Network Working Group
Request for Comments: 2825
Category: Informational

Internet Architecture Board (IAB)
L. Daigle, Editor

May 2000

A Tangled Web: Issues of I18N, Domain Names, and the Other Internet protocols

Status of this Memo

Copyright Notice

Abstract

The goals of the work to "internationalize" Internet protocols include providing all users of the Internet with the capability of using their own language and its standard character set to express themselves, write names, and to navigate the network. This impacts the domain names visible in e-mail addresses and so many of today's URLs used to locate information on the World Wide Web, etc. However, domain names are used by Internet protocols that are used across national boundaries. These services must interoperate worldwide, or we risk isolating components of the network from each other along locale boundaries. This type of isolation could impede not only communications among people, but opportunities of the areas involved to participate effectively in e-commerce, distance learning, and other activities at an international scale, thereby retarding economic development.

There are several proposals for internationalizing domain names, however it it is still to be determined whether any of them will ensure this interoperability and global reach while addressing visible-name representation. Some of them obviously do not. This document does not attempt to review any specific proposals, as that is the work of the Internationalized Domain Name (IDN) Working Group of the IETF, which is tasked with evaluating them in consideration of the continued global network interoperation that is the deserved expectation of all Internet users.

This document is a statement by the Internet Architecture Board. It is not a protocol specification, but an attempt to clarify the range of architectural issues that the internationalization of domain names faces.

1. A Definition of Success

The Internationalized Domain Names (IDN) Working Group is one component of the IETF's continuing comprehensive effort to internationalize language representation facilities in the protocols that support the global functioning of the Internet.

In keeping with the principles of rough consensus, running code, architectural integrity, and in the interest of ensuring the global stability of the Internet, the IAB emphasizes that all solutions proposed to the (IDN) Working Group will have to be evaluated not only on their individual technical features, but also in terms of impact on existing standards and operations of the Internet and the total effect for end-users: solutions must not cause users to become more isolated from their global neighbors even if they appear to solve a local problem. In some cases, existing protocols have limitations on allowable characters, and in other cases implementations of protocols used in the core of the Internet (beyond individual organizations) have in practice not implemented all the requisite options of the standards.

2. Technical Challenges within the Domain Name System (DNS)

In many technical respects, the IDN work is not different from any other effort to enable multiple character set representations in textual elements that were traditionally restricted to English language characters.

One aspect of the challenge is to decide how to represent the names users want in the DNS in a way that is clear, technically feasible, and ensures that a name always means the same thing. Several proposals have been suggested to address these issues.

These issues are being outlined in more detail in the IDN WG's evolving draft requirements document; further discussion is deferred to the WG and its documents.

3. Integrating with Current Realities

Nevertheless, issues faced by the IDN working group are complex and intricately intertwined with other operational components of the Internet. A key challenge in evaluating any proposed solution is the analysis of the impact on existing critical operational standards

which use fully-qualified domain names [RFC1034], or simply host names [RFC1123]. Standards-changes can be effected, but the best path forward is one that takes into account current realities and (re)deployment latencies. In the Internet's global context, it is not enough to update a few isolated systems, or even most of the systems in a country or region. Deployment must be nearly universal in order to avoid the creation of "islands" of interoperation that provide users with less access to and connection from the rest of the world.

These are not esoteric or ephemeral concerns. Some specific issues have already been identified as part of the IDN WG's efforts. These include (but are not limited to) the following examples.

3.1 Domain Names and E-mail

As indicated in the IDN WG's draft requirements document, the issue goes beyond standardization of DNS usage. Electronic mail has long been one of the most-used and most important applications of the Internet. Internet e-mail is also used as the bridge that permits the users of a variety of local and proprietary mail systems to communicate. The standard protocols that define its use (e.g., SMTP [RFC821, RFC822] and MIME [RFC2045]) do not permit the full range of characters allowed in the DNS specification. Certain characters are not allowed in e-mail address domain portions of these specifications. Some mailers, built to adhere to these specifications, are known to fail when on mail having non-ASCII domain names in its address -- by discarding, misrouting or damaging the mail. Thus, it's not possible to simply switch to internationalized domain names and expect global e-mail to continue to work until most of the servers in the world are upgraded.

3.2 Domain Names and Routing

At a lower level, the Routing Policy Specification Language (RPLS) [RFC2622] makes use of "named objects" -- and inherits object naming restrictions from older standards ([RFC822] for the same e-mail address restrictions, [RFC1034] for hostnames). This means that until routing registries and their protocols are updated, it is not possible to enter or retrieve network descriptions utilizing internationalized domain names.

3.3 Domain Names and Network Management

Also, the Simple Network Management Protocol (SNMP) uses the textual representation defined in [RFC2579]. While that specification does allow for UTF-8-based domain names, an informal survey of deployed implementations of software libraries being used to build SNMP-compliant software uncovered the fact that few (if any) implement it.

This may cause inability to enter or display correct data in network management tools, if such names are internationalized domain names.

3.4 Domain Names and Security

Critical components of Internet public key technologies (PKIX, [RFC2459], IKE [RFC2409]) rely heavily on identification of servers (hostnames, or fully qualified domain names) and users (e-mail addresses). Failure to respect the character restrictions in these protocols will impact security tools built to use them -- Transport Layer Security protocol (TLS, [RFC2246]), and IPsec [RFC2401] to name two.

Failure may not be obvious. For example, in TLS, it is common usage for a server to display a certificate containing a domain name purporting to be the domain name of the server, which the client can then match with the server name he thought he used to reach the service.

Unless comparison of domain names is properly defined, the client may either fail to match the domain name of a legitimate server, or match incorrectly the domain name of a server performing a man-in-the-middle attack. Either failure could enable attacks on systems that are now impossible or at least far more difficult.

4. Conclusion

It is therefore clear that, although there are many possible ways to assign internationalized names that are compatible with today’s DNS (or a version that is easily-deployable in the near future), not all of them are compatible with the full range of necessary networking tools. When designing a solution for internationalization of domain names, the effects on the current Internet must be carefully evaluated. Some types of solutions proposed would, if put into effect immediately, cause Internet communications to fail in ways that would be hard to detect by and pose problems for those who deploy the new services, but also for those who do not; this would have the effect of cutting those who deploy them off from effective use of the Internet.

The IDN WG has been identified as the appropriate forum for identifying and discussing solutions for such potential interoperability issues.

Experience with deployment of other protocols has indicated that it will take years before a new protocol or enhancement is used all over the Internet. So far, the IDN WG has benefited from proposed solutions from all quarters, including organizations hoping to

provide services that address visible-name representation and registration -- continuing this process with the aim of getting a single, scalable and deployable solution to this problem is the only way to ensure the continued global interoperation that is the deserved expectation of all Internet users.

5. Security Considerations

In general, assignment and use of names does not raise any special security problems. However, as noted above, some existing security mechanisms are reliant on the current specification of domain names and may not be expected to work, as is, with Internationalized domain names. Additionally, deployment of non-standard systems (e.g., in response to current pressures to address national or regional characterset representation) might result in name strings that are not globally unique, thereby opening up the possibility of "spoofing" hosts from one domain in another, as described in [RFC2826].

6. Acknowledgements

This document is the outcome of the joint effort of the members of the IAB. Additionally, valuable remarks were provided by Randy Bush, Patrik Faltstrom, Ted Hardie, Paul Hoffman, and Mark Kosters.

7. References

[RFC821] Postel, J., "Simple Mail Transfer Protocol", STD 10, RFC 821, August 1982.

[RFC822] Crocker, D., "Standard for the Format of ARPA Internet Text Messages", STD 11, RFC 822, August 1982.

[RFC1034] Mockapetris, P., "Domain Names - Concepts and Facilities", STD 13, RFC 1034, November 1987.

[RFC1123] Braden, R., "Requirements for Internet Hosts -- Application and Support", STD 3, RFC 1123, November 1989.

[RFC2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[RFC2409] Harkins, D and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

[RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

[RFC2246] Dierks, T. and C. Allen, "The TLS Protocol Version 1.0", RFC 2246, January 1999.

[RFC2459] Housley, R., Ford, W., Polk, W. and D. Solo, "Internet X.509 Public Key Infrastructure Certificate and CRL Profile", RFC 2459, January 1999.

[RFC2579] McCloghrie, K., Perkins, D., Schoenwaelder, J., Case, J. and M. Rose, "Textual Conventions for SMIv2", RFC 2579, April 1999.

[RFC2622] Alaettinoglu, C., Villamizar, C., Gerich, E., Kessens, D., Meyer, D., Bates, T., Karrenberg, D. and M. Terpstra, "Routing Policy Specification Language (RPSL)", RFC 2622, June 1999.

[RFC2826] IAB, "IAB Technical Comment on the Unique DNS Root", RFC 2826, May 2000.

8. Author's Address

Internet Architecture Board

EMail: iab@iab.org

Membership at time this document was completed:

Harald Alvestrand
Ran Atkinson
Rob Austein
Brian Carpenter
Steve Bellovin
Jon Crowcroft
Leslie Daigle
Steve Deering
Tony Hain
Geoff Huston
John Klensin
Henning Schulzrinne

9. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.