

Online Registries: The DNS and Beyond...

BY ESTHER DYSON

Digital Object Architecture and the Handle System

Both the DNS and the phone system offer ways to reach specific Internet locations or phone numbers, at a given time. Telephone numbers change, as do the devices they're attached to; website content changes and moves, from, say "Today's hot tip" to "Last week's highlights" to "Archive," or from "Team member" to "CEO" to "emeritus" to "who he?" Owners of websites change over time, whether through corporate reorganizations or the sale of a catchy domain name from one speculator to another (or perhaps to an actual website operator); people switch (or accumulate) e-mail addresses.

But the people and content you can reach these ways persist through time, even as site names and numbers change. How can we register persistent things in a more persistent way? One man with an answer is Bob Kahn, a veteran of the Internet. He and Vint Cerf, now chairman of ICANN and SVP of Internet Architecture and Technology at MCI, jointly developed TCP/IP for the US Defense Department's Advanced Research Projects Agency (DARPA) in the early 1970s.

In 1986, Kahn founded the nonprofit Corporation for National Research Initiatives (CNRI) in order to pursue the notion of a national information infrastructure. (Cerf had left DARPA to join MCI in 1982; he rejoined Kahn as vice president of CNRI from 1986 to 1994 before returning to MCI.) Over the last two decades CNRI has worked with government, educational and private organizations on a variety of R&D initiatives, and provides the secretariat for the IETF. It patents its R&D where it can, says Kahn, but primarily for defensive purposes – to ensure that its work is available to anyone, retains its integrity and remains unrestricted by other patents.

One of CNRI's leading initiatives is the "digital object architecture" and its implementation through the system of unique, persistent identifiers it calls handles. The Handle System is somewhat like the DNS, in that it provides a registry to find things online. But it has important differences, says Kahn: "While the DNS registers machines [with domain names], the Handle System registers digital objects. The DNS served its purpose well, but now we have something better and different, with persistence

{ continued on page 2 }



The conversation starts here.

Permission has been granted by EDventure Holdings Inc. to post this excerpt from Release 1.0. It may not be reproduced or redistributed, in whole or in part, or excerpted without the prior consent of EDventure Holdings Inc. Copyright © 2003, EDventure Holdings Inc.

Reprinted with permission from **Release 1.0**® (ISSN 1047-935X), which is published monthly except for a combined July/August issue by EDventure Holdings Inc., 104 Fifth Avenue, New York, NY 10011-6987; 1 (212) 924-8800; fax, 1 (212) 924-0240; www.edventure.com. It covers the worlds of information technology and the Internet, including wireless communications, security, business models, online services, tracking systems, identity management and other unpredictable topics. . .and the policy issues they raise.

EDITOR: Esther Dyson
(edyson@edventure.com)

PUBLISHER: Daphne Kis
(daphne@edventure.com)

MANAGING EDITOR: Christina Koukkos
(christina@edventure.com)

CONTRIBUTING WRITERS: Eric Dean
(ecdean99@hotmail.com),
Dan Gillmor (dan@gillmor.com),
John Hagel (john@johnhagel.com),
JC Herz (jc@joysticknation.com),
Clay Shirky (clay@shirky.com),
Jeff Ubois (jeff@ubois.com)

CIRCULATION MANAGER: Beckie Jankiewicz
(beckie@edventure.com)

SYSTEMS MANAGER: Geoff Clarke
(geoff@edventure.com)

CONSULTING EDITOR: Bill Kutik
(bill@kutik.com)

Copyright © 2003, EDventure Holdings Inc. All rights reserved. No material in this publication may be reproduced without written permission; however, we gladly arrange for reprints, bulk orders or site licenses. Subscriptions cost \$795 per year in the US, Canada and Mexico; \$850 overseas.

and location independence. The digital object architecture is a reconceptualization of the Internet to deal with specific information objects, instead of just flows of packets between servers. It takes transport out of the picture [although you still need it underneath].”

History that hasn't happened. . .yet?

The Handle System has been operational since 1994, and growing steadily, with somewhere upwards of 10 million handles now registered. Says Kahn, “In the late 90s, Jon Postel [who oversaw the DNS until ICANN took over] and I had come to an agreement to consider using the Handle System to replace or at least augment the DNS, but he couldn't make the decision by himself. . .” And then came ICANN, and Postel's death. The existing system persisted.

“Understanding the digital object architecture is a bit like understanding the Internet,” Kahn continues. “When they finally got it, a lot of people started to do interesting things with the Internet. We hope the same thing will happen with handles. The idea is that you can invest in content independent of platforms and of where you store or it or make it available: Anyone can find it, with certainty, not with the ambiguity of a search engine or dependency on some website.”

Although the implementation is complex, its use is not. The result is as simple as, say, a Pobox.com or alum.XXX.edu e-mail account/ID that you can keep for life (updating the information behind your unchanging e-mail ID to point to the new address). In the same way, for example, you might use a handle for your blog, so that you could keep the same handle and update its location information (rather than mailing all your friends) when you switch from one blog-hosting service to another.

Handles work seamlessly with the DNS: For human use, they can be embedded in URLs and resolved through a proxy server (such as CNRI's hdl.handle.net), so that they look just like a “regular” URL: e.g. HTTP://HDL.HANDLE.NET/4263537/1001, which currently resolves to the URL, HTTP://WWW.CNRI.RESTON.VA.US/FCC/REPLYTOCOMMENTS-FCC.PDF.

What's different from the DNS is that the Handle System's location independence, including identifiers that have no meaning (and therefore do not need to be changed, when, say, an owner changes its name or an object changes its location). That provides for persistence as well. For example, if CNRI decided to rename the document above, "RejectedPetitions," for example, or to put it elsewhere on the Net, the handle wouldn't change – and it would still resolve to the same object at its new URL (assuming proper updating of the handle record).

Thus, the Handle System obviates the DNS problem of broken URL links. That problem is a side-effect of the fact that the URL is used as an identifier and typically specifies the local file name at a location where you once were able to find that object.

Resolved but not seen

Handles generally are not meant to be seen by humans; they usually hide behind clickable buttons in applications. When a browser enabled with a handle plug-in encounters a handle, it resolves the handle via the registry to an appropriate URL or IP address for that object. Other, non-browser applications can interact directly with the Handle System using software routines that CNRI makes available for download from its site. When a handle server resolves an identifier, it returns a handle record which contains owner-specified information for the object, which could be location information, a PKI key, or a variety of different kinds of information or functions about the object such as links to digital rights management or purchase information. (It's very flexible!) It's up to the user or the user's software to extract the appropriate information from the handle record. In short, the Handle System supports application interactions rather than just queries.

The organization side

Currently, CNRI runs the handle servers, but the system is designed so that any entity can set up its own registry and run its own handle servers without continued involvement from CNRI.

Like the DNS, the Handle System is logically a single registry, but unlike the DNS it is not hierarchical. It can scale and still operate "flat." While both the DNS and the Handle System can be partitioned into subset registries that are managed separately, TLDs and second-level domains are hierarchical subsets, while the Handle System has local handle servers that interconnect horizontally.

Like a second-level domain name, a handle has two parts, before and after a slash (vs. a dot in a domain name). But again, there's a difference. If the handle is in the global registry, the entire handle is resolved. If the handle is stored only in a local registry, set up independently, the part of the handle before the slash may be used to determine which local registry can resolve the entire handle.

Because the handles are simply arbitrary numbers, about as exciting as IP addresses, there's none of the vying for "space" that afflicts the DNS. "It's a green-eyeshade job to allocate them," says Kahn. "You do have to pick a registry, or run your own," but that just means using the global handle registry or selecting from a list of registrars certified by the International DOI Foundation (IDF, see below) or another organization. . .or taking a risk on whatever registry you choose. Whether or not it's as simple as that, the IDF stands ready to provide oversight of one part of the Handle System – and if users don't like the IDF's approach, they are free to start their own registries and their own governance system.

One of the Handle System's most valuable uses could be for managing a public key infrastructure or for identity management more generally. Or it could provide a valuable degree of persistence and indirection for ENUM. "I've talked to people connected with ENUM," says Kahn, "but the implementation decisions involved political considerations, and they are reluctant to open them again at this point." The Handle System could also provide an effective underpinning for AutoID's ONS.

However, the Handle System lacks the visibility of the DNS; after all, it is so open that no one has seen much edge in marketing it until now. In addition to the IDF, a number of organizations including the Department of Defense Technical Information Center (DTIC) and the Library of Congress are running their own handle servers as internal registries. And groups in several countries, including CNNIC (which also handles China's .cn TLD), have been experimenting with it on a national basis.

But there is now a broader, outward-facing implementation, overseen by the IDF. Created by major players in the publishing industry, the IDF oversees the most visible, external registry in the Handle System, for what it calls "Digital Object Identifiers," or DOIs, a branded, trademarked version of handles. The IDF board includes the Association of American Publishers, The International Publishers Association, and various publishers including John Wiley & Sons and Springer Verlag as well as technology companies including Microsoft and Hewlett-Packard. IDF performs the sorts of functions ICANN does for the DNS, including certifying

registrars for its local registry (but unlike ICANN, it does not control the global registry). It also licenses the DOI trademark and establishes policies for use of DOIs. The DOI registrars include CrossRef, a consortium for journal publishers; The [UK] Stationery Office, a recently privatized government bureau; the Copyright Agency Ltd. in Australia; Enpia Systems, a South Korean scholarly publishing group; MEDRA, the EU-funded Multilingual European DOI Registration Agency, comprising publishers' associations in France, Germany, Italy and Spain; and the leading for-profit, commercial entry, Content Directions, Inc.

Content Directions, Inc.: Objects get respect

Of the seven registrars for DOIs certified by the IDF, Content Directions, Inc. (CDI) is the most active commercially. The story begins in 1996, when CDI's founder and CEO, David Sidman, was director of New Publishing Technologies at John Wiley & Sons and an active participant in the Association of American Publishers (AAP). The AAP was looking for a persistent registry system for content – more specifically, for a way to publish and sell high-value content electronically without losing it to piracy. The scientific community in particular was pushing science journal publishers to find a way to put their very high-value content online in a controllable way. They wanted it online to facilitate sharing. They wanted it controllable both to help ensure proper credit in the academic world and also for the usual commercial reasons of protecting intellectual property.

The new digital numbering scheme had to be more granular than, say, the ISBN (International Standard Book Number) codes used to identify books, for references to scientific articles and other such content sections rather than entire books or journals. Also, it had to be permanent. Scientific journals still refer to Einstein's 1905 paper on Special Relativity, for example. The system also had to handle very high volumes, of both registrations and references. Millions of books are published in the US each year alone; if you count chapters, articles and the like, the Lexis-Nexis database lists over 4 billion [sic] publications. Finally, it had to handle any language and alphabet in which things might be published.

The Handle System was the only system that met all the requirements. The AAP – and Sidman (then at Wiley) – worked closely with CNRI on an implementation. The first registrar was CrossRef, a non-profit consortium of about 200 scientific journal publishers. “They happily used it for cross-referencing all their articles,” says

Sidman, “but I saw the potential for a broader range of uses. The killer application for DOI is not online footnote resolution, but rather full e-commerce, eventually including digital rights management.” So in 2000 he left Wiley to start CDI, which is now a leading commercial registrar of handles.

Digital objects and digital rights

CDI has used the Handle System’s extensibility to develop what it calls “MultiLinks”: basically, a specific use of the DOI system to manage and present a variety of functions and contents that a single DOI can link to. . . including the content itself, related content (author information, for example), as well as links to a copyright administration server or to retailers’ online shopping carts. (Note that the digital rights administration server may need an identifier for the person who wants access to the content. . .another example of the ubiquitous need for identity management.)

CONTENT DIRECTIONS, INC. INFO
Headquarters: Brooklyn, NY
Founded: August 2000
Employees: 6 full time, 10 contractors “plus a halo of 40 to 50 consultants who implement systems”
Key metric: customers include McGraw Hill, Thomson Learning, Corbis, RAND and Penguin Putnam
Funding: “\$several million” from angels including former Dun & Bradstreet chairman Bob Weissmann, hedge fund wizard D.E. Shaw and Marty Kahn
URL: www.contentdirections.com

Publishers and other owners of content pay an initial setup fee for registration of the objects, including the MultiLinks, and then they pay for maintenance of the DOIs and their MultiLinks in the global directory, including, for example, constantly changing links to retailers or to related content. The pricing is negotiated; publishers typically pay in the tens or hundreds of thousands of dollars per year, says Sidman.

Some big names in the publishing business have signed up with CDI thus far: McGraw-Hill broke the ice last year by embracing DOIs for all its published content. Since then, Thomson Learning, Corbis (Bill Gates’s stock photo agency), RAND, Penguin Putnam and others have followed suit. Even the AAP itself has embraced DOI and uses CDI as the registrar for its association journal and other publications.

But Sidman has eyes on more: “The flexibility of the MultiLinks [and of the underlying Handle System] just opens whole worlds of applications.” CDI and a large, unnamed music company have initiated a pilot under which CDI will register the label’s offerings for several major artists, with three kinds of objects – for the artists, for their albums, and for individual songs. Their goal is to improve the online content experience, to make it easier and more fun for customers to hear their favorite artists and look things up about them online at the same time, find related performances. . .and to support all this, of course, without losing control over their copyrighted material.

“We’re getting a good hearing at the other music companies too,” says Sidman. “We didn’t want to venture into that world without a solid success behind us, but now we’re ready.” And perhaps the music companies are ready, too. But in the meantime, there’s Gracenote (PAGE 29), a do-it-yourself registry of music published on CD.

What the content providers really want is to make another sale when friends pass each other a song or a book or a magazine article. The vendors’ dream of user-oriented, flexible digital rights management, of the sort that will require a registry like the DOI registry, is precisely to make it easy for that secondary sale to happen. The question, of course, is whether friends are really willing to be sales agents. When Juan sends Alice a copy of something, he is showing off what he already bought, just as he might drive by her house in his new red Infiniti. How likely would he be to drive by her house in his new car if the manufacturer tried to collect a royalty from Alice when he did?

Finally, CDI is also talking to a maker of networking equipment, which is looking to use DOIs to track individual pieces of equipment, allowing various parties to see or update information about the individual item’s history – manufacture, journey through a variety of sales channels, configuration, repair and maintenance. The idea is to manage after-sales support more effectively – much the same idea as motivates many of the RFID applications covered in our June 2003 issue – and above in the section on ONS. ■ R 1.0



Content Directions, Inc.
558 9th Street
Brooklyn, NY 11215

Web site: <http://www.contentdirections.com>
Live customer examples: <http://dx.doi.org/10.1220/demo21>

David Sidman, CEO
phone: (212) 792-1847
fax: (718) 768-7542
email: dsidman@contentdirections.com