자원단위 도메인이름의 도출

A Derivation of Resource Level Domain Names

한 영 석*
Young S. Han

요 약

콘텐츠나 데이터 중심의 입력 인터페이스가 점차 확대되는 방향으로 웹응용이 진화하고 있다. 이용자는 웹사이트 보다는 사이트에 저장된 서비스와 정보에 더 관심이 많다. 콘텐츠 중심의 연산을 효과적으로 구현하기 위하여, 도메인 개념의 범위를 두 가지 방향에서 확장하였다. 자원단위의 접근에 적합한 이름구조를 제안하였고, 도메인의 의미는 다양한 함수의 영역이 되도록 허용하였다. 불필요한 최상위 확장자를 제거함으로써 자원을 위한 도메인으로서의 효율성을 높였으며 또한 호스트에 대한 독립성을 가진 도메인의 가능성을 열었다. 다양한 기능의응용이 적어도 개념적으로는 중간 매개체를 거치지 않고 도메인에 의해서 직접 호출 될 수 있다. 본 논문에서 제시한 도메인의 확장 개념은 인터넷진화에 새로운 토대가 될 수 있는 기술적 변화를 예상할 수 있다.

Abstract

The access to web applications has continuously evolved toward the adoption of content centric interfaces. Services and information items are much of user's concern than web sites. For an efficient accommodation of increasing needs of content oriented operations, the scope of domain names is extended in two directions. A naming scheme is derived that is suitable for resource level access units and the semantics of domain names is conceived as very flexible functions. First extension regards the removal of unnecessary TLD's of domain names for resource level binding. Second extension makes the resource level domain names rich in functional binding and consequently various applications can be directly triggered by the invocation of domain names without intervening medium. These extensions may lay a certain direction for the future internet evolution.

☞ keyword : DNS, Domain names, URN, HFN, Naming Architecture

1. INTRODUCTION

In most cases, people begin web services by entering a query into browser input area or search windows and this signifies the vital role of text based interfaces in online applications. Search services by portal sites maintain dominant market influences with higher effectiveness than domain names that once was an iconic status of online services as a firsthand access method.

The need for a resource oriented naming scheme is not new and motivated the works on URN (Universal Resource Name). Of many arguments supporting the URNs one is the persistence of names. Names are more reliable as means of communication as they last over the lifespan of the resources [6, 9]. URNs are one of the three components of URI (Uniform Resource Identifier) naming architecture and is designed with the expectation to be used for machines [5, 10] thereby making them less readable for humans.

Balakrishnan [1] raised many problems of current domain architecture and arguedthat there should be three levels of name resolution. One of the problems is that the internet does not provide a

* 정 회 원 : 수원대학교 정보미디어학과 교수
yshan@suwon.ac.kr
[2007/05/30 투고 - 2007/06/21 심사 - 2007/10/01 심사완료]

mechanism for directly and persistently naming data and services [1]. Their proposal includes naming independent of the hosts on which data and services reside and a similar issue has inspired researches on naming architecture such as URN [6].

A notion of resource names for human readability has not been successfully addressed. HFN (Human Friendly Name) has been a vague concept that can be used in various contexts. For example, Real Name system was proposed as a human friendly URN scheme [8]. The ideas of URN may appeal to the today's resource oriented service patterns, but URN lacked commercialmerits. Because it was mainly for machine processing the naming syntax makes it hard to read and verbalize names degrading the readability and the efficiency of communication.

It is certain that domain names were not meant for resource mappings at the time of conception of domain names, but the domain names of present days are situated in the resource centric online environments in which individual resource items and users are the units to be referenced and distributed.

By the thesis of this paper the rigidity of top level domains (TLD) is claimed to be the source of limitations to improve domain services continuously in an evolutionary system. Some commercial attempts to free the TLD space for unlimited creation of new TLD's are found in some countries of Europe and Korea [4, 12]. Their efforts are now gaining more and more consensus from the local communities. The paper tries to make a justification for the adoption of the open TLD space.

Another proposal for the empowerment of domain names is about fully functional domain names and domain names are arbitrary functions whose side effects are not confined into returning IP addresses or web pages. The reason behind this is that there exist far more various contexts in which resources are presented and associated with human needs and obtaining IP addresses or executing a small set of predefined schema with given names are not enough to fulfill what can be readily made for the betterment of users.

Current domain name services are not affected in any respect by the proposed extensions. Instead resourcedomain names are made possible in the same syntax as existing domain names withoutcausing significant level of confusion. When properly implemented the resource level domains can be a solution to some of the problems that motivated URN research and the naming architecture by Balakrishnan [1].

The service ID layer assumes a flat numbering system over resource items where each resource is assigned a number. The assigned numbers are then mapped into another flat numbers that uniquely correspond to endpoints or hosts. The endpoint numbers are then associated with IP addresses. In fact the resource domain as is proposed in this paper can be used as a replacement of the user interface layer of Balakrishnan's model [1]. The incorporation of the service ID and host ID with the resource domains is not directly dependent on each other because they are brought up by different motivations. For instance, service ID and host ID are employed to achieve some dependency of name binding from the network protocols and details. The resource domain in this paper regards the efficiency of naming design.

There are issues on the semantics of IP addresses, as well. Understanding the semantic of IP begins at the differentiation of identifiers and locators. Historically IP addresses are considered as both identifiers and locators and are expected to be unique in space and time [2]. The IP addresses,

however, are not globally unique and are semantically weak or ad hoc in current practice [2].

The semantic issues of IP addresses share concerns of semantic problems of domain names as they both care about the persistence and uniqueness. Unlike the case of IP addresses domain names can be restructured to become data and service centric and so attain the persistence property for data accesses. As to the domain semantics host independence of names is one another interest, which is not a direct point of argument of this paper. This paper tries to establish a claim that extensions of current domain strings can be safe and serve as resource domains. Another study can show that the resource level domains can be made to work independently of the hosts that contain the resources.

2. SYNTACTIC ELEVATION FOR RESOURCE LEVEL DOMAIN NAMES

Most domain names refer to host names with network identifications such as IP addresses. At the time of the invention of domain names, identifying hosts with convenience would suffice the goal of DNS (Domain Name System). That was when the most valuable resources were servers and hosts themselves. With the advent of web services, the online environment has experienced fundamental changes in all aspects.

A notable transition of emphasis from servers to individual data and service units reshaped the patterns of online interactions between users and service providers. Many people care for smaller units of information instead of sites or servers. The direction that makes resource units more central in online environments seems to be permanent with no reason to go back to the previous host centric days.

Domain names can refer to resources often smaller than servers such as data, services, and web pages that are called resource units in this paper, but the syntax of domain names seems to be a conceptualbarrier in making use of them as resource identifiers. The relaxation of the syntactic rigidity will make extended utilization of domain names.

The function of domain names takes name strings and returns network addresses pointing to hosts and resource units as is shown below.

$$F(D) = D \quad H + R$$

where F(D) is a function of domain name D, H is a set of hosts, and R is a set of resource units.

Domain names consist of domain strings delimited by '.'and each domain indicates its target resource type. Right most domains are either one of two categories representing the highest level of classification. They are known as gTLD (general Top Level Domain) and ccTLD (country code Top Level Domain). TLD's serve several goals including geopolitical differentiation and market segmentation. By the nature of domains, sub-domains constituting a domain name are categorized as follows.

$$D := (DR .)^* (DH .)+ (DC)+$$

where D is a domain name, D_R represents resource names, $D_H$corresponds to host names, and D_C regards categorical names. Syntactic descriptions can be found in [3, 7]. Sub-domains are rich in semantics but they are not typed or strongly categorized. The above definition of domain names further characterizes sub-domains according to their semantic roles.

Some domain names are meant to refer to resources with respect to the sites specified in the names. When domain names are targeted for resources, they are called resource level domain name (RDN) in this paper. RDNs, however, are notas popular as domain names (DN in short) for host references. The length of domain names may be one factor causing the less frequent uses of the resource level domain names. People may have difficulties

to memorize at least three sub domain strings constituting a RDN. For this and other reasons, a naming system called URN (Universal Resource Name) for resource naming has been long speculated and is not getting publicacceptances, yet. The URN protocol is not supported by popular browsers and DNS networks. The syntax of URN poses extra burden to the ordinary users who are familiar with domain names and the names are not designed for human reading making URNs hardly acceptable as a human interface.

It is to be shown that there exist good reasons to extend the popular syntax of domain names to meet the needs of resource centric activities that current domain system cannot handle well. Intuitively by lessening the number of sub domains we can make domain names more applicable for resource mapping because short names are easier to memorize and deliver in communications.

A simple observation leads to the notion that categorical sub-domains are unnecessary for resource level domains. Domain names (D) for resource mapping is a space by Cartesian product of resource domain, host domain, and category domain. Because of the regulations by ICANN the domain practices may not be congruent with the conceptual space given as follows.

D = R H C

The volume of resource, host, and category domains differ significantly that the size of R measures up to the order of online content units and the size of C is order of a few hundreds. The notable difference between R and C itself does not imply that category domain is not useful. Resource factor (Rfactor) is defined to be the ratio of resource volume per host.

$$Rfactor = \frac{\frac{R(t)}{H(t)}}{C(t)} = c\frac{R(t)}{H(t)}$$

where $R(t)$, $H(t)$, and $C(t)$ are functions of volume in time. $C(t)$ returning the number of category domains at time $t$may show very slow increase of the operational domains since only handful number of TLDs are administered by ICANN. With no doubt $R(t)$ on the other hand is a function in which the resource units subject to domain registration rapidly grows each day at a speed that cannot be matched with that of $C(t)$'s growth.

Rfactor indicates the resource density for hosts, and shows the relative account of resource and host volumes in time. If *Rfactor* is overly high, the chance of resource collisions because of the similarity and ambiguity of resource strings increases as well. As the *Rfactor* is maintained at a reasonably low level, $C(t)$will not produce meaningful effects, and thus categorical domains are redundant at least in resource level domain names (RDNs).

Discriminative power of a domain name depends on the discriminative quality of each sub-domain constituting the domain name. The discriminative quality of a domain name $Q(d)$ is defined as following.

$$Q(d) = \frac{Rv(r)+Rv(h)+Rv(c)}{L(r)+L(h)+L(c)} = \frac{Rv(r)+Rv(h)}{L(r)+L(h)+L(c)} \le \frac{Rv(r)+Rv(h)}{L(r)+L(h)}$$

where d D, r R, h H, c C, $Rv(r)$ is a relevance measurement of r, and $L(r)$ is a cost of r.

The cost of sub-domain string is what plays a negative role in handling the string from a human perspective. The length of a sub-domain string can be used as the cost measurement. Unfamiliarity of expressions can be another source of cost. More than one cost source can be used together in the cost function.

The relevance measurement $Rv()$ shows how positively the sub-domain string contributes to the handlingof the RDN. Semantic similarity of the sub-domain string with the target resource is a basis of the relevance measurements. With some exceptions such as '.shop' and '.name'all ccTLDs and most gTLDs are hardly relevant to the target resource even when the domain names of those TLDs contain resource sub-domains.

Since the relevance of categorical sub-domains is weak

with respect to the resources, it is safe to assume $Rv(c)$ is near zero. Consequently the use of categorical sub-domains degrades the qualityof resource level domain names. As far as domain names for resource mapping are concerned, the categorical sub-domains are not useful. The quality of RDNs is better without category domains. There is no evidence that domain names other than RDNs will be affected in meaning and syntax by the RDNs.

3. EXTENDED SEMANTICS

Resource level domain names (RDNs) in a relaxed syntax make more room for new applications. The semantics of RDNs is strongly dependent on the characteristics of resources the names represent. In the case of host level domain names the names mean host addresses. Similarly the semantics of RDNs may be a mapping into the addresses of hosts providing the resources. In the case of conventional domain names (DN), various functions of host addresses obtained by the domain names are embodied as schema called URI (Universal Resource Identifier).

The RDNs deal with the variety of content and services that can be presented on line. Instead of developing numerous schema corresponding to the applied services of RDNs we regard a RDN as a scheme itself and each RDN can be a unique function with or without any arguments. As RDN's mapping space is diversified the semantics of RDNs is a generalization of that of DNs.

Following the conventions of denotational method [11], the semantics of domain names can be described with semantic domains and functions. Assuming C, D, A, H, R, and S are semantic domains prescribing the scope of semantic objects, C is a set of schema, D is a set of domain names, A is a set of argument expressions, H is a set of hosts, R is a set of resources, and S is a set of functional services. Arguments C and A are fed to domain names that are regarded as functions returning a value in a joint space of H, R, and S. Semantic function is given as follows.

$$DS : Q \rightarrow D \times C \times A \rightarrow H + R + S$$
$$D(C, A) = A \times C \rightarrow H + R + S$$

The implication of the semantics defined as above is that domain names are function names taking arguments and they become a medium of particular function interface. The domain names mapped into host addressesare also fabricated into invocation of particular functions through extensive utilities such as URL and other schema. Semantically the target domain is a set of addresses, thus the functional invocation is made only through web pages in the case of URLs.

As for the RDNs there is no intermediary means required such as web pages at all. Inputting a RDN may shut down a machine or turn on a light. What is achieved by extending the semantics of domain names is a heightened level of the flexibility and generality of textual interface to the functional events as well as information access.

Semantic elevation is different from scheme extensions because the schema do not alter the nature of semantics of names. Each name can be a unique scheme of its own in the semantic elevation model of names. The concept of semantic extension can bring out new classes of applications in mobile services as well as online interfaces.

One another observation of RDNs is that the flexibility of TLDs and the direct association of names with resourceslead to the independence of names on the hosts on which the resources reside. This must be a meaningful result because the independence implies the persistence of names that is pursued in many previous studies [1, 6].

4. CASE STUDY

The syntactic and semantic elevation of domain nameshas been put under experimentation with the naming services by the most dominantservice providers in Korea. A similar syntactic variation is found in a European company [12] that provides a

DNS root service supporting domain names with relaxed TLDs. The European service maintains the identical semantics of conventional domain services.

The applied instance of RDN in Korea fully implements the syntacticelevation by allowing TLDs of relaxed categories andthe script is limited to Korean character set (Hangul). The Korean RDN service realizes the elevated semantics as well. The RDN is allowed to be functional, so the result or side-effect of a RDN is more than mere presentation of web pages. When the RDN input is made into the address bar of web browsers, the presence of web pages is inevitable, and thus in this regard the behavior resembles that of inputting URLs in the address bar. The RDN, however, is different because it does not necessarily involve web pages in creating some side-effects or functional outcomes. Nor is it considered as a page call like URLs are.

As an example of the RDN service in Korea, a domain name in the elevated syntax may look like the following.

an-implementation-of-domain-names.paper

The rightmost domain is a TLD "paper" that is specialized for the registration of papers. This particular TLD distinguishes itself from well known gTLDs like ".shop" and ".name" because each ".paper" relates a paper and each ".shop" is supposed to be a site. Such distinction may reside in the scope of semantics. More semantically relaxed example from conventional domain names is given by the following.

an-implementation-of-domain-names.paper/Hi! Cheers···from Mike.

A message is given as a parameter of a functional domain name that contains the contact address of a recipient of the message. In this example the domain name as a function determines the means of message transfer and other factors necessary for messaging. The input like above does not need to involve web servers or presume any response to the message sender. The message function, however, can be easily implemented with the use of web servers as well.

Unlike conventional domain names, the owner of the name may associate each name with auniquely customized function. When a RDN with functional relaxation is used in the address bar of web browser, the address bar becomes command line that leads to side effects beyond returns of web page.

toms-house.lighting/ 155

The above example illuminates the potential application in which one can control the brightness of lighting system of one's house. This particular example may presume a method of gaining access rights to the house lighting system.

Functional RDNs are polymorphic innature. A particular RDN performs different functions each of which becomes active in matching argument patterns or in application settings. The same input of RDN query may exhibit different outcomes depending on the contextof query interfaces such as web browsers, email address window, or mobile command line.

The Korean version of functional RDN adopts Hangul character set as a script language and this implements the Human-Friendliness that is often mentioned as a separate layer of domain interfaces [10] . RDN with human friendliness captures the scopes of URN and HFN.

TLDs in the Korean RDN service are grouped into three categories: commercial contents, noncommercial contents, and personal contents. Some of those RDN TLDs under service are shown in table 1.

(Table 1) Sample list of RDN TLDs.

Group	RDN TLDs
commercial use	.map .job .village .game .flower .patent .stock .free .call .sale .news .trade .tool .resort .music .best .dental .bakery .pup .eat .land .lease .building .Seoul .Busan .samsung .hyundai .LG .KT
noncommercial use	.college .school .association .meeting .church .police .diet .government .donation
personal use	.ceo .professor .chairperson .painter .writer .teacher .engineer .manager .cook .scholar .leader .issue .lesson .ucc .play .resume .humor .paper .picture .home .cartoon .why .friend .beautiful

What is notable in table 1 is that TLD can be proper nouns and even verbs that are hard to make sense in the light of conventional TLD space. The TLDs of RDN service in Korea come with a message protocol in which each RDN becomes a message recipient of the message expression following the RDN. The messages are forwarded to mobile message box or email. Certainly a message query using RDN does not need to involve an invocation of pages, but it is rather regarded as a direct function call.

There arises a question on whether a RND TLD can be allowed to support more than one of the uses shown in Table 1. Another vital issue regards the design of RDN TLDs as a consistent body of ontology comprising resource categories that as a whole achieve an efficient representation of resource space. Those issues belong to the construction of the government architecture for RND TLDs.

With these new breeds of TLDs domain practices must be quite different from what they are. Resource level names will be more common and new applications including RFID/EPC related services may be converged into names making name services pivotal in online world once again.

5. CONCLUSIONS

RDN (Resource level Domain Name) is weakly URN with the syntax of domain names, and thus it is a version of HFN for URN because RNDs are for human. It is shown that DN (Domain Name) can be more useful if an efficient representation of names for resource level subjects is accommodatedso that various TLDs may be easily afforded and still the new naming scheme does not alter the nature of current uses of domain names.

Conventional domain names include secondary domain and primary domain within their naming structure. The secondary domains are usually rich in meaning and more specific indicating particular resource types. The flexible generation of TLDs indicating resource categories implies the secondary domains replace the primary domains that often bear little relevance to the resource categories.

Through the expansion of TLD space economical representation is achieved making the resource bearing names meaningful in practice, and the claim that severe confusion may arise because of the expanded TLD space must be prejudiced by simple postulate that more TLDs cause more confusion. Confusion occurs because users have more difficulties to recollect names once they remembered because of the meaningless TLDs tailing every name such as '.com.' The fact that meaningless TLDs are the source of confusion is not addressed in this paper and is left to the reader's judgment.

Another thesis of this paper is that a domain

name can be a unique function that results in arbitrary side effects. A domain name can be considered as a command. This particular idea is useful with RDNs because RDNs refer to unit level of resources or services and will widen the scope of applications of domain names as a central text interface to numerous applications of today and future.

When RDNs are written in scripts other than ASCII, RDNs become IDNs (International Domain Names) with no extra assumptions. From the speculations of this paper, internationalized RDNs are most ideal for international users who actively participate into the dynamic evolutionary process of online world.

REFERENCES

[1] Balakrishnan, H., Lakshminarayanan, H., Karthik, and Ratnasamy, S., A Layered Naming Architecture for the Internet, Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications SIGCOMM 2004.

[2] Carpenter B., IPv4 Address Behaviour Today, RFC 2101, February, 1997.

[3] Crocker, David H, Standard for the Format of ARPA Internet Text Messages. RFC 822, August, 1982.

[4] jumhangul.com, About Contents Domain. www.jumHangul.com, 2007.

[5] Kunze, J, Functional Recommendations for Internet Resource Locators, RFC 1736, February, 1995.

[6] Moats, R, URN Syntax, RFC 2141, May, 1997.

[7] Mockapetris, P, Domain Names-Implementation and Specification, RFC 1035, November, 1987.

[8] Nicolas, Popp, The Realname system: a Human Friendly Naming Scheme. Internet Draft. draft-popp-realname-hfn-00.txt, September, 1998.

[9] Sollins, K., and Masinter, L., Functional Requirements for Uniform Resource Names, RFC 1738, December, 1994.

[10] Sollins, K., Architectural Principles of Uniform Resource Name Resolution, RFC 2276, January, 1998.

[11] Pagan, G., "Formal Specification of Programming Languages: A Panoramic Primer", Prentice Hall, New Jersey, 1981.

[12] Unified Root, About Unified Root. www.unifiedroot.com, 2007.

◐ 저 자 소 개 ◑

한 영 석(Young S. Han)

1987년 University of Central Oklahoma 컴퓨터학과 졸업(학사)
1989년 University of Oklahoma 컴퓨터학과 졸업(석사)
1995년 한국과학기술원 전산학과 졸업(박사)
1996~현재 수원대학교 정보미디어학과 교수
관심분야 : 인공지능, 정보검색, 도메인, 계산음악
E-mail : yshan@suwon..ac.kr