Scaling Information Infrastructure: The Case of Next-Generation IP in the Internet

Eric Monteiro

Department of Information and Computer Sciences, Norwegian University of Science and Technology, Trondheim, Norway

An information infrastructure has to scale, and hence change, as it expands. This creates a dilemma. The expansion fuels new patterns of use, which require changes, while on the other hand, the diffusion of and investments in the information infrastructure have a strong, conservative influence—the intertia of the installed base. The changes required to implement the scaling have to be in small steps. An information infrastructure is not "changed," but rather it undergoes transitions. These transitions are highly involved sociotechnical negotiations. This article is based on a case study of the efforts to change the Internet Protocol (IP) in the Internet to facilitate further growth. The revision of IP is the most serious challenge to the continued scaling of the Internet during its nearly 30 years of existence.

Keywords actor-network theory, inertia of the installed base, information infrastructure, scaling, transition strategy, Internet

Infrastructure technology lasts for many years and needs to scale in response to new patterns of use and new services. This is also true for an information infrastructure. An information infrastructure has to scale, hence change, to meet new requirements stemming from its growth. The problem, however, is how to accomplish this.

The aim of this article is to develop a deeper understanding of the problems and challenges associated with making changes in order to scale an information infrastructure. At the core of this lies a dilemma. On the one hand, the expanding information infrastructure supporting a growing population of users and new services accumulates pressure to make changes, but, on the other hand, this has to be balanced against the conservative influence of the huge, already installed base of elements of the information infrastructure. There is simply no way to accomplish abrupt changes to the whole information infrastructure requiring any kind of overall coordination (e.g., so-called flag days) because it is "too large for any kind of controlled rollout to be successful" (Hinden, 1996, p. 63). A feasible way for an information infrastructure to change is through a (more or less) smooth, near-continuous *transition* from one phase to another. To develop a firmer understanding of exactly how large changes can be made, when they are appropriate and where, and in which sequence they are to be implemented is of vital concern when establishing a National Information Infrastructure (IITA, 1995).

The empirical basis of this article is a case study of the revision of the Internet Protocol (IP) in the Internet. Its revision was a direct response to the problems of scaling the Internet: "Growth is the basic issue that created the need for a next-generation IP" (Hinden, 1996, p. 62). The IP forms the core of the Internet in the sense that most services, including the World Wide Web, e-mail, ftp, telnet, archie, and WAIS, build upon and presuppose IP. Revising IP is the most difficult and involved change ever made to the Internet during its nearly 30 years of existence. It provides an unique opportunity to study the problems of scaling large information infrastructures.

The remainder of the article is organized as follows. The second section explains and motivates why focusing on scaling in information infrastructure design is vital. The notion of a transition strategy is elaborated. The third section gives a brief outline of the social and bureaucratic organization of the Internet. It also identifies key design principles within the Internet. The function of the IP is sketched. The next three sections present the revision of

Received 5 November 1997; accepted 21 February 1998.

Address correspondence to Eric Monteiro, Department of Information and Computer Sciences, Norwegian University of Science and Technology, N-7034 Trondheim, Norway. E-mail: Eric.Monteiro@ifi.ntnu.no.

The Information Society, 14:229–245, 1998

0197-2243/98 $12.00 + .00

the IP through three phases. First the early problem formulation is retraced, roughly covering the period from the late 1980s until July 1992. Then the elaboration of the scaling problem is described together with how criteria were worked out for selecting among the alternative solutions during the 2 years from July 1992 until July 1994. The sixth section outlines the process following the decision about the protocol, a process spanning from July 1994 until today. The seventh section analyzes and discusses further three key issues regarding scaling information infrastructures that surface in the IP case: how abstract design principles always need to be appropriated to a given context, how the design of IP should be recognized as the design of infrastructure rather than an artifact, and whether the institutionalized practice of pragmatic Internet design will survive the challenges of the future. The final section contains concluding remarks.

Methodologically, the case study is a historical reconstruction based on several sources. The Internet keeps a truly extensive written record of most of its activities, an ideal source for empirical studies related to the design of the Internet. I have used the archives for IETF (see the third section for an explanation of acronyms) meetings including BOFs, working group presentations at IETF meetings (ftp://ds.internic.net/ietf/ and http://www.ieft.org), RFCs (ftp://ds.internic.net/rfc/), minutes from IPng directorate meetings (ftp://Hsdndev.harvard.edu/pub/ipng/directorate.minutes/), e-mail lists for big-internet (ftp://munnari.oz.au/big-internet/list-archive/) and several working groups (ftp://Hsdndev.harvard.edu/pub/ipng/archive/), Internet drafts (ftp://ds.internic.net/internet-drafts/), IESG membership (http://ietf.org/iesg.html#members), IAB minutes (http://info.internet.isi.edu:80/IAB), IAB membership (http://www.iab.org/iab/members.html), and information about IAB activities (http://www.iab.org/iab/connexions.html). The archives are vast, many thousands of pages of documentation in total. The big-internet e-mail list, for instance, receives on the average about 200 e-mails every month. As a supplement, I have conducted in-depth semistructured interviewing lasting about 2 hours with two persons involved in the design of the Internet (Alvestrand, 1996; Eidnes, 1996). One of them is director for one of the so-called areas within IETF and a member of the IESG.

THE NEED FOR TRANSITION STRATEGIES

The Dilemma of Scaling

Scaling an information infrastructure is neither trivial nor automatic. When expansion exceeds given limits, the information infrastructure needs to evolve, that is, change, to cater for further scaling.[1] This concern has been a cornerstone during the continued development of the Internet: "From its conception, the Internet has been, and is expected to remain, an evolving system whose participants regularly factor new requirements and technology into its design and implementation" (RFC, 1994a, p. 6). This concern for facilitating an evolving, constantly changing infrastructure is not restricted only to the Internet. It carries over to the National Information Infrastructure initiative as well. In a report by the Information Infrastructure Technology and Application working group, the highest level National Information Infrastructure technical committee, it is pointed out:

> We don't know how to approach scaling as a research question, other than to build upon experience with the Internet. However, *attention to scaling as a research theme is essential* and may help in further clarifying infrastructure needs and priorities.... It is clear that limited deployment of prototype systems will not suffice. (IITA, 1995, emphasis added)

Contributing to the problems of making changes to an information infrastructure is the fact that it is not a self-contained artifact. It is a huge, tightly interconnected yet geographically dispersed collection of both technical and nontechnical elements. Because the different elements of an information infrastructure are so tightly interconnected, it becomes increasingly difficult to make changes when it expands. The inertia of the installed base increases as the information infrastructure scales as is the case with the Internet: "The fact that the Internet is doubling in size every 11 months means that the cost of transition... (in terms of equipment and manpower) is also increasing" (IPDECIDE, 1993). But changes, and significant ones, are called for.

The scaling of an information infrastructure is accordingly caught in a dilemma. It is a process where the pressure for making changes that ensure the scaling has to be pragmatically negotiated against the conservative forces of the economical, technical, and organizational investments in the existing information infrastructure, the installed base. A feasible way to deal with this is for the information infrastructure to evolve in a small-step, near-continuous fashion with respect to the inertia of the installed base (Grindley, 1995; Hanseth et al., 1996; Neumann & Star, 1996; Star & Ruhlender, 1996). Between each of these evolutionary steps there has to be a *transition strategy*, a plan that outlines how to evolve from one stage to another. The controversies over a transition strategy are negotiations about *how* big changes can—or have to—be made, *where* to make them, and *when* and in which *sequence* to deploy them.

The aim of this article is to contribute to our understanding of scaling of information infrastructures through a study of the transition from one version of the IP to the next. It is fair to say that the Internet has never had a more serious challenge to its scaling than the revision of IP. This is because the dilemma outlined here has never been more pressing. The explosive growth of the Internet is generating a tremendous pressure for making changes, changes that are so fundamental that they need to be made at the core, that is, in IP. At the same time, these changes are likely to have repercussions on an Internet that has never been as huge and has never exhibited a stronger inertia of the installed base.

Related Research

The growing interest for information infrastructure has produced a rich variety of studies and analyses of information infrastructures. Surprisingly enough, there do not exist many studies about the Internet that try to spell out in some detail how the design process actually takes place. There exist several overviews of the historical development of the Internet, but they contain little or no evidence of how or why various design decisions came about (see, e.g., Hauben & Hauben, 1996; Lo, 1996). Abbate (1994) represents an exception. Here the underlying design visions of two competing alternatives for networking, namely, IP and the one developed within the telecommunication community (called X.25 by the CCITT[2]), are uncovered. Hanseth et al. (1996) discuss the structure of the tension between change and stability in information infrastructures with illustrations from the Internet and the Open Systems Interconnection (OSI) of the International Standardisation Organization (ISO). Hanseth (1996) analyzes the nature of the installed base through illustrations of a variety of cases, including the Internet. A highly relevant area of research regarding the Internet is to unwrap the design culture within the Internet. This, however, seems to be a completely neglected area. The few studies related to cultural aspects of the Internet focus on others than the designers—for instance, Turkle (1995) on MUD users and Baym (1995) on Usenet users.

Star et al. seek through a series of studies to develop a deeper understanding of what an infrastructure technology, as opposed to an artifact, is. Star and Ruhlender (1996) discuss the adoption and use of an information infrastructure supporting a community of researchers. Neumann and Star (1996) discuss the multiplicity of viewpoints of an information infrastructure. Bowker and Star (1994) describe how standardized classification schemes in medicine highlight some aspects of the phenomenon while downplaying others. Similarly, Jewett and Kling (1991) develop a notion of infrastructure that is to capture the many hidden resources that need to be mobilized to get an information system to actually be used.

Lehr (1992) points to the bureaucratic and procedural differences in the way standardization bodies organize their work. These are argued to play an important role for the outcome, namely, the standards. For instance, the OSI effort represents a clear-cut alternative to the evolutionary approach underlying an emphasis on transition strategies. OSI is designed monolithically from scratch, that is, with a total disregard for existing information infrastructures, the installed base. It has been fiercely criticized for exactly this (Hanseth et al., 1996; Rose, 1992c).

The literature on large technical systems is illuminating in describing how infrastructures are established but tends to bypass how to facilitate changes (Summerton, 1994). A particularly relevant contribution is Hughes (1983), which gives a historical account of the electrification of the Western world around the turn of the century. Hughes's work is important but it does not address the dilemmas of scaling explicitly. It would be illuminating to reinterpret his account by focusing on the issues of transition strategies, the role of the installed base, and gateway technologies.

Recently, there has been attention to development strategies suitable for information infrastructures (Kahin & Abbate, 1995). These strategies do not deal with scaling but address issues such as the role of government intervention and industrial consortia.

Grindley (1995) argues for the importance of the installed base of products. This emphasizes the need for products to be backward compatible; that is, that they interoperate with earlier versions of the product. In other words, this protects the installed base of earlier versions of the product. Backward compatibility plays the same role for products as transition strategies for information infrastructures (Hinden, 1996).

A possible strategy to support a scaling information infrastructure different from enhancing a smooth transition is the use of gateways. This is a potentially useful mechanism for scaling information infrastructure that plays practically no role in the Internet. It is pretty much written off within the Internet community (Eidnes, 1996; Stefferud & Pliskin, 1994). Using gateways is interesting but is not pursued in this article.[3] Consult, for instance, David and Dunn (1988) for a discussion of gateway technologies from an economic perspective. Hughes (1983) describes several employments of gateways in the history of electrification, and within the digital library initiative there is some attention to the role of gateways.

THE INTERNET

The Organization and the Standards

The Internet is one of several information infrastructures. Two other open, global ones are OSI and EDIFACT.[4] The Internet has, of course, a number of historically contingent features that distinguish it and make it difficult, if not impossible, to reproduce intentionally. Still, a number of key lessons may be learned from studying the Internet. The term "Internet" has three different meanings that sometimes need to be distinguished. It may denote (1) the collection of certified standards, (2) the physical network itself, and (3) the procedural and bureaucratic organization of the standardization process that produces, revises, and scraps the Internet standards.

The physical network relies on the implementation and deployment of relevant the Internet standards. These standards are dynamically worked out (see later discussion) and number about 200 standards today. These standards include the specification of communication protocols, which basically are organized in a hierarchy in which higher-level protocols rely on predefined functions of the lower-level ones (see Hanseth et al., 1996, for a more precise description). There are three levels, where IP forms the bottom layer, TCP the middle, and the application level is the topmost one. The application level includes the World Wide Web, e-mail, WAIS, archie, and ftp. The current version of IP is version 4 (written IPv4) and dates back to 1981. Next-generation IP, written IPng, was used to refer to the forthcoming revision of IPv4.

The Internet community consists, in principle, of everybody with access to the Internet (in the sense of meaning 2) (RFC, 1994a). Participation in the e-mail discussions, either general ones or those devoted to specific topics, is open to anyone who submits an e-mail request in the way specified (see http://www.ietf.org). The Internet community may participate in the three yearly meetings of the Internet Engineering Task Force (IETF). The IETF dynamically decides to establish and dismantle working groups devoted to specific topics. These working groups do much of the actual work of developing proposals. At the IETF meetings design issues are debated. It is furthermore possible to organize informal forums called BOFs ("birds of feather") at these meetings.

IETF nominates candidates to both the 13-member Internet Advisory Board (IAB) and the 10-member Internet Engineering Steering Group (IESG). The IETF, the IESG, and the IAB constitute the core institutions for the design of the Internet. Their members are part-time volunteers. In principle, they have distinct roles: The IETF is responsible for actually working out the proposals, the IESG for managing the standardization process, and the IAB for the overall architecture of the Internet together with the editorial management for the report series within the Internet called Requests For Comments (RFC). In practice, however, the "boundaries of the proper role for the IETF, the IESG and the IAB are somewhat fuzzy" as the current chair of the IAB admits (Carpenter, 1996). It has proven particularly difficult, as vividly illustrated in the case described next, to negotiate how the IAB should exercise its role and extend advice to the IESG and the IETF about the overall architecture of the Internet protocols. In the discussion in the seventh section of this article, the roles and interests of the IETF, the IESG, and the IAB are elaborated further.

Design Principles in the Internet

The Internet has so far scaled. It has proved remarkably flexible, adaptable, and extendable. It has undergone a substantial transformation—constantly changing, elaborating, or rejecting its constitutive standards—during its almost 30-year history of existence. The historic ability of the Internet to scale is not accidental but closely linked to deep-seated views on design. It is illuminating to try to make the underlying design principles of the Internet as explicit as possible. Few are made explicit by the community itself, so this involves an element of interpretation. Clarifying these principles will be relevant later when presenting the case of IP. An important aim of this article is to inquire in some detail into how these principles unfold in practice when confronted with the challenges of revising IP to achieve further scaling of the Internet. As will become evident later, even if the design principles are widely accepted, their "applications" in specific cases are anything but clear-cut (Suchman, 1987). To learn about how issues of scaling are handled in the Internet, it is accordingly not sufficient to refer to the programmatic design principles. They tell little. It is necessary to study how they get appropriated.

A key source for identifying design principles shared by the vast majority of the Internet community is the ones embedded in the procedural arrangements for developing the Internet standards. The standards pass through three phases that explicitly aim at interleaving the development of the standard with practical use and evaluation:

> These procedures are explicitly aimed at recognizing and adopting generally accepted practices. Thus, a candidate specification is implemented and tested for correct operation and interoperability by multiple independent parties and utilized in increasingly demanding environments, before it can be adopted as an the Internet Standard. (RFC, 1994a, p. 5)

During the first phase (a Proposed Standard), known design problems should be resolved but no practical use is required. In the second phase (a Draft Standard), at least two independent implementations need to be developed and evaluated before the standard may pass to the final phase, to be certified as a full the Internet Standard. This process is intended to ensure, beyond obviously improving the functionality, that the protocols are lean and simple, and that they are compatible with the already installed base of networks. The Internet standards are to function in a multivendor environment, that is, achieve "interoperability by multiple independent parties" (RFC, 1994a, p. 5). The current suggestion for an IPng reached the status of a Proposed Standard on 17 November 1994. It has not yet become a Draft Standard.

LATE 1980s TO JULY 1992

Framing the Problem

During the late 1980s there was a growing concern that the success of the Internet, with its accelerating adoption, diffusion, and development, was generating a problem (RFC, 1995, p. 4). No one had ever anticipated the growth rate of the Internet. The design of the Internet was not capable of handling this kind of growth for very long. The Internet is designed so that every node (e.g., a server, PC, printer, or router) has an unique address. The core of the problem was considered to be that IPv4 has a 32-bit, fixed-length address. Even though 32 bits might theoretically produce 2^{32} different identifiers, which is a very significant number, the actual number of available identifiers is dramatically lower. This is because the address space is hierarchically structured: Users, organizations, or geographical regions wanting to hook onto the Internet are assigned a set of unique identifiers (a subnetwork) of predetermined size. There are only three available sizes to choose from, the so-called class A, B, or C networks. The problem, then, is that class B networks are too popular. For a large group of users, class C is too small. Even though for many, class C would suffice, they are assigned the next size, class B, which is 256 times larger than class C. In this way, the problem of fixed-length IPv4 addresses gradually got reformed into the problem of exhausting class B networks. At the August 1990 IETF meeting it was projected that class B space would be exhausted by 1994, that is, fairly soon (RFC, 1995, p. 4). This scenario produced a profound sense of urgency. Something had to be done quickly. The easy solution of simply assigning several class C networks to users requiring somewhat more than class C size but much less than class B was immediately recognized to cause another, equally troublesome, problem. As the backbone routers in the Internet, the nodes that decide which node to forward traffic to next need to keep tables of the subnets; this explosion of the number of class C networks would dramatically increase the size of the routing tables, tables that already were growing disturbingly quickly (RFC, 1995). Even without this explosion of class C networks, the size of routing tables was causing severe problems as they grew 50% more quickly than hardware advances in memory technology.

During the early 1990s, there was a growing awareness regarding the problems associated with the continued growth of the Internet. It was also recognized that this was not an isolated problem but, rather, involved issues including assignment policies for networks, routing algorithms, and addressing schemes. There was accordingly a fairly clear conception that there was a problem complex but a poor sense of how the different problems related to each other, not to mention their relative importance or urgency. In response, the IETF in November 1991 formed a working group called Routing and Addressing (ROAD) to inquire more closely into these matters.

Appropriating the Problem

By November 1992 the ROAD group had identified two of the problems (class B exhaustion, routing table explosion) as the most pressing and IP address exhaustion as less urgent:

> Therefore, we will consider interim measures to deal with Class B address exhaustion and routing table explosion (together), and to deal with IP address exhaustion (separately). (RFC, 1992, p. 10)

The two most pressing problems required quick action. But the ROAD group recognized that for swift action to be feasible, changes had to be limited, as the total installed base cannot change quickly. This exemplifies a, if not *the*, core dilemma when extending infrastructure technologies. There is pressure for changes—some immediate, others more long-term, some well understood, others less so—that need to be pragmatically balanced against the conservative influence of the inertia of the installed base. This dilemma is intrinsic to the development of infrastructure technology and is accordingly impossible to resolve once and for all. On the one hand, one wants to explore a number of different approaches to make sure the potential problems are encountered, but on the other hand one must at some point settle for a solution in order to make further progress. It makes more sense to study specific instances of the dilemma and see how they are pragmatically negotiated in every case. A necessary prerequisite for this kind of judgment is a deep appreciation of and understanding for exactly how the inertia of the installed base operates.

In the discussions around IPng, the Internet community exhibited a rich understanding of the inertia of the installed base. It was clearly stated that the installed base was not only technical but included "systems, software, training, etc." (Crocker, 1992) and that:

> The large and growing installed base of IP systems comprises people, as well as software and machines. The proposal should describe changes in understanding and procedures that are used by the people involved in internetworking. This should include new and/or changes in concepts, terminology, and organization. (RFC, 1992, p. 19)

Furthermore, the need to order the required changes in a sequence was repeatedly stated. To be realistic, only small changes can be employed quickly. More substantial ones need to be instituted through a gradual transition.

> The [currently unknown] long-term solution will require replacement and/or extension of the Internet layer. This will be a significant trauma for vendors, operators, and for users. Therefore, it is particularly important that we either minimize the trauma involved in deploying the short- and mid-term

> solutions, or we need to assure that the short- and mid-term solutions will provide a smooth transition path for the long-term solutions. (RFC,1992, p. 11)

So much for the problem in general. How does this unfold in specific instances? Is it always clear-cut what a "small" as opposed to "large" change is, or what a "short-term" rather than "mid-" or "long-term" solution is? The controversy over CIDR and C# illustrates the problem.

CIDR Versus C#

Instead of rigid network sizes (such as class A, B, and C), the ROAD working group proposed employing CIDR ("Classless Inter-Domain Routing"). CIDR supports variable-sized networks (Eidnes, 1994). It was argued to solve many of the problems and that the disruptions to the installed base were known:

> CIDR solves the routing table explosion problem (for the current IP addressing scheme), makes the Class B exhaustion problem less important, and buys time for the crucial address exhaustion problem.... CIDR will require policy changes, protocol specification changes, implementation, and deployment of new router software, but it does not call for changes to host software. (RFC, 1992, p. 12)

At this stage, the CIDR solution to the most pressing problems was not well known, as Fuller's (1992) question to the big-internet mailing list illustrates: "But what is 'CIDR'?" Nor was it unanimous (Chiappa, 1992).

Furthermore, alternatives to CIDR existed that had several proponents. One was C#, which supported a different kind of variable-sized networks. The thrust of the argument for C#, perfectly in line with the fidelity of the installed base, was that it required fewer changes:

> I feel strongly that we should be doing C# right now. It's not new, and it's not great, but its very easy—there's nothing involved that takes any research, any developments, or any agreements not made already—just say "go" and the developers can start getting this into the production systems, and out into the field. I don't think that CIDR can be done quite that quickly. (Elz, 1992)

The discussions surrounding the different short-term solutions for the IP-related problems show broad consensus for paying respect to the installed base. The CIDR versus C# debate amounts to a judgment about exactly how much change to the installed base is feasible within a certain time frame. This judgment varied, producing disagreement and personal frustration. At the same time, the closing down of the controversy and deciding on CIDR illustrates the widespread belief that the need to move on overrides "smaller" disagreements:

> I do feel strongly that it is far more important that we decide on one, and *DO IT*, than continue to debate the merits for an extended period. Leadtimes are long, even for the simplest fix, and needs are becoming pressing. So, I want to see us *quickly* decide (agreement is probably too much to ask for :-) on *one* of the three options and *get on with it*!... I will say that I am extremely, deeply, personally, upset with the process that encouraged the creation of the C# effort, then stalled it for months while the Road group educated themselves, leaving the C# workers in the dark, etc., etc. (Chiappa, 1992)

The immediate steps, including deployment of CIDR, were to buy some time badly needed to address the big problem of IP address exhaustion. How to solve the problem was a lot less clear and the consequences were expected to be a lot bigger and cause "significant trauma for vendors, operators, and for users" (RFC, 1992, p. 11).

The Big Heat

At this stage in late 1992, there already had been proposed four solutions to the problem. One solution, called CLNP (see fifth section), was acknowledged to have a certain amount of support but was not accepted (RFC, 1992, p. 13). Unable to vouch for any one specific solution, the IESG only outlined a process of exploration that, they hoped, would lead to a solution. Central to this decision was a judgment about exactly how urgent it was to find a solution. As will become clear later in this article, this was a highly controversial issue. The IESG position was that there still was some time:

> The IESG felt that if a decision had to be made *immediately*, then "Simple CLNP" might be their choice. However, they would feel much more comfortable if more detailed information was part of the decision. The IESG felt there needed to be an open and thorough evaluation of any proposed new routing and addressing architecture. The Internet community must have a thorough understanding of the impact of changing from the current IP architecture to a new one. The community needs to be confident that we all understand which approach has the most benefits for long-term internet growth and evolution, and the least impact on the current Internet. (RFC, 1992, p. 14)

In parallel with the work of the ROAD group, and apparently poorly aligned with it, the IAB proposed its own plan for the next generation IP (IAB, 1992a). It was dubbed version 7, written IPv7. This plan of July 1992 opposed the recommendations of the ROAD group and IESG regarding the long-term problem of exhausting IPv4 address space. It produced an unprecedented heated debate during the summer of 1992. The debate focused both on the contents of IAB's solution and on the decision process producing the plan.

The crucial element of the IAB plan for IPv7 was the endorsement of one of the four available solutions, namely, CLNP. The thrust of the argument was appealing to the ideals of the Internet design: CLNP existed and people had some experience with it, so why not build upon it? Again, the controversy was not about abstract principles—they are unanimously accepted—but about how to apply the principles to a difficult situation. Hence, the IAB (1992a, p. 14) argued that:

> Delaying by a few more months in order to gather more information would be very unlikely to help us make a decision, and would encourage people to spend their time crafting arguments for why CLNP is or is not a better solution than some alternative, rather than working on the detailed specification of how CLNP can be used as the basis for IPv7.

The IAB plan for IPv7 thus made a different judgment about the available time for the Internet community to search for alternatives than the IESG IPng plan (RFC, 1992).

The decisive measures taken by the IAB, settling for a solution rather than continuing to quarrel, were praised by a number of people (Braun, 1992; Rekhter & Knopper, 1992), particularly those close to the commercial interests of the Internet. This support for swift action rather than smooth talk was mixed with a discontent about letting the fate of the Internet be left to designers with little or no interest or insight into "reality." A particularly crisp formulation of this position was submitted to the big-internet mailing list shortly after the IAB's decision (Rekhter & Knopper, 1992):

> We would like to express our strong support for the decision made by the IAB with respect to adopting CLNP as the basis for V7 of the Internet Protocol. It is high time to acknowledge that the Internet involves significant investment from the computer industry (both within the US and abroad), and provides production services to an extremely large and diverse population of users. Such an environment dictates that decisions about critical aspects of the Internet should lean toward conservatism, and should clearly steer away from proposals whose success is predicated on some future research. While other than CLNP proposals may on the surface sound tempting, the Internet community should not close its eyes to plain reality—namely that at the present moment these proposals are nothing more than just proposals; with no implementations, no experience, and in few cases strong dependencies on future research and funding. Resting the Internet future on such a foundation creates an unjustifiable risk for the whole Internet community. The decision made by the IAB clearly demonstrated that the IAB was able to go beyond parochial arguments (TCP/IP vs CLNP), and make its judgments based on practical and pragmatic considerations. Yakov Rekhter (IBM Corporation) and Mark Knopper (Merit Network)

One of the founding fathers of the Internet, Vint Cerf (1992), agreed initially with the IAB that in this case one should organize the efforts rather than fragment them:

> The CLNP specification is proposed as the starting point for the IPv7 both to lend concreteness to the ensuing discussion (I hope this does NOT result in concrete brickbats being hurled through MIME mail....!!) and to take advantage of whatever has already been learned by use of this particular packet format.

But the majority of the Internet was appalled. In the heated debate on the big-internet mailing list, a number of people spoke about "shocked disbelief," "a disastrous idea," "shocked," "dismayed," "strongly disagree," and "irresponsible." The general feeling was clear. The frustration with the decision was obviously very much influenced by the oblique way the IAB had reached its decision, thus broaching deep-seated concerns about participatory, quasi-democratic decision-making processes in the Internet. Bracketing the frustration about the decision process itself, the controversies circled around different views and interpretations of praised design principles.[5] In other words, even though there can be said to be near full consensus among the Internet community regarding concerns about continuity, installed base, transition and so forth (cf. earlier discussion), the application to specific contexts is regularly contested. The debate over IAB's IPv7 illustrates this in a striking way.

Abstract Design Principles Meet the Real World

The main reason, IAB argued, that it favored CLNP was that it was necessary for the Internet to find a solution very soon (IAB, 1992a, p. 14). CLNP is a protocol that "is already specified, and several implementations exist" so it "will avoid design of a new protocol from scratch, a process that would consume valuable time and delay testing and deployment" (IAB, 1992, p. 10).

The concern for practical experience is deep, and the CLNP solution of the IAB appealed to this. Furthermore, it paved the road for interoperability, another key principle in the Internet. Interoperability is recognized to be the end result of a process of stabilization:

> I think that relying on highly independent and distributed development and support groups (i.e., a competitive product environment) means that we need a production, multi-vendor environment operating for awhile, before interoperability can be highly stable. It simply takes time for the engineering, operations and support infrastructure to develop a common understanding of a technology. (Crocker, 1992)

While acknowledging this design principle, the IAB (1992b) in its Kobe declaration of June 1992 explained its IPv7 decision and argued that for IP an exception had to be made:

> We believe that the normal IETF process of "let a thousand (proposals) bloom," in which the "right choice" emerges gradually and naturally from a dialectic of deployment and experimentation, would in this case expose the community to too great a risk that the Internet will drown in its own explosive success before the process had run its course.

The principal difference was the pragmatic judgment of the amount of time and resources available to work out a revised IP protocol. The IESG's judgment is a head-on disagreement with the IAB's judgment. In addition, more indirect strategies for challenging the IAB were employed. One important line of argument aimed at questioning the experience with CLNP: Did it really represent a sufficiently rich source of experience?

> There does exist some pieces of an CLNP infrastructure, but not only is it much smaller than the IP infrastructure (by several orders of magnitude), but important pieces of that infrastructure are not deployed. For example the CLNP routing protocols IS-IS and IDRP are not widely deployed. ISIS (Intra-Domain routing protocol) is starting to become available from vendors, but IDRP (the ISO inter-domain routing protocol) is just coming out of ANSI. As far as I know there aren't any implementations yet. (Tsuchiya, 1992)

And more specifically, there was the question of whether the amount and types of experience were enough to ensure interoperability:

> While there certainly are some implementations and some people using [CLNP], I have no feel for the scale of the usage or—more importantly—the amount of multi-vendor interoperability that is part of production-level usage. Since we have recently been hearing repeated reference to the reliance upon and the benefits of CLNP's installed base, I'd like to hear much more concrete information about the nature of the system-level shakeout that it has already received. Discussion about deployment history, network configuration and operation experience, and assorted user-level items would also seem appropriate to flesh out the assertion that CLNP has a stable installed base upon which the Internet can rely. (Crocker, 1992)

Interoperability resulting from experience in stable environments presupposes a variety of vendors. CLNP was associated with one specific vendor, DEC, as succinctly coined by Crowcroft (1992): "IPv7 = DECNET Phase 5?" (DECNET is DEC's proprietary communication protocols.) Hence, the substance of the experience with CLNP experience was undermined, as Crocker (1992) illustrates:

> So, when we start looking at making changes to the Internet, I hope we constantly ask about the _real_ experience that is already widely available and the _real_ effort it will take to make each and every piece of every change we require.... References to the stability of CLNP leave me somewhat confused.

Gaining experience by keeping certain parts stable is a design principle (cf. earlier discussion). But some started challenging the very notion of stability. They started questioning exactly what it took for some part to be considered "stable." An important and relevant instance of this dispute was IPv4. Seemingly, IPv4 has been stable for a number of years, as the protocol was passed as an Internet Standard in 1981 without subsequent changes. But even if the isolated protocol itself has been unchanged for 15 years, have there not been a number of changes in associated and tightly coupled elements? Is it, then, reasonable to maintain that IPv4 has been stable?

> How long do we think IP has been stable? It turns out that one can give honestly different answers. The base spec hasn't changed in a very long time. On the other hand, people got different implementations of some of the options and it was not until relatively recently that things stabilized. (TCP Urgent Pointer handling was another prize. I think we got stable, interoperable implementations universally somewhere around 1988 or 89.) (Crocker, 1992)

> I still don't see how you can say things have been stable that long. There are still algorithms and systems that don't do variable length subnets. When were variable length subnets finally decided on? Are they in the previous router requirements?... So things are STILL unstable. (Tsuchiya, 1992)

This is an important argument. It will be addressed later. In effect, it states that the IP cannot be considered an isolated artifact. It is but one element of a tightly intertwined collection of artifacts. It is this collection of artifacts—this infrastructure—that is to be changed. A shift of focus from the artifact to infrastructure has far-reaching repercussions on what design is all about.

A highly contested issue was exactly which problems CLNP allegedly solved and whether these were in fact the right ones. A well-known figure in the Internet (and OSI) community, Marshall Rose, was among the ones voicing concern that it "is less clear that IPv7 will be able to achieve route-aggregation without significant administrative overhead and/or total deployment" (Rose, 1992a). After the storm of protests against IAB, combining objections against CLNP with IAB's decision process, one of the Internet's grand old men, Vint Cerf, reversed the IAB decision at the IETF in July 1992:

> Vint Cerf Monday morning basically retracted the IAB position. They are now supporting the IESG position, and he said that the IAB has learned not to try and enforce stuff from above.... Apparently Vint did a strip tease until he took off his shirt to reveal an "IP over everything" T-shirt underneath. (Medin, 1992)

The overall result of the hot summer of 1992 was that a plan to explore and evaluate proposals was worked out (RFC, 1992). By this time it was clear that "forcing premature closure of a healthy debate, in the name of 'getting things done', is *exactly* the mistake the IAB made" (Chiappa, 1992).

JULY 1992 TO JULY 1994

Let the Thousand Blossoms Bloom; or, Negotiating the Available Time

The situation by July 1992 was this. The IESG recommendation (RFC, 1992) of June 1992 calling for proposals drowned in the subsequent controversy over IAB's IPv7 plan. As the dramatic July 1992 IETF meeting led by Vint Cerf decided to reject the IAB plan, the IESG plan (RFC, 1992) was accepted and so a call for proposals for IPng was made at the meeting itself.

The problem now was to organize the effort. Central to this was, again, the issue of time: How urgent were the changes, how many different approaches should be pursued, and at which stage should one move toward a closing?

The plan by IESG formulated in June 1992 and revised a month later at the IETF meeting was shaped according to a definite sense of urgency. But it was far from panic. IESG declined to accept the problem as one merely of timing. So even though "At first the question seemed to be one of timing" (RFC, 1992, p. 14), the IESG was calm enough to hold that "additional information and criteria were needed to choose between approaches" (RFC, 1992, p. 14). Still, the suggested timetables and milestones clearly mirror a sense of urgency. The plan outlines phases of exploring alternatives, elaborating requirements for IPng and a pluralistic decision process—all to be completed within 5 months, by December 1992 (RFC, 1992, p. 15). As it turned out, this timetable was to underestimate the effort by a factor of more than four. It eventually took more than 2 years to reach the milestone the IESG originally had scheduled for late 1992.

The IESG feared fragmenting the effort too much by spending an excessive amount of time exploring many different proposals. This argument, as illustrated earlier, led Vint Cerf to initially go along with the IAB IPv7 plan that focused on CLNP. At this stage in July 1992, four proposals existed (called CNAT, IP Encaps, Nimrod, and Simple CLNP; see RFC, 1995, p. 11). This was, according to the IESG, more than sufficient, as "in fact, our biggest problem is having too many possible solutions rather than too few" (RFC, 1992, p. 2).

Following the call for proposals in July, three additional proposals were submitted during the autumn of 1992, namely, the P Internet Protocol (PIP), the Simple Internet Protocol (SIP), and TP/IX (RFC, 1995, p. 11). So by the time the IESG had planned to close down on a single solution, the Internet community was facing a wider variety of proposals than ever. Seven proposed solutions existed by December 1992.

Preparing Selection Criteria

In parallel with, and fueled by, the submission of proposals, there were efforts and discussions about the criteria for selecting proposals. As it was evident that there would be several to chose from, there was a natural need to identify a set of criteria that, ideally, would function as a vehicle for making a reasonable and open decision.

The process of working out these criteria evolved in conjunction with, rather than prior to, the elaboration of the solutions themselves. From the early sketch in 1992, the set of criteria did not stabilize into its final form as an RFC until the IPng decision was already made in July 1994 (RFC, 1994c). It accordingly makes better sense to view the process of defining a set of selection criteria as an expression of the gradual understanding and articulation of the challenges of an evolving infrastructure technology like the Internet. Neither working on the proposals themselves nor settling for selection criteria was straightforward. The efforts spanned more than 2 years, involving a significant number of people. The work and discussions took place in a variety of forms and arenas, including IETF meetings and BOFs, several e-mail lists, working groups, and teleconferencing. In tandem with the escalating debate and discussion, the institutional organization of the efforts was changed. This underscores an important but neglected aspect of developing infrastructure technology, namely, that there has to be a significant flexibility in the institutional framework not only (the more well-known challenge of) flexibility in the technology. It would carry us well beyond the scope of this article to pursue this issue in any detail, but let me indicate a few aspects. The Internet establishes and dismantles working groups dynamically. To establish a working group, the group only has to have its charter mandated by the IETF. In relation to IPng, several working groups were established (including ALE, ROAD, SIPP, TUBA, TACIT, and NGTRANS; see ftp://Hsdndev.harvard.edu/pub/ipng/archive/). As the explorative process unfolded during 1993, there was a sense of an escalating rather than diminishing degree of clarity:

> The [IPDECIDE] BOF [about criterias at the July 1993 IETF] was held in a productive atmosphere, but did not achieve what could be called a clear consensus among the assembled attendees. In fact, despite its generally productive spirit, it did more to highlight the lack of a firm direction than to create it. (RFC, 1994b, p. 2)

In response to this situation, Gross, chair of the IESG, called for the establishment of an IPng "area," an ad hoc constellation of the collection of relevant working groups with a directorate (for which he suggested the leadership of himself). At a critical time of escalating diversity, the IESG thus institutionalized a concerting of efforts. The changes in the institutional framework for the design of the Internet are elaborated in the discussion in the seventh section of this article.

Returning to the heart of the matter, the contents of solutions and the criteria, there were many variations. The rich and varied set of criteria mirrors the fact that many participants in the Internet community felt that they were at a critical point in time, and that important and consequential decision had to be made in response to a rapidly changing outside world. Hence, the natural first aim of formulating a tight and orderly set of criteria was not possible:

> This set of criteria originally began as an ordered list, with the goal of ranking the importance of various criteria. Eventually,... each criterion was presented without weighting. (RFC, 1994c, p. 2)

> The goal was to provide a yardstick against which the various proposals could be objectively measured to point up their relative strengths and weaknesses. Needless to say, this goal was far too ambitious to actually be achievable. (SELECT, 1992)

To get a feeling for the kinds of considerations, types of arguments, and levels of reflection about the problem, a small selection of issues are elaborated that relate to this article's core question of how to make changes to infrastructure technology in order to scale.

Market-Orienting the Internet

One issue concerned the role of and extent to which market forces, big organizations, and user groups should be involved. Of course, none objected to their legitimate role. But exactly how influential these concerns should be was debated. Partly, this issue had to do with the fact that historically the Internet has been dominated by individuals with a primary interest in design. There has until fairly recently not been much attention to the commercial potential of the Internet among the community itself. This is clearly changing now (Hinden, 1996). The economic and commercial repercussions of the Internet were debated, as, for instance, the IPDECIDE BOF at the July 1993 IETF confirmed that "IETF decisions now have an enormous potential economic impact on suppliers of equipment and services" (IPDECIDE, 1993). There was widespread agreement that the (near) future would witness a number of influential actors, both in terms of new markets and in terms of participants in the future development of the Internet:

> Remember, we are at the threshold of a market-driven environment.... Large-scale phone companies, international PTTs and such, for example, as they discover that there is enough money in data networking worth their attention. A major point here is that the combination of the IETF and the IAB really has to deliver here, in order to survive. (Braun, 1992)

Market forces were recognized to play an important, complementary role:

> [The] potential time frame of transition, coexistence and testing processes will be greatly influenced through the interplay of market forces within the Internet, and any IPng transition plan should recognize these motivations. (AREA, 1994a)

Still, there was broad consensus that the Internet community should take the lead. At one of the earliest broad, open hearings regarding selection criteria, the IPDECIDE BOF at the July 1993 IETF, it was forcefully stated that "'letting the market decide' (whatever that may mean) was criticized on several grounds [including the fact that the] decision was too complicated for a rational market-led solution" (IPDECIDE, 1993). Nevertheless, the increasing tension between the traditional the Internet community of designers and commercial interest surfaced. Several pointed out that the Internet designers were not in close enough contact with the "real" world. "The Internet community should not close its eyes to plain reality" (Rekhter & Knopper, 1992). This tension between users, broadly conceived, and designers did not die out. It was repeatedly voiced:

> Concerns were expressed by several service providers that the developers had little appreciation of the real-world networking complexities that transition would force people to cope with. (IPDECIDE, 1993)

> More bluntly, I find it rather peculiar to be an end user saying: we end users desperately need [a certain feature] and then sitting back and hearing non-end-users saying, "No you don't." (Fleichman, 1993)

Stick or Carrot?

Still, the core problem with IPng concerned how large changes could (or ought to) be made, where, how, and when to make them—in other words, the transition strategy broadly conceived.

On the one hand, there were good reasons for making substantial changes to IPv4. A number of new services and patterns of use were expected, including real-time, multimedia, asynchronous transfer mode, routing policy, and mobile computing. On the other hand, there was the pressure for playing it reasonable safe by focusing only on what was absolutely required, namely, solving the addressing space and routing problems. This was recognized as a dilemma:

> There was no consensus about how to resolve this dilemma, since both smooth transition and [new services like for instance] multimedia support are musts. (IPDECIDE, 1993)

It was pointed out earlier that balancing the pressure for changes against the need to protect the installed base is an intrinsic dilemma of infrastructure technology. In the case of IPng, this was amplified by the fact that the core requirements for IPng, namely, solving the routing and address space problems, were invisible to most users. They were taken for granted. Hence, there were few incentives for users to change. Why would anyone bother to change to something with little perceived added value? In the final version of the selection criteria, addressing this dilemma is used to guide all other requirements:

> We have had two guiding principles. First, IPng must offer an internetwork service akin to that of IPv4, but improved to handle the well-known and widely understood problems of scaling the Internet architecture to more end-points and an ever increasing range of bandwidths. Second, it must be desirable for users and network managers to upgrade their equipment to support IPng. At a minimum, this second point implies that there must be a straightforward way to transition systems from IPv4 to IPng. But it also strongly suggests that IPng should offer features that IPv4 does not; new features provide a motivation to deploy IPng more quickly. (RFC, 1994c, pp. 3–4)

It was argued that the incentives should be easily recognizable for important user groups. Hence, it was pointed out that network operators were so vital that they should be offered tempting features such as controlling "load-shedding and balancing, switching to backup routers" (NGREQS, 1994). Similarly, the deep-seated aversion for application platform interfaces, that is, tailor-made interfaces for specific platforms, was questioned. Despite the fact that "the IETF does not 'do' [application platform interfaces]" (RFC, 1995, p. 39), the IESG finally recommended that an exception should be made in the case of IPng. This was because it met the pressing need for tangible incentives for a transition to IPng (RFC, 1995, p. 5).

The Internet Is an Infrastructure, Not an Artifact

A large number of requirements were suggested and debated. They included topological flexibility, mobile communication, security, architectural simplicity, unique identifiers, risk assessment, network management, variable-length addresses, and performance (RFC, 1994c). Besides addressing perceived and anticipated requirements, the requirements might have repercussions on the whole infrastructure, not only IPng.

It was repeatedly pointed out that IPng was not only about revising one self-contained element of the Internet. It was about changing a core element of an infrastructure with tight and oblique coupling to a host of other elements in the infrastructure:

> Matt Mathis pointed out that different proposals may differ in how the pain of deployment is allocated among the levels of the networking food chain (backbones, midlevels, campus nets, end users). (SELECT, 1992)

> I would strongly urge the customer/user community to think about costs, training efforts, and operational impacts of the various proposals and PLEASE contribute those thoughts to the technical process. (Crocker, 1992)

This well-developed sense of trying to grasp how one component, here IPng, relates to the surrounding components of the information infrastructure is a principal reason for the Internet's success until now.

New features are included to tempt key users to change. But the drive toward conservatism is linked to one of the most important design principles of the Internet, namely, to protect the installed base. It is of overriding importance:

> The transition and interoperation aspects of any IPng is *the* key first element, without which any other significant advantage won't be able to be integrated into the user's network environment. (e-mail from B. Fink to SIPP mailing list, cited by Hinden, 1996)

This appeal for conservatism is repeated ad nauseam. The very first sentence of RFC (1996), describing the transition mechanisms of IPv6, reads: "The key to a successful IPv6 transition is compatibility with the large installed base of IPv4 hosts and routers" (p. 1). The pressure for holding back and declining features that might disturb the installed base is tremendous.

"Applying" the Principles

A rich and varied set of proposed requirements was worked out. Still, it is not reasonable to hold that the decision was made by simply "applying" the abstract selection criteria to the different proposals for IPng. Despite the fact that the resulting requirements (RFC, 1994c) with 17 criteria were "presented without weighting" (RFC, 1994, p. 3), a few themes were of overriding importance (IPDECIDE, 1993). At this stage, draft requirements had been suggested for more than a year and seven candidates existed, but the requirements were "too general to support a defensible choice on the grounds of technical adequacy" and "had so far not gelled enough to eliminate any candidate" (RFC, 1994c). The concern for sharper criteria prevailed. It was repeated as late as in March 1994, only 2 months before the decision was made:

> One important improvement that seemed to have great support from the community was that the requirements should be strengthened and made firmer—fewer "should allows" and the like and more "musts." (AREA, 1994b)

The core concern focused on making transition from IPv4 to IPv6 as smooth, simple, and uncostly as possible. A few carrots were considered crucial as incentives for a transition, primarily security:

> What is the trade-off between time (getting the protocol done quickly) versus getting autoconfiguration and security into the protocol? Autoconfiguration and security are important carrots to get people to use IPng. The trade-off between making IPng better than IP (so people will use it) versus keeping IPv4 to be as good as it can be. (NGDIR, 1994)

Other requirements were to a large extent subordinate or related to these. For instance, autoconfiguration, that is, "plug and play" functionality, may be viewed as an incentive for transition.

The collection of proposed IPng solutions had evolved, joined forces, or died. As explained earlier, there was tight interplay between the development of the solutions and the criteria. The real closing down on one solution took place during May–July 1994. In this period, there were extensive e-mail discussions, but more importantly, the IPng Directorate organized a 2-day retreat on 19–20 May 1994 at BigTen with the aim of evaluating and reworking the proposals (Knopper, 1994). Through this and the subsequent IETF in July 1994, an IPng solution was decided upon.

Showdown

By the spring of 1994, three candidates for IPng existed, namely, CATNIP (evolving from TP/IX), SIPP (an alliance between IPAE, SIP, and PIP), and TUBA (evolving from Simple CLNP). A fourth proposal, Nimrod, was more or less immediately rejected for being too unfinished and too much of a research project.

CATNIP was "to provide common ground between the Internet, OSI, and the Novell protocols" (RFC, 1995, p. 12). The basic idea of CATNIP for ensuring this was to have the Internet, OSI, and Novell transport layer protocols (e.g., TCP, TP4 and SPX) run on to of any of the network layer protocols (IPv4, CLNP, IPX- or CATNIP). The addressing scheme was borrowed from OSI.

A primary objection against CATNIP that surfaced during the BigTen retreat was that it was not completely specified (Knopper, 1994; RFC, 1995, pp. 14–15). Beyond the obvious problems with evaluating an incomplete proposal, this illustrates a more general point made earlier and illustrated by Alvestrand (1996), area director within IETF: "The way to get something done in the Internet is to work and write down the proposal." Despite appreciation for the "innovative" solution, there was skepticism toward the "complexity of trying to be the union of a number of existing network protocols" (RFC, 1995, p. 15).

The TUBA solution was explicitly conservative. Its principal aim was to "minimize the risk associated with the migration to a new IP address space" (RFC, 1995, p. 13). This would mean "only replacing IP with CLNP" (RFC, 1995, p. 13) and letting "existing Internet transport and application protocols continue to operate unchanged, except for the replacement of 32-bit IP[v4] addresses with larger addresses" (RFC, 1995, p. 13). CLNP is, as outlined earlier, OSI's already existing network layer protocol. Hence, the core idea is simply to encapsulate, that is, wrap up, TCP in CLNP packets. The evaluation of TUBA acknowledged the benefits of a solution making use of the "significant deployment of CLNP-routers throughout the Internet" (RFC, 1995, p. 16), that is, a solution paying respect to an installed base. Similar to the arguments outlined in the fourth section of this article regarding the IAB's IPv7 plan to build IPng on CLNP, "There was considerably less agreement that there was significant deployment of CLNP-capable hosts or actual networks running CLNP" (RFC, 1995, p. 16). The worries—"including prejudice in a few cases" (RFC, 1995, p. 16)—about the prospects of losing control of the Internet by aligning IPng with an OSI protocol were deep-seated.

SIPP was to be "an evolutionary step from IPv4 ... not ... a radical step" (RFC, 1995, p. 12). SIPP doubles the address size of IP from 32 to 64 bits to support more levels of addressing hierarchy and a much greater number of addressable nodes. SIPP does not, in the same way as CATNIP or TUBA, relate to non-Internet protocols.

The reviews of SIPP were favorable. SIPP was praised for its "aesthetically beautiful protocol well tailored to compactly satisfy today's known network requirement" (RFC, 1995, p. 15). It was furthermore pointed out that the SIPP working group had been the most dynamic one in the previous year, producing close to a complete specification. Still, it was definitely not a satisfactory solution. In particular, the transition plans (based on the encapsulation suggestion originally in IPAE) were viewed as "fatally flawed" (Knopper, 1994). A number of reviewers also felt that the routing problems were not really addressed, partly because there was no way to deal with topological information and aggregation of information about areas of the network.

In sum, there were significant problems with all three proposals. Because CATNIP was so incomplete, the real choice was between TUBA and SIPP. Following the BigTen evaluation retreat, Deering and Francis (1994), co-chairs of the SIPP working group, summarized the BigTen retreat to the SIPP e-mail list and proposed to build upon suggestions that came out of it. Particularly important, they suggested to "change address size from 8 bytes [= 64 bits, the original SIPP proposal] to 16 bytes [= 128 bits] (fixed-length)" (Deering & Francis, 1994). This increase in address length would buy flexibility to find better solutions for autoconfiguration, more akin to the TUBA solution. These suggestions were accepted by the SIPP working group, which submitted the revised SIPP (version 128 bits) to the IPng Directorate together with a new but incomplete transition plan inspired by TUBA. This was accepted in July 1994 as the solution for IPng, finally ready to be put on the ordinary standards track of the Internet.

JULY 1994 TO TODAY

Finished at Last—Or Are We?

By the summer of 1994, a recommended candidate for IPng was found. It was called IPv6. It was put on the standard track (cf. earlier description) and was made a Proposed Standard in November 1994. One could accordingly be tempted to think that it was all over, that one had found a way that secured the future of the Internet. This, however, is not quite the case, not even today. There is a fairly well-founded doubt about "whether IPv6 is in fact the right solution to the right problem" (Eidnes, 1996). There are two reasons for this, both to be elaborated later:

- There was—and still is—a considerable degree of uncertainty about how to conduct full-scale testing.
- Even if the IPng protocol itself was completed, a number of tightly related issues were still unresolved, most importantly, a transition strategy.

Full-Scale Testing

A core element of the Internet design principles, which could be said to be the realization of the Internet pragmatism, is the emphasis on practical experience and testing of any solutions (RFC, 1994a). Although this principle is universally accepted within the Internet community, the point is that as the installed base of the Internet expands, so do the difficulties of actually accomplishing large-scale, realistic testing. So again, how should the principle of realistic testing be implemented for IPng? This worry was voiced fairly early on:

> It is unclear how to prove that any proposal truly scales to a billion nodes.... Concern was expressed about the feasibility of conducting reasonably-sized trials of more than one selected protocol and of the confusing signals this would send the market. (IPDECIDE, 1993)

The problem of insufficient testing is important because it undermines the possibility of establishing interoperability (IPDECIDE, 1993): "It is also difficult to estimate the time taken to implement, test and then deploy any chosen solution: it was not clear who was best placed to do this."

Current deployment of IPv6 is very slow. Implementations of IPv6 segments, even on an experimental basis, hardly exist (Eidnes, 1996). Even though the phases that a standard undergoes before becoming a full Internet Standard may take as little as 10 months, a more realistic projection for IPv6 is 5 years (Alvestrand, 1996). The upgrading of IPv6 to a Draft Standard requires testing well beyond what has so far been conducted. As the Internet expands, full-scale testing becomes more cumbersome. Some within the IETF see an increasingly important role for noncommercial actors, for instance, research networks, to function as early test beds for future the Internet Standards (Alvestrand, 1996). The U.S. Naval Research Laboratory had implemented an experimental IPv6 segment by 1 June 1996 as part of their Internet working research. The Norwegian research network, which traditionally has been fairly up front, expects to start deployment of IPv6 during 1997.

Unresolved Issues

At the time when the IPng protocol was accepted on the standards track, several crucial issues were still not completed. At the November 1994 IETF immediately following the IPng decision, it was estimated that 10–20 specifications were required (AREA, 1994b). Most importantly, a transition strategy was not in place. This illustrates the point made earlier, namely, that the actual design decisions are not derived in any straightforward sense from abstract principles. Besides a transition strategy, the security mechanisms related to key management had not been—and, indeed, still are not—completed.

A core requirement for IPng was to have a clear transition strategy (RFC, 1995). The SIPP (version 128 bits) was accepted as IPng without formally having produced a clear transition strategy because the concerns for facilitating a smooth transition were interwoven with the whole process, as outlined earlier. There was a feeling that it would be feasible to work out the details of the transition mechanisms based on the IPng protocol. It was accordingly decided by the IPng Directorate just prior to the BigTen retreat to separate transition from the protocol.

In response to the lack of a complete transition strategy, informal BOFs (NGTRANS and TACIT) were held at the November 1994 IETF. TACIT was a working group formed during the spring of 1994; NGTRANS was established as a working group shortly after the November 1994 IETF. Both TACIT and NGTRANS were to address the issue of a transition strategy, but with slightly different focus. NGTRANS was to develop and specify the actual, short-term transition mechanisms, leaving TACIT to deal with deployment plans and operational policies (NGTRANS, 1994). The available time for a transition was to be "complete before IPv4 routing and addressing break down" (Hinden, 1996, p. 62). As a result of the deployment of CIDR (cf. earlier discussion), it was now estimated that "IPv4 addresses would be depleted around 2008, give or take three years" (AREA, 1994b).

From drafts sketched prior to the establishment of NGTRANS and TACIT, the work with the transition strategy was completed to the stage of an RFC only by April 1996 (RFC, 1996).

The transition mechanisms evolved gradually. It was recognized early on that a cornerstone of the transition strategy was a "dual-stack" node, that is, host or router. A dual-stack node implements both IPv4 and IPv6 and thus functions as a gateway between IPv4 and IPv6 segments. Dual-stack nodes have the capability to send and receive both IPv4 and IPv6 packets. They enforce no special ordering on the sequence of nodes to be upgraded to IPv6, as dual-stack nodes "can directly interoperate with IPv4 nodes using IPv4 packets, and also directly interoperate with IPv6 nodes using IPv6 packets" (RFC, 1996, p. 4).

Progress was also made on closely related elements of an IPv6 infrastructure. The bulk of the IPv4 routing algorithms were reported to be working also for IPv6 routers, a piece of pleasant news in November 1994 (AREA, 1994a, p. 4).

The additional key transition mechanism, besides dual-stack nodes, was IPv6 over IPv4 "tunneling." This is the encapsulation, or wrapping up, of an IPv6 packet within an IPv4 header in order to carry them across IPv4 segments of the infrastructure. A key element to facilitate this is to assign IPv6 addresses that are compatible to IPv4 addresses in a special way. (The IPv4-compatible IPv6 address has its first 96 bits set to zero and the remaining 32 bits equaling the IPv4 address.)

DISCUSSION

Rule-Following Versus Reflective Practitioners

A striking aspects of the IPng effort is the difference between abstract design principles and the application of these to situated contexts. A considerable body of literature, on both a theoretical and an empirical basis, has pointed out how human action always involves a significant element of situated interpretations extending well beyond predefined rules, procedures, methods, or principles (Suchman, 1987). That designers deviate from codified methods and textbooks is likewise not news (Curtis et al., 1988; Vincenti, 1990). Still, times when the manner of deviation from, application to, or expectation from design principles is made the subject of a fairly open and pluralistic discussion are rare. It is not merely the case that the actual design of the Internet does not adhere strictly to any design principles. This should not surprise anyone. More surprising is the extent to which the situated interpretations of the design principles are openly and explicitly discussed among a significant portion of the community of designers.

When outlining different approaches to systems design or interdisciplinarity, the engineering or technically inclined approach is commonly portrayed as quite narrow-minded (Lyytinen, 1987). The Internet community is massively dominated by designers with backgrounds, experiences, and identities stemming from the technically inclined systems design. The design process of IPng, however, illustrates an impressively high degree of reflection among the designers. It is not at all narrow-minded. As outlined earlier, there are numerous examples of this, including crucial ones such as how the installed base constrains and facilitates further changes, the new role of market forces, and the balance between exploring alternatives and closing down.

Aligning Actor-Networks

The majority of members of the Internet community have a well-developed sense of what they are designing. They are not designing artifacts but tightly related collections of artifacts, that is, an infrastructure. When changes are called for (and they often are), they do not change isolated elements of the infrastructure. They facilitate a transition of the infrastructure from one state to another.

> Key to understanding the notion of transition and coexistence is the idea that any scheme has associated with it a cost-distribution. That is, some parts of the system are going to be affected more than other parts. Sometimes there will be a lot of changes; sometimes a few. Sometimes the changes will be spread out; sometimes they will be concentrated. In order to compare transition schemes, you *must* compare their respective cost-distribution and then balance that against their benefits. (Rose, 1992b)

In the vocabulary of actor-network theory (Callon, 1991; Latour, 1991), this insight corresponds to recognizing that the huge actor-network of the Internet—the immense installed base of routers, users' experience and practice, backbones, hosts, software, and specifications—is well aligned and to a large extent irreversible. To change it, one must change it into another equally well-aligned actor-network. To do this, only one component (or very few components) of the actor-network can be changed at a time. This component then has to be aligned with the rest of the actor-network before anything else can be changed. This gives rise to an alternation over time between stability and change for the various components of the information infrastructure (Hanseth et al., 1996).

The crucial but neglected insight of infrastructure design is well developed in the Internet community, and the IPng case contains several illustrations of it: the difference between short-term and long-term solutions, the debate over CIDR versus C#, and concerns regarding transition mechanisms. The failure to really appreciate this is probably the key reason why the otherwise similar and heavily sponsored OSI efforts have yet to produce anything close to an information infrastructure of the Internet's character (Rose, 1992c). Hanseth et al. (1996) compared the OSI and the Internet efforts more closely.

An actor-network may become almost impossible to change by having the components accumulating too much irreversibility and becoming too well aligned with each other (Hughes, 1983). The components of the actor-network become locked into one another in a deadly dance where none succeeds in breaking out. This is not rare with infrastructure technologies. Grindley (1995) described the collapse of closed operating systems along these lines, without employing the language of actor-network theory. The operating systems were too conservative. They were locked into each other by insisting that new versions be backward-compatible with earlier ones and by tailoring a large family of applications to run on only one operating system. The danger that something similar could happen to the Internet is increasing as the infrastructure expands because the "longer it takes to reach a decision, the more costly the process of transition and the more difficult it is to undertake" (IPDECIDE, 1993).

Obviously, there are no generic answers to how much one should open an infrastructure technology to further changes or when to close down on a solution that addresses at least fairly well understood problems—or when simply to keep the old solution without changes for the time being. The Internet has pursued and developed what seems a reasonably sound, pragmatic sense of this problem: "Making a reasonable, well-founded decision earlier was preferred over taking longer to decide and allowing major deployment of competing proposals" (IPDECIDE, 1993).

Striking a balance between stability and change has to date been fairly successful. Whether this level of openness and willingness to be innovative will suffice to meet future challenges remains to be seen. It is anything but obvious.

But What About the Future?

The institutionalized framework of the Internet is under a tremendous—and a completely new kind of—pressure. This is partly due to the fact that the majority of users come from sectors other than the traditional ones. The crucial challenge is to preserve the relatively pluralistic decision process that involves a significant fraction of the community when confronted with situations calling for pragmatic judgment.

> So there it is: politics, compromise, struggle, technical problems to solve, personality clashes to overcome, no guarantee that we'll get the best result, no guarantee that we'll get any result. The worst decision-making system in the world except for all the others. (Smart, 1992)

But only a minority of today's Internet community has acquired the required sense of pragmatism about the Internet. There are signs that indicate a growing gulf between the traditional design culture and the more commercially motivated ones (Rekhter & Knopper, 1992).

The core institutions of the Internet are the IETF, the IESG, and the IAB (described earlier). Despite the fact that the IAB members are appointed from the IETF, the IAB was—especially during the heated debate over the Kobe declaration—poorly aligned with the IESG and the IETF. How, then, can the interests of the IAB seemingly differ so much from those of the IESG and the IETF? I point out a couple of issues I believe are relevant to working out an explanation.

Even if the IAB today is recruited from basically the same population as the IESG and the IETF, this has not always been the case (Kahn, 1994). The bulk of the current members of the IAB come from the computer and telecommunication industry (eight), two from universities, one from a research institute, and one from manufacturing industry. Seven are based in the United States and one each in Australia, Britain, Canada, the Netherlands, and Switzerland (Carpenter, 1996). The IAB struggled until fairly recently, however, with a reputation of being too closed (IAB, 1990). The minutes of the IAB were not published until 1990. In addition, the IAB was for some time "regarded as a closed body dominated by representatives of the United States Government" rather than the traditional designers of the IETF and the IESG (Carpenter, 1996). In connection with the Kobe declaration, this legacy of the IAB was made rhetorically use of and hence kept alive: "Let's face it: in general, these guys [from IAB] do little design, they don't code, they don't deploy, they don't deal with users, etc., etc., etc." (Rose, 1992b).

The programmatically stated role of the IAB to advise and stimulate action—rather than direct—has to be constantly adjusted. As Carpenter (1996), the IAB chair, stated, "The IAB has often discussed what this means ... and how to implement it." It seems that the IAB during recent years has become more careful when extending advice in order not to have it misconstrued as a direction. The controversy over the Kobe declaration was an important adjustment of what it is to mean when the IAB provides advice: "The most important thing about the IAB IPv7 controversy [in the summer of 1992] was *not* to skip CLNP. It was to move the power from the IAB to the IESG and the IETF" (Alvestrand, 1996).

The last few years have witnessed a manyfold increase in the IETF attendance, even if it seems to have stabilized during the last year or so. Many important elements of the future of the Internet, most notably those related to Web technology, are developed outside the Internet community in industrial consortia dealing with the HTML protocol family, HTTP, Web browsers, and electronic payment. It is not clear that all of the standards these consortia develop will ever get on the Internet standards track. These consortia might decide to keep them proprietary. Still, a key consortium like the World Wide Web consortium lead by Tim Berners-Lee has gained widespread respect within the Internet community for the way the standardization process mimics that of the Internet (see http://www.w3.org/pub/WWW). As the organization of the Internet standardization activities grows, so does the perceived need to introduce more formal, bureaucratic procedures closer to those employed within the OSI: "The IETF might be able to learn from ISO about how to run a large organization: 'mutual cultural infection' might be positive" (IAB, 1993).

An important design principle within the Internet is the iterative development of standards that combine practical testing and deployment with the standardization process (cf. discussion in third section of this article). This principle is getting increasingly more difficult to meet, as the IP revision makes painfully clear. There is a growing danger that the Internet standardization process may degenerate into a more traditional, specification-driven approach. Noncommercial actors—for instance, research networks—have an important role to play to function as a test bed for future standards (Alvestrand, 1996).

CONCLUSION

To learn about the problems of scaling information infrastructure, we should study the Internet. With the escalating use of the Internet, making changes required for scaling becomes increasingly more difficult. The Internet has never faced a more challenging task regarding scaling than its revision of IP. After years of hard work, most people reckon that IPv6 will enhance further scaling of the Internet. But even today, there is a reasonably well-founded doubt about this. We have yet to see documented testing of IPv6 segments. The real asset of the Internet is its institutionalized practice of pragmatically and fairly pluralistically negotiating design issues. Whether this will survive the increasing pressure from new users, interest groups, commercial actors, and industrial consortia remains to be seen.

NOTES

1. Changes are made for a number of reasons, including those motivated primarily by the need to facilitate further scaling. Changes related to scaling are, accordingly, a subset of the total changes. There is not always a clear-cut distinction between changes related to scaling and changes made for other reasons. As scaling in itself is "invisible" it does not add functionality, it only avoids a breakdown (cf. fifth section); other changes are aligned with those directly aimed at scaling in order to provide incentives for making changes. I consider changes made in this way are only indirectly related to scaling. It is accordingly a matter of perspective rather than principle whether these changes regard scaling.

2. CCITT is the international standardization body for standards within telecommunication.

3. The notion of a gateway is, perhaps surprisingly, not clear. It is used in different ways. In particular, it may be used as a mechanism to implement a transition strategy (Stefferud & Pliskin, 1994). It is then crucial that the gateway translates back and forth between two infrastructures in such a way that no information is lost. Dual-stack nodes and "tunneling" (see sixth section) are illustrations of such gateways. But gateways more generally might lose information, as, for instance, the gateway between the ISO X.400 e-mail protocol and the e-mail protocol in the Internet (Simple Mail Transfer Protocol). Within the Internet community, however, only gateways of the latter type are referred to as "gateways." The former type is regarded as a transition mechanism. And it is this latter type of gateway that is not seriously considered within the Internet community. The reasons for this lack of interest in gateways that lose information—and hence are "imperfect"—within the Internet seems to be a drive toward "purity" in the design (Eidnes, 1996). As this purity is likely to be increasingly difficult to maintain in the future, it would be interesting to investigate more closely the role of and attitudes toward gateways within the Internet.

4. EDIFACT stands for Electronic Data Interchange (EDI) in Administration, Commerce, and Transportation and is a United Nations standard for defining EDI messages.

5. Alvestrand (1996) suggests that had it not been for the clumsy way IAB announced its decision, many more would probably have gone along with the CLNP solution.

REFERENCES

Abbate, J. 1994. The Internet challenge: Conflict and compromise in computer network. In *Changing large technical systems*, ed. J. Summerton, pp. 193–210. Boulder, CO: Westview Press.

Alvestrand, H. 1996. Interview, October.

AREA. 1994a. IPng Area summary. [ftp://ds.internic.net/ietf/94dec/area.ipng.94nov.txt]]

AREA. 1994b. IPng Area summary, reported by S. Bradner and A. Mankin at December IETF. [ftp://ds.internic.net/ietf/94dec/area.ipng.94dec.txt]

Baym, N. K. 1995. From practice to culture on Usenet. In *The cultures of computing*, ed. S. L. Star, pp. 29–52. Oxford: Blackwell.

Bowker, G., and Star, S. L. 1994. Knowledge and infrastructure in international information management. Problems of classification and coding. In *Information acumen. The understanding and use of knowledge in modern business*, ed. L. Bud-Frierman, pp. 187–213. Routledge.

Braun, H.-W. 1992. E-mail to big-internet, 3. July. [ftp archive at munnari.oz.au/big-internet/list-archive]]

Callon, M. 1991. Techno-economic networks and irreversibility. In *A sociology of monsters. Essays on power, technology and domination*, ed. J. Law, pp. 132–161. Routledge.

Carpenter, B. 1996. What does the IAB, anyway? [http://www.iab.org/iab/connexions.html]

Cerf, V. 1992. E-mail to big-internet, 6 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Chiappa, N. 1992. E-mail to big-internet, 17 June. [ftp archive at munnari.oz.au/big-internet/list-archive]

Crocker, D. 1992. E-mail to big-internet, 26 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Crowcroft, J. 1992. E-mail to big-internet, 3 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Curtis, B., Krasner, H., and Iscoe, N. 1988. A field study of the software design process for large systems. *Communications of the ACM* 31(11):1268–1287.

David, P., and Dunn, J. A. 1988. The economics of gateway technologies and network evolutions. *Information Economics* 3:165–202.

Deering, P., and Francis, S. 1994. E-mail to the SIPP mailing list, 31 May. [ftp://Hsdndev.harvard.edu/pub/ipng/archive/sipp/]

Eidnes, H. 1994. Practical considerations for network addressing using CIDR. *Communications of the ACM* 37(8):46–53. Special issue on the Internet technology.

Eidnes, H. 1996. Interview, October.

Elz, R. 1992. E-mail to big-internet, 29 June. [ftp archive at munnari.oz.au/big-internet/list-archive]

Fleichman, E. 1993. E-mail to IPng-mailing list, 14 December. [ftp archive at Hsdndev.harvard.edu/pub/ipng/mailing.list/]

Fuller, V. 1992. E-mail to to big-internet, 27 March. [ftp archive at munnari.oz.au/big-internet/list-archive]

Grindley, P. 1995. *Standards, strategy, and policy. Cases and stories.* New York: Oxford University Press.

Hanseth, O. 1996. Developing information infrastructure: Cultivating the installed base. PhD dissertation, University of Gothenburg, Sweden.

Hanseth, O., Monteiro, E., and Hatling, M. 1996. Developing information infrastructure standards: The tension between standardization and flexibility. *Science, Technology & Human Values* 21(4): 407–426.

Hauben, M., and Haugen, R. The Netizen's netbook. [http://www.columbia.edu/~hauben/netbook]

Hinden, R. M. 1996. IP next generation overview. *Communications of the ACM* 39(6):61–71. Earlier version appeared as draft-hinden-ipng-overview-00.txt from October 1994.

Hughes, T. P. 1983. *Networks of power. Electrification in Western society 1880–1930*. Baltimore, MD: Johns Hopkins University Press.

IAB. 1990. Minutes from the IAB meeting, 3 January. [http://info.internet.isi.edu:80/IAB/IABmins.900103]

IAB. 1992a. IP version 7, the Internet Draft, July. [ftp archive at Hsdndev.harvard.edu/pub/ipng/archive/iab/]

IAB. 1992b. A summary of the IAB's proposals in response to the work of the ROAD group, Kobe, June, circulated to the big-internet list on 1 July by L. Chapin, chair of the IAB.

IAB. 1993. Minutes from the IAB meeting, 30 March. [http://info.internet.isi.edu:80/IAB/IABmins.930330]

IITA. 1995. Interoperability, scaling, and the digital library research agenda. Technical report, Information Infrastructure Technology and Applications (IITA) working group. [http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html]

IPDECIDE. 1993. BOF held at July IETF at Amsterdam. [ftp://ds.internic.net/ietf/93jul/ipdecide-minutes-93jul.txt]

Jewett, T., and Kling, R. 1991. The dynamics of computerization in a social science research team: A case study of infrastructure, strategies, and skills. *Social Science Computer Review* 9(2): 246–275.

Kahin, B., and Abbate, J., eds. 1995. *Standards policy for information infrastructure.* Cambridge, MA: MIT Press.

Kahn, R. E. 1994. The role of the government in the evolution of the Internet. *Communications of the ACM* 37(8):15–19. Special issue on Internet technology.

Knopper, M. 1994. Minutes from IPng Directorate retreat, BigTen Conference Center, Chicago, 19–20 May. [ftp://Hsdndev.harvard.edu/pub/ipng/recomendation.references/knopper94b]

Latour, B. 1991. Techology is society made durable. In *A sociology of monsters. Essays on power, technology and domination*, ed. J. Law, pp. 103–131. London: Routledge.

Lerh, W. 1992. Standardization: Understanding the process. *Journal of the American Society for Information Science* 43(8):550–555.

Lo, C. 1996. Internet History site. [http://lcweb.loc.gov/global/internet/history.html]

Lyytinen, K. 1987. Two views of information modeling. *Information and Management* 12(1):9–19.

Medin, M. 1992. E-mail to big-internet, 15 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Neumann, L. J., and Star, S. L. 1996. Making infrastructure: The dream of a common language. *Proc. Participatory Design Conference (PDC) '96*, Computer Professionals for Social Responsibility (CPSR).

NGDIR. 1994. IPng Directorate open meeting held at July IETF in Toronto, reported by S. Bradner. [ftp://ds.internic.net/ietf/94jul/ngdir-minutes-94jul.txt]

NGREQS. 1994. IPng requirements BOF held at March IETF in Seattle, WA, reported by J. Crowcroft. [ftp://ds.internic.net/ietf/94mar/ngreqs-minutes-94mar.txt]

NGTRANS. 1994. Working group summary at the December IETF. [ftp://ds.internic.net/ietf/94dec/ngtrans-tacit-minutes-94dec.txt]

Rekhter, Y., and Knopper, M. 1992. E-mail to big-internet, 3 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

RFC. 1992. IESG deliberations on routing and addressing. RFC 1380, IAB and IESG, November.

RFC. 1994a. The Internet standards process—Revision 2. RFC1602, IAB and IESG, March.

RFC. 1994b. A direction for IPng. RFC 1719, IAB and IESG, December.

RFC. 1994c. Technical criteria for choosing IP the next generation (IPng). RFC 1726, IAB and IESG, December.

RFC. 1995. The recommendation for the IP next generation protocol. RFC1752, IAB and IESG, January.

RFC. 1996. Transition mechanisms for IPv6 hosts and routers. RFC1933, IAB and IESG, April.

Rose, M. T. 1992a. E-mail to big-internet, 7 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Rose, M. T. 1992b. E-mail to big-internet, 8 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Rose, M. T. 1992c. The future of OSI: A modest prediction. *IFIP Transactions C* C-7:367–376.

SELECT. 1992. Selection criteria BOF held at the November IETF in Washington, DC, reported by P. Almquist. [ftp://ds.internic.net/ietf/92nov/select-minutes-92nov.txt]

Smart, B. 1992. E-mail to big-internet, 15 July. [ftp archive at munnari.oz.au/big-internet/list-archive]

Star, S. L., and Ruhlender, K. 1996. Steps toward an ecology of infrastructure: Borderlines of design and access for large information spaces. *Information Systems Research* 7(1).

Stefferud, E., and Pliskin, J. 1994. To tunnel or translate, that is the question. *Proc. Upper Layer Protocols, Architectures and Application*, ed. M. Medina and N. Borenstein. IFIP.

Suchman, L. 1987. *Plans and situated actions. The problem of human-machine interaction.* Cambridge: Cambridge University Press.

Summerton, J. 1994. *Changing large technical systems.* Boulder, CO: Westview Press.

Tsuchiya, P. 1992. E-mail to big-internet, 25 June. [ftp archive at munnari.oz.au/big-internet/list-archive]

Turkle, S. 1995. *Life on the screen: Identity in the age of the Internet.* New York: Simon & Schuster.

Vincenti, W. G. 1990. *What engineers know and how they know it. Analytical studies from aeronautical history.* Baltimore, MD: John Hopkins University Press.