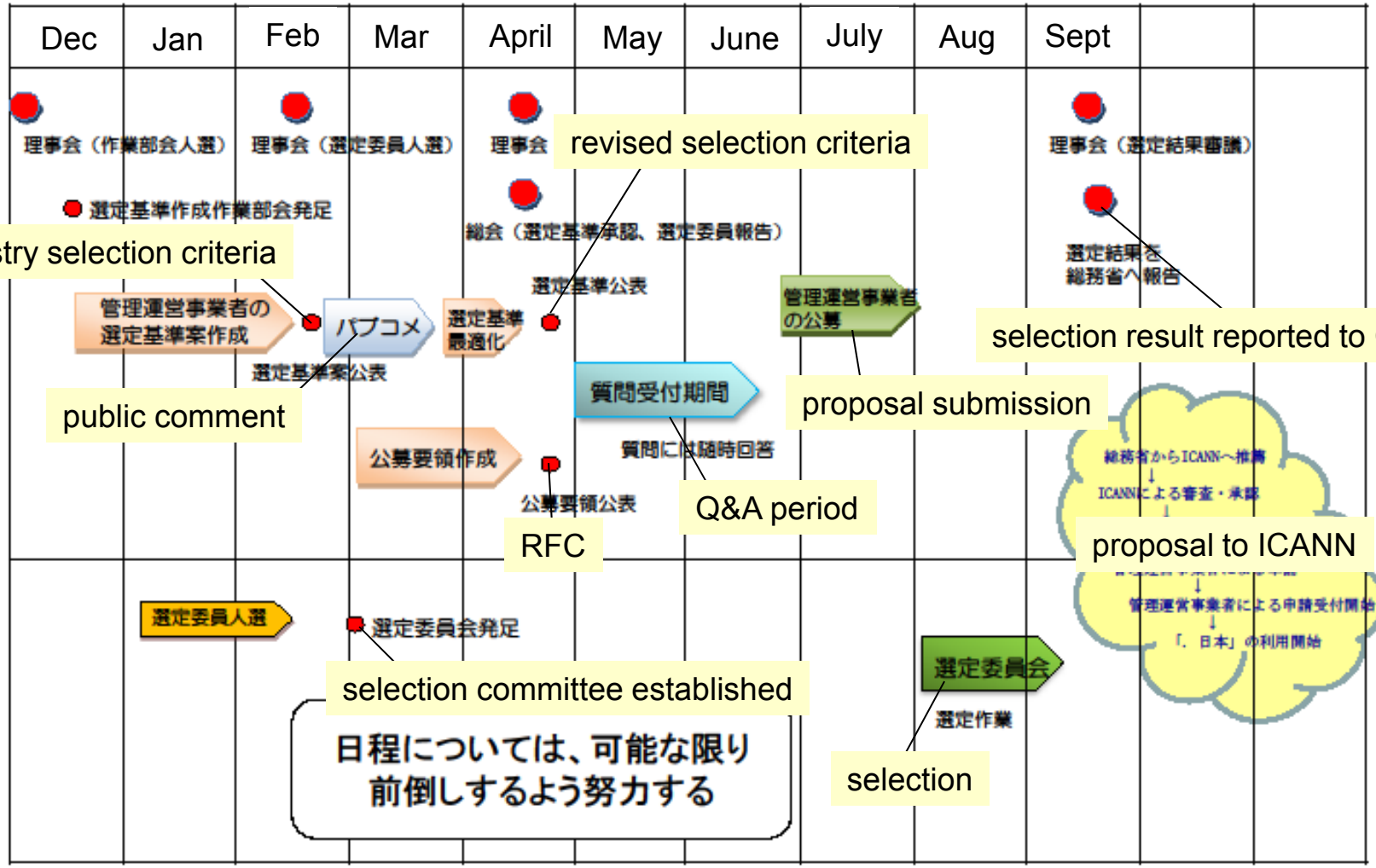
Elements for IDN Deployment



English Alphabets in Daily Life

- in Japan
  - English alphabets are everywhere
    - newspapers
    - magazines
    - company names
    - brand names
    - ...
  - 'JP' is very popular and easy to remember for most Japanese people
  - but ....
    - English 'words' are not everywhere
- in other area
  - the situation may be very different
  - IDN usage may explode when IDN TLDs become available

Schedule of IDN ccTLD in Japan



from http://jidnc.jp/wp-content/uploads/2010/02/B3-schedule.pdf

Search Navigation

Search Engines overwhelm Domain Names?

- ways to access to web sites
  - world-wide (by some statistics)
    - 2/3 accesses are via URL typing
    - 1/7 accesses are via search engines
  - in Japan (by some statistics)
    - more than 80% are via search engines

- some said
  - "Japan: URL's Are Totally Out"
    - CABEL'S BLOG LOL (Monday, March 24, 2008)
    - http://www.cabel.name/2008/03/japan-urls-are-totally-out.html

'Addressing' and 'Naming'

- communication over the Internet is becoming more and more important in our daily life
  - like bloodstream over blood vessel
- how to identify the other party in communication
  - physical position of device/host [physical semantics]
  - name of individual or organization [social semantics]
  - name of information [social semantics]
- how names in social life are presented ?
  - strings expressed in everyday languages/scripts
- how names in communication should be presented?
  - strings expressed in everyday languages/scripts
    - very important for people far from English-based life
  - names are in various media (not limited to the Internet)

Information Source vs. Information Itself

- which is more important in getting information
  - sometimes
    - information source is more important
    - if target information is reliable only when it's from a reliable source
  - usually
    - identification of target information is more important
    - if such information comes from reliable information source
- getting information
  - with addressing or naming
    - identifying the information source and then getting information

      vs.
  - without addressing or naming
    - getting the information itself without identifying the source

How to Get the Information without Addressing or Naming

- how to get to the target information
  - identifying the target by narrowing down the set of relevant information
    - using the structured categorization of data

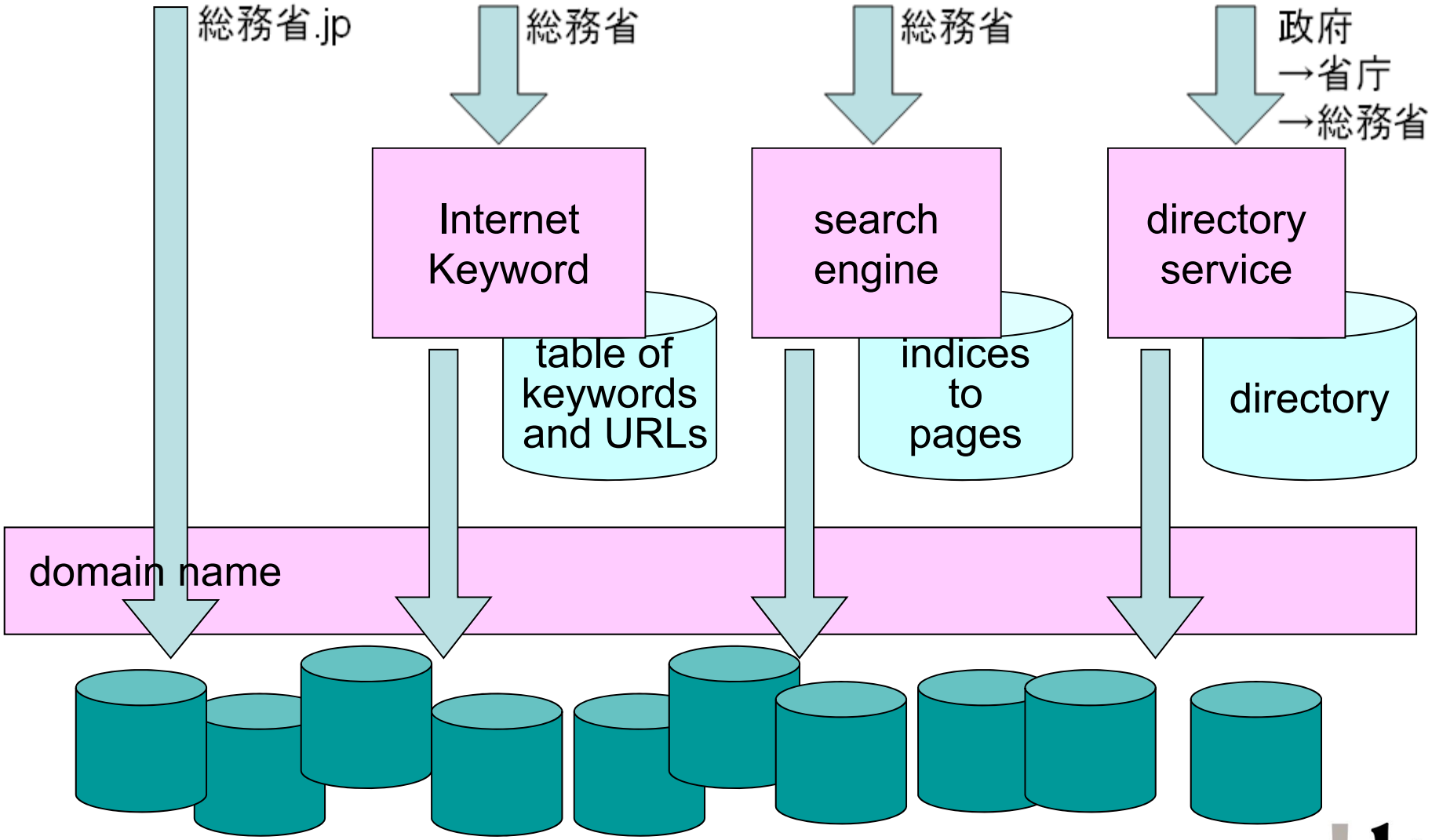
      ==> directory service
    - using the attributes of flatly scattered data

      ==> search engine
- what language/script should be used for narrowing down?
  - strings expressed in everyday languages/scripts
  - very important for people far from English-based life

Technologies to get the information using everyday languages/scripts

- by specifying information source
  1. IDN (Internationalized Domain Name)
     - Chinese Domain Name
     - Korean Domain Name
     - Japanese Domain Name
     - :
  2. internet keyword
     - CNNIC
     - Netpia
     - JWord
     - :
- by searching the information itself
  3. search engine
  4. directory service

Relations among Four Technologies



Viewpoints of Internet Users

	uniqueness in the whole world	uniqueness in a specific world	no need of pre-knowledge	easiness of input	low cost for information source	linkage from Web or e-mail text
Internet Keyword		X		X		
IDN	X				X	X
search engine			X	X	X	(X)
directory service			X		X	

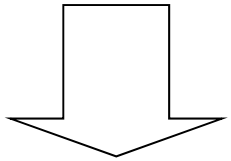
Difference in Technical Aspects

- as standardized protocol
  - IDN technologies are standard and need no standardization on top of it
  - IDNs can be used just as ASCII domain names
- as Web pages
  - search engines and directory services are provided as ordinary Web pages
- as plug-ins of Web browsers
  - some Internet keyword services are provided using plug-ins to Web browsers
- as partially-devised DNS servers
  - some Internet keyword services are provided by altered DNS in ISPs or domain name registries

Which is the Best Technology?

- right method for the right occasion
  - know the information source?
  - want information only from the source?
  - favorite way in getting to information?
  - which is the key to search? - name or category

    :



- each service coexists
  - users use the technologies(=services) case by case basis

Future Direction

using everyday languages/scripts

- pin-point access
  - access to information from already-known source
    - value raised for reliable information from reliable source
    - more secure access using technologies such as DNSSEC
- information collection
  - getting information
    - evolution of search algorithm and categorization
  - processing information
    - dynamic collection/screening/processing of multiple sourced information
- screening based on reliability
  - authentication
  - reputation

Q & A