When Is a "Little in the Middle" OK? The Internet's End-to-End Principle Faces More Debate

Greg Goth

Among technopolitical idealists, the organizational structure guiding the Internet's development stands as one of the last collegial meritocracies. For almost two decades, the Internet Engineering Task Force's unofficial creed of "rough consensus and running code" has well served those who contribute to protocol development and the Internet at large.

However, the once-small group of engineers concerned mainly with ensuring that research facilities were reliably networked has become an ad hoc group, numbering up to 2,000 at its meetings in recent years, and critics complain that introducing new standards can resemble wading through knee-deep muck.

It should come as no surprise, then, that the occasional renegade technology will attempt to shoulder its way onto the global network without the blessing of the IETF, its affiliated bodies of experts, the Internet Engineering Steering Group, and the Internet Architecture Board or of the Internet Corporation for Assigned Names and Numbers, the nonprofit corporation
responsible for coordinating certain Internet technical functions, including the management of the Internet domain name system.

**TESTING THE LIMITS**

The most recent attempt at launching such a technology was SiteFinder, a search engine run by top-level registry VeriSign, which manages the .com and .net domains for ICANN. SiteFinder redirected misspelled user queries to a revenue-driven page owned by VeriSign rather than informing the user that no such domain existed, as specified in DNS protocols, and gained a large amount of unfavorable publicity in the autumn of 2003.

SiteFinder was operational for only a few weeks. Several Internet service providers complained that the technology interfered with spam filters and email servers. And prominent Internet technologists were nearly unanimous in criticizing VeriSign for apparently flouting its assigned role as a registry in launching SiteFinder and then springing it on a totally unprepared network. VeriSign suspended the search engine's operation but also hinted it might relaunch it.

A VeriSign spokesman said the company had not made a decision to restart SiteFinder, but if it did, it would give the community 60 to 90 days' notice. The company had no other comment.

However, leading Internet architects contend there is no way for VeriSign to legitimize SiteFinder at any time. Other similar technologies have waited several years for IETF and IAB approval, and VeriSign's contention that SiteFinder is a service users want has no bearing on its legitimacy or lack thereof, they say.

Beyond the specific argument of what VeriSign can do as a registry, the episode seemed to serve as the latest chapter in one of the enduring debates of Internet governance—at what layers of the network should specific technologies be allowed? And what is the proper way to introduce them?

**WHERE IS THE EDGE, ANYWAY?**

The central contention by VeriSign's critics is that SiteFinder violates not only its contract as a registry, but also the Internet's prevailing principle of end-to-end robustness as stated in RFC 1958 (www.faqs.org/rfcs/rfc1958.html). Essentially, the principle calls for stability above all in the Internet's core, while encouraging application layer innovations at the edge.
However, locating the exact spot where the middle ends and the edge begins has gotten more
difficult as more services and users come online. Compounding the theoretical quandary is the
economic potential for companies seeking to introduce a killer app somewhere between a
sender and receiver of any given online interaction.

"These days, the term 'end to end' is taken to mean 'direct,' with no intermediaries between the
end systems," says longtime IETF contributor Dave Crocker. "It is a myth, and always has been.

"What the end-to-end reference is really about is to demand as little as possible from the
infrastructure and instead [place] the 'intelligence' at the edges. This is sometimes
misunderstood to mean that everything must be dumped onto the user's computer. It doesn't.
The meaning of edge is fluid. So, for example, the user's local area network servers might
present the edge as far as the public Internet is concerned. What is essential is to have the
technical architecture designed to permit dumping as much as possible onto the end user's
system. This then gives them the opportunity to offload work onto local systems, but to do this
as a local decision."

An example of such a technology is Network Address Translation routers and firewalls, which
you can deploy from within an enterprise configuration down to the level of a residential
broadband setup connecting several devices to the network. The NAT router in the residential
setting, for example, supplies the computers in the home with internal Internet Protocol
addresses that are different than the IP address issued to the router by the user's ISP. While end-
to-end purists dislike NAT boxes because they alter core IP addresses, the sheer number of new
devices and limited address space under IPv4 has compelled many service providers and
enterprises to use them.

**BREAKING NEW GROUND**

IETF communities are debating numerous other intelligence technologies. One of the most
debated of these technologies is called Open Pluggable Edge Services, or OPES, which would
essentially let intermediary proxy servers deliver customized services while assuring either the
user or the originator of a given piece of content (and sometimes both parties) that what they
authorized actually transpired. Projected OPES services might include language transcoders or
differently configured user devices programmed to access the same body of content without
requiring the content provider to maintain separate copies for PDAs, PCs, and wireless phones.

When first proposed three years ago, OPES drew hostility and skepticism from within the IETF
and from outside organizations such as the Center for Democracy and Technology, a well-
established technology policy think tank. The concept attracted stinging comments from within the IETF calling it, among other things, "an abomination." The IAB also subjected OPES to uncommonly close scrutiny before letting it proceed.

However, those involved in steering OPES through the IETF process say the group's history proves that the traditional consensus process benefits those with vested interests as well as the public Internet at large.

"There were some times where some folks, maybe including myself, were quite frustrated, and some folks even jumped off the train," says Markus Hofmann, cochairman of the OPES working group. "However, looking back, I have to say it was worthwhile doing that. When we started discussions in the group, they were very, very broad and people had a lot of great or crazy—it depends on who you ask—ideas. And if the group had gone forward at that stage, I think it would have been very dangerous. We would have specified things that got us into those traps the IAB pointed out to us. I think the group as a whole learned to understand the major concerns and dangers that were out there."

Hofmann also notes the OPES controversy broke ground for the entire IETF.

"I also believe that what we went through was a pioneering effort—the first time there were serious discussions in the IETF on how we should handle folks or groups or efforts that want to do something in the middle of the network. I think that's also a reason it took longer. I would hope and assume that for future work or work coming up and being discussed right now that's a little bit along these lines, that the lessons we learned through OPES and the considerations the IAB put out for OPES will be very helpful for discussions on other work and will help speed that up."

John Klensin was chairman of the IAB during the early OPES discussions and says it was indeed a touchpoint for the end-to-end debate, but that the openness with which the concept was discussed had positive results.

"They came to the IETF and initiated an open discussion, into which there was a lot of input from people who wanted to make it better, and people who wanted not to let it happen, and people who had all kinds of concerns and directions. It became a very heated, but ultimately very constructive, discussion about those issues and boundaries. What VeriSign did was to put the thing up and yell 'Surprise!' A lot of people who were angered by VeriSign were more outraged that this was put up without output rather than [by] what happened. I think that's the wrong view, but that's what happened."
COMPETING PHILOSOPHIES

One of Klensin's recent ideas, on how best to deploy internationalized email addresses, has also become fodder in the end-to-end discussion.

With a standard for internationalized domain names recently agreed upon (in RFC 3490; www.faqs.org/rfcs/rfc3490.html), the engineering community has begun discussing how to best apply the same properties to the left side of the @ symbol: the user name. In early discussions, two ideas seemed most prominent: one, advanced by Klensin, would involve mailtransfer agents, notably mail servers, in forwarding mail. Another competing idea would rely more on the mail applications on client machines, or mail user agents. End-to-end purists contend Klensin's concept calls for unnecessary infrastructure changes away from the edge, notably in MTAs located between the originating and receiving MTA.

"When the design expects things to be done by intermediaries, then there is a permanent requirement for infrastructure," Crocker says. "With an end-system-oriented architecture, there is immediate utility, as soon as the first end systems, or their agents, adopt the technology. With infrastructure-oriented architecture, it usually requires large-scale adoption of the infrastructure before anyone gains utility. Notably, proposals for changing existing systems to use substantially different architectures need to pay attention to end-user incentives. John's proposal seeks to make Internet mail relaying a fundamentally different activity than it is today. The incentives for the existing half-billion users of email need to be considered rather carefully, as does their transition to the new architecture, as does our history of achievement in effecting such changes." The history of such changes, he argues, is not very successful.

Yet Klensin, who admits in his Internet draft that relay MTAs that haven't been upgraded could bounce internationalized email, says the roots of the debate are long standing, and understanding the context of probable deployment is critical.

Rather than envisioning free global email services incurring additional infrastructure costs for relay MTAs, which he calls "rotten economics," Klensin says the bulk of internationalized email, at least in early deployments, will likely be done on an intra-enterprise level and within language groups. However, he also says that we could be in for a prolonged debate.

"I wish I knew how we would resolve it, but ultimately it's a battle about competing philosophies. It may be a case where nobody is right, and independent of that, it's not clear how one resolves it."
If VeriSign attempts to relaunch SiteFinder or a similar technology, ICANN Chairman Vint Cerf says the ensuing debate might not hinge on technology at all, but rather on interpretations of exactly what a registry service is.

"I think VeriSign takes the view this particular thing, among others, is not a registry service, and thus ICANN has no jurisdiction. And I don't agree with that particular conclusion for this particular function. To directly affect the design and specifications of the DNS is clearly involving the registry function."

Crocker says even if VeriSign claims ICANN doesn't have jurisdiction on technology that alters the registry function, it can't operate unfettered.

"There is a pleasant, esoteric debate about the placement of ultimate authority for a top-level domain," Crocker says. "The debate is whether it is the US government or whether it is ICANN. Nowhere in that debate is there any serious consideration that anyone who is delegated operational responsibilities for a registry in fact owns it."

Klensin says he doesn't consider the SiteFinder debate to be a serious attack on RFC 1958, but rather on jurisdiction's gray areas.

"It's an attack on ICANN and an attack on anybody who tries to constrain [VeriSign CEO] Stratton Sclavos's behavior," Klensin says. "There were a number of companies that had clever ideas on how to get the DNS to do weird things four or five years ago that are no longer in business because they discovered their technology was never accepted by the community and was essentially quarantined. Stratton is in a little different position because he has what VeriSign has tended to assume is a stranglehold on some key resources (the .com and .net registries). That is a legal policy question about whether that stranglehold is in fact irrevocable. It's possible to speculate that VeriSign has decided it's in their business interests to test the boundaries of that group of questions and, even if they lose, they'll know the answers."