Workshop Report: Future Directions in Network Architecture (FDNA-03)

Steven Bauer, Xiaowei Yang
{bauer,yxw}@mit.edu

Introduction

The Future Directions in Network Architecture (FDNA) Workshop, a one day workshop held in conjunction with the ACM Sigcomm 2003, provided a forum for participants to predict and consider the architectural underpinnings of future networks and the evolving Internet. The workshop was very well attended, attracting over eighty participants. A total of 48 papers were submitted to the workshop. Joint submission to the workshop and the general Sigcomm 2003 conference was permitted, and roughly half of the papers were dual submissions. Of these submissions, nine full papers and six short talks were selected for presentation. Speakers of the full papers gave 20-25 minute talks, with 5-10 minutes for questions. Short talk speakers were allotted 10 minutes for their talks with 5 minutes for questions. The presented full papers are published in the Sigcomm 2003 consolidated workshop proceedings, available from the ACM Digital Library.

Workshop Purpose

The purpose of the workshop was to convene researchers interested in the future of network architecture. The architecture of a network specifies the high level principles and structures that guide its design, especially the engineering of protocols and algorithms, and the interaction of different functional components. Advances in architecture come along two distinct fronts: 1) identification of new fundamental structuring principles and 2) new guidelines for making decisions about the functional decomposition and modularity of a system.

The current Internet architecture has been remarkably successful as the underpinning of a global, general-purpose, decentralized data communication network. Architectural decisions made 30 years ago have allowed the Internet to quickly support new applications and adapt to dramatic changes in technology. The requirements underlying the Internet architecture, however, have changed significantly since the 1970's.

New architectural requirements reflect ever growing government and commercial demands, increasingly complex security concerns, and changing user expectations and needs. Further, new classes of networks - sensor-nets, highly mobile ad-hoc nets, overlays, and others - have come into existence. These networks have very different design goals, operating requirements, and implementation environments than those imagined for traditional network architectures.

As the network community works to accommodate these new requirements, all too often the coherence of the Internet's architectural design is eroded by a patchwork of narrow technical embellishments. The result is an ever increasing complexity accompanied by a loss of functionality and extensibility. Given these pressures, revisiting the architectural principles of large general-purpose networks is appropriate and necessary.

The workshop addressed these issues in a series of four sessions. Talks included ideas on new routing architectures, new network abstractions, and new techniques and approaches for a variety of existing network problems. Each session sparked numerous interesting discussions and debates. The sessions are summarized in the following sections. We outline the talks and the discussions that followed.

Session 1: Motivations

This session addressed ways to design architectures that flexibly accommodate diverse concerns and requirements. Recognizing that various parties participate in a network for diverse reasons, this session's speakers considered how different participant's motivations can be leveraged in architectural designs to more efficiently accomplish system goals. The session introduced attempts to reason about and analyze different requirements in making architectural decisions. Finally, architectural influences of the changing nature of network technologies was considered.

Addressing Reality: An Architectural Response to Demands on the Evolving Internet

Presented by David Clark (MIT LCS)

Clark presented three key tenets that he contended should guide the evolution of the Internet in its next generation: 1) design for change 2) design for controlled transparency and trust and 3) respect the centrality of the tussle space by accepting conflicts of interest in the technology. Clark reviewed each of these tenets and then discussed their implications for future networks.

Clark explained that designing for change implies taking explicit architected action to preserve the ability to change and evolve a technology. This may imply making sacrifices in performance and efficiency. The challenge for architects is how to preserve generality and evolvability while minimizing costs.

Designing for controlled transparency is a tenet born from the recognition that users often have different requirements for the network. Sometimes the network should appear completely transparent and pass all packets but other times the network should mediate some communications. Clark postulated that the requirements often may be dictated by how much users trust each other. Users want the network itself to prevent certain traffic from reaching them if they don't trust the source. The key to this design principle is that the transparency of the network is controllable by the end user.

The third tenet was respecting the centrality of the "tussle space" in designing architectures. This tenet prescribes that architects should not attempt to resolve inevitable conflicts of interest that arise in a technology, but should instead architect to allow the conflicts to occur naturally within the coherently designed structure.

Clark explored implications of these design tenets in detail in the rest of the talk. The examples he employed came from a reexamination of the Internet and consideration of future requirements such as mobility, non-general purpose networks such as sensor nets, and evolving security needs.

Clark considered the issue of packet-switched versus circuit-switched networks and concluded that neither model is ideal. He contended that the fine-grained multiplexing achievable with packets has passed the test of time, however, we are missing an architecture for aggregates. This missing piece often triggers the erroneous call to replace packets with circuit abstractions. The lack of aggregates has lead to naming and managing them using lower level mechanisms such as MPLS. Aggregates instead should be first class objects in the network architecture.

Clark argued for a reconsideration of the "stateless faith" of Internet architects. He argued that we must accept that non-general networks will be attached to the edge of a general Internet core in the future. We must architect for this, likely necessitating application level state in the network. Clark argued that we need to "design the future" not "drift from the present."

The discussion after the talk initially focused on how architectural evolution occurs. Clark observed that incentives often control this process but technologies can be positioned appropriately by observing the interests of key players. Clark emphasized that we must only standardize the things seen on the wire and we can not and should not dictate how applications are built. He stated thought that it is crucial that network architects provide better guidance to application designers than has been done in the past.

The next series of questions probed at why Clark had limited his focus to the type of networks that he had. He was asked why multicast was not discussed. Clark indicated that the research group did not reach any consensus on multicast.

Other audience members questioned how Clark saw more exotic networks like Delay Tolerant Networks (DTNs). One person, for instance, wanted to know why Clark had assumed that the sender and receiver were static entities. A different network model could allow the networks to actively find appropriate or available recipients that were unknown or maybe didn't even exist when the packet was sent. Clark questioned the level of the networking stack that DTN style problems actually manifest themselves and suggested that the problems posed by DTNs might better be addressed at higher levels of the stack.

A Real Options Framework Illustrating the Economic Value of the End-2-End Argument

Presented by Mark Gaynor (Boston University)

Gaynor's talk presented a real options framework for understanding the economic value of competing system designs. He applied his techniques to network architectures, since, he noted, architecting networks in the past has relied more on "art" than "science". Gaynor is attempting to change this by providing a framework in which architectural questions can be reasoned about formally. He stated that he is seeking to model the long term impact of architectural design decisions.

He conducted his analysis using an extensions of options called real options, which he described briefly. Options are an economic construct that models uncertainty, flexibility and choice. Real options extends the theory of options to non-financial assets. Gaynor has employed real

options to evaluate the end-to-end architecture of the Internet and compare the value of competing technologies such as packet-switched to circuit-switched networks, of SIP compared to Megaco for VoIP, and WiFi compared to cellular networks.

Assuming that centralized and non-modular systems are ”cheaper” than distributed modular ones, Gaynor accessed the relative worth of systems in a given market context. He demonstrated quantitatively that the experimentation allowed by modular system is worthwhile when market uncertainty is sufficiently high. The intuition behind his results is that the benefits of a modular system outweigh the costs during periods when the market requirements of a system are not sufficiently well understood.

The discussion first focused on the models he employed to justify his conclusions. Gaynor acknowledged that putting real numbers into his models is difficult, but argued that his work provided a framework for a structured debate to occur. A questioner asked if he employed the Black-Scholes model of valuing options. Gaynor indicated that he did not, instead using a binominal model.

The discussion then centered on whether flexibility was actually good for providers, since it often seemingly empowers consumers. Gaynor’s response was to argue that the flexibility was indeed in the long term interest of providers. Even in the short term there are market opportunities if the flexibility exists. I-mode services in Japan were discussed as an example of modularity and flexibility being beneficial for providers.

Peer-to-Peer Network Architectures: The Next Step (Harnessing the Symbiosis of Altruism and Selfishness)

Presented by Peter Triantafillou (University of Patras)

In this talk, Triantafillou argued that too often peer-to-peer networks are viewed as networks of homogeneous nodes. However, the computational, bandwidth, and storage resources available to nodes often varies considerably. The heterogeneous nodes are distinguished by their behaviors; this is the key distinction between Triantafillou’s work and others. Some nodes are ”selfish” while other nodes are ”altruistic” and work for the benefit of the larger network. Whether or not this is true ”altruism” or nodes are in actuality deriving some benefit from their behavior is inconsequential. The point is that differences in behavior exist and can be exploited to the benefit of the network.

Triantafillou’s premise is that one can identify altruistic peers and leverage them to improve network functions. He suggested, for instance, improving routing lookups and routing functions by concentrating them at the altruistic nodes. He also suggested that one can isolate nodes behaving selfishly if the behavioral differences can be effectively identified.

While he did not have time in the short talk to provide the algorithmic or protocol details of the architecture, Triantafillou argued the general case for his design. His larger point was that the inherent structure and behavior of network nodes were important to how future peer-to-peer architectures are built.

After the talk, the discussion focused on the need to carefully balance exploiting good behavior with the potential that such exploitation could have the negative impact of discouraging desired behavior. An alternative hypothesis could be that one should architect the network to encourage the desired behavior. Triantafillou restated his position that behavioral differences exist for a variety of reasons and could be actively leveraged to improve a network.

Beyond Hosts and Routers: Some Architectural Principles for Future Mobile Networks

Presented by Robert Hancock (Siemens/Roke Manor Research, U.K.)

Hancock contended that current network architectures exhibit very high rigidity; they have fixed function modularity and asymmetric interfaces. In the future, he expects arbitrarily complex user scenarios where the ”host vs. infrastructure” division is no longer appropriate. Instead multiple nested/joined groups of cooperating devices will exist. He argued that the rigid approach does not scale to accommodate the requirements that mobile and multihomed nodes will have.

For these reasons, Hancock argued that multihoming and network composition are important organizing principles for future mobile networks. However, there are many challenging technical issues remaining in addressing these areas. The fundamental challenge in multihomed hosts is the need to separate identity and addresses. Nodes also need to negotiate for services to satisfy their diverse set of requirements.

To address these future problems, Hancock argued that control plane protocols will need an architectural framework of their own. Further the requirement for the architecture to support multihoming and composition will both blur the host/router distinction and spur a need for networks, rather than just nodes, to be nameable first class objects. Such advances will also serve to encourage more universal bi-directional interfaces.

The questions during the discussion time focused on clarifying Hancock’s position. One questioner wanted to

know if there was a fundamental new challenge as he felt that systems already successfully employed multiple radios for a variety of tasks. Hancock responded that the new challenges were 1) the methods to sew together identities and transport and 2) mechanisms to discover and negotiate services available in other networks. He stated that it is really difficult to get the interfaces right.

The discussion then turned to the properties that such networks would have. One participant noted that resilience to failure is an interesting property of such future networks; Hancock agreed.

Reconsidering the Wireless LAN Platform with Multiple Radios

Presented by Victor Bahl (Microsoft Research)

This talk was a demonstration of how changing technologies are influencing future networking architectures. The evolving technology in this case was radios for mobile or wireless devices. Radio technology is becoming increasingly cheap with predictions for even more price drops in the future. Some radios can be bought for under $5.00 apiece according to Bahl.

This evolution has lead Bahl's research group to consider how to make wireless networks more robust by exploiting multiple radios on a device. Most often this is multiple different radios that have different performance characteristics for different operations. Bahl noted that many people would claim that the real future lay in software radios. He preemptively countered that software radios are not a viable option even a number of years out due to the significant cost of the hardware required.

Bahl demonstrated a number of different ways that multiple radios can be used to significantly improve performance metrics such as power consumption. He addressed the question of how one could take a systems level approach to determine how to use multiple radios.

A variety of specific technical questions were asked of Bahl after his talk. One participant wondered whether Bahl was considering multiple and different radios only or were their applications for multiple and not different radios? Bahl responded that they were most often leveraging the differences between types of radios to accomplish their goals. Questioning where Bahl had placed the virtualization layer for the radios, one participant complained that he would lose valuable information that he would need to make certain decisions. Bahl indicated that the questioner could simply put his logic at the virtualization layer instead of above it.

Session 2: New Abstractions

Conventional networking relies upon a well-understood set of abstractions. These abstractions simplify the reasoning about network architectures and provide a common understanding and way of discussing networks. When the networking community considers new functionality, we often rely upon our conventional abstract models. During this session presenters challenged this conventional wisdom by presenting new abstractions for organizing network concepts. These abstractions potentially allow us to conceive of different ways of organizing networks, thus enabling new functionality and better prospects for evolving the architecture in the future.

Plutarch: An Argument for Network Pluralism

Presented by Andrew Warfield (University of Cambridge)

Warfield explained that IP's philosophy is to enable inter-networking by homogenizing the network and transport layers. By standardizing the middle, layers above and below are free to evolve. The fast growth of IP has demonstrated the power of this hourglass-shaped architecture. Warfield then proceeded to challenge this underlying philosophy. He listed problems associated with the current IP architecture, including, functional deficiencies (e.g., lack of support for mobility), scaling issues, the lack of freedom to innovate, and the ever increasing heterogeneity of networks with different capabilities, e.g., sensor networks, ad hoc networks, and wireless networks, he argued that it is less and less obvious that a ubiquitous transport and name resolution protocol is the right solution for the future.

In contrast, Plutarch emphasizes less homogeneity, and states that an inter-networking architecture must allow communications between dissimilar networks without mandating a standardized data path. There are two fundamental concepts in Plutarch: Contexts and Interstitial Functions (IFs). A Context is an area of the network that is homogeneous to some extent. It serves two purposes. First, contexts serves as descriptors for composing end-to-end services; Second, contexts describe communication mechanisms with which end points might use for session setups. Interstitial Functions exist at the borders between contexts. The primary function of IFs is to allow data with different naming and addressing schemes to cross contexts. The contemporary realizations of IFs include NAT boxes.

Warfield gave an example to illustrate the concept of

Plutarch. In the example, a user attempted to connect from a GPRS laptop to a sensor net via the Internet. In Plutarch, three stages happen before communication. First, a distributed and decentralized search service, not DNS, is used to resolve name=value pairs for addresses. Second, as the query results may include chained contexts, two IFs, one at the border of the sensor net and the Internet, and the other at the border of the Internet and the GPRS network, would be instantiated and installed. Finally, applications bind to the newly created chained context. Communication can then be established.

Warfield described several future directions for his research, including scalable name lookup services, correct semantics of IFs, fault tolerance, and garbage collection.

During the discussion time, concerns were raised about the degree of complexity in Plutarch since Plutarch requires complicated IFs to translate between contexts. The solution to avoid the complexity seems to be to standardize data representation.

The next comment came from a audience member who shared his experience in the design process of Application Level Framing. He stated that the idea of having states in the middle of the net and allowing asynchronous data transfer starts to break down when mechanisms must be added to handle failures. For example, what happens if an intermediate buffer is filled up? The complexity gradually builds up. Finally, he stated that he gave up the idea and decided to let applications handle end-to-end state.

Another participant observed that there is a commonality between DTNs and Plutarch. Warfield agreed but noted that the Plutarch approach is to make the IFs generic and not architect the system to handle disconnected networks. Finally, in response to a security question, Warfield noted that security remained an interesting, but largely unexplored issue and was something they hoped to explore more thoroughly in the future.

Designing for Scale and Differentiation

Presented by Karen Sollins (MIT LCS)

In her talk, Sollins formalized the idea of grouping and subdivision using an architectural abstraction called a ”region”. The region abstraction is being explored as a general organizing principle that can be leveraged to improve network functionality. The need for this new organizing principle arises because the network is becoming increasingly heterogeneous.

Sollins explained a region is defined by a set of characteristics or invariants, for example addresses or AS numbers, but perhaps much more complicated groupings. Regions come into existence in many ways. An entity acquires a region membership through explicit insertion. Insertion implies applying invariants to the entity. Insertion into a region may fail if the invariants on the entity do not hold. This does not imply human involvement and may be automated. Since a region has a logical boundary, when an entity cross a region boundary, states may be changed, and interested parties may be notified.

Instantiations of regions include Autonomous Systems, DNS, and security regions defined by firewalls. Sollins and her students have been working on a variety of region-related projects, including garbage collection in regions, inter-region information exchange, region adaptation, and security. Several important lessons have been learned from these projects. First, manageable exchange of information about different abstractions of region invariants can have significant performance benefit. Second, understanding where tradeoffs are necessary is critical. Third, much more work is needed on questions of region creation and membership.

During the discussion Sollins was asked why people have recently started to think philosophically about large scale heterogeneous problems after the community has been silent for so long on this topic. Sollins posited that it was a political problem; the community needed to get the basic things done first. Before, we focused on TCP performance and multicast, now Sollins claimed is a time to step back and think about what we are trying to do in a much more general way.

The next questioner asked how communication between regions would occur. Was the right idea is to pick the least common denominator at the translation points or to have complicated translation points. Sollins answered that in reality, both approaches would exist. Perhaps the best that could be done was to architect for the two approaches to coexist and make them as efficient as possible. Sollins further elaborated that the two approaches were different and could not be compared directly.

A questioner challenged Sollins whether we really want all these ”sandboxes” (regions), or would common standards be better. Sollins responded that the question might be answered by economic arguments. She did not think we can predict what will be the shared functionality. What we want to build are systems where we can have architected heterogeneity.

Building an Internet Control Plane

Presented by Tim Gibson (US Army/DARPA)

Gibson addressed the need for an Internet control plane. He argued that the original design assumptions are not true for today’s Internet. He argued that we should not

treat the network as a black box any more.

Gibson's solution is to have a control plane in the network that end hosts can employ to configure network resources. His control plane has three building blocks. First, the infrastructure provides performance information and end hosts can query the infrastructure for the information. Second, end hosts are authoritatively identified. Third, end hosts pass instructions to infrastructure for their traffic, such as requesting alternate network paths. Gibson emphasized that the control plane is not a knowledge plane as it does not assume applications that rely upon cognitive or AI techniques.

The discussion centered on questions of scale and functionality required from such a control plane. Gibson was asked whether he has considered scale issues; if end hosts all send queries, will the system scale? Gibson said the research is a work-in-progress and he is more focused on providing services to end hosts and has not given much thought to scalability issues.

One audience member observed that ultimately what we learn from the data plane is often better than what we can get from a control plane. So maybe we do not actually need a control plane at all in the conventional sense of the word.

Finally, a comment was made on the spectrum of service semantics covered by the control plane. The audience member noted that there are some things we can easily know and some things we simply cannot know. If the control plan has to provide an absolute guarantee on the answers it provides, then the protocol would likely become overly complex.

A Virtual Internet Architecture

Presented by Joe Touch (UCS ISI)

Touch presented a short talk on a virtual Internet architecture (VIA). A virtual Internet is a network composed of virtual hosts (VHs), virtual routers (VRs), and virtual links (VLs). VHs and VRs are connected by encapsulated tunnels (VLs). It provides at least the same services as the Internet Architecture but in a virtual context. Touch emphasized that this is a first principle extension to the existing Internet, and not just a patch, or an interim solution. A virtual Internet facilitates the incremental deployment of new services. As an example, Touch argued if we are going to build a different kind of forwarding scheme, without a Virtual Internet, we have to rebuild everything to support the forwarding.

A VIA emulates the Internet, yet decouples services from their base networks. It supports recursion, i.e., some of VRs are VS networks themselves. A combination of BGP and ARP, called "BARP", enables recursion. It also supports revisitation, i.e., a node can be in the same overlay more than one time. Currently, running code on FreeBSD, Linux, and Cisco exists to demonstrate the concept of VIA.

During the discussion time, Touch was questioned as to why revisitation is needed. His response was that that it is needed for the same reason we need multiple processes. This point was debated further in more detailed discussion. Another member asked Touch how resource partitioning was accomplished in a VIA. Touch said that a node would first ask the operating system and then ask the network for resources.

Session 3: Routing

In this session, two papers presented new routing architectures. These new routing architectures potentially support functionality that cannot be achieved in the current Internet architecture. The other talk of the session presented work ongoing in the development of a logic for global routing that would facilitate reasoning about protocols and policies.

BANANAS: An Evolutionary Framework for Explicit and Multipath Routing in the Internet

Presented by Shivkumar Kalyanaraman (Rensselaer Polytechnic Institute)

Kalyanaraman opened the routing session with a talk on "BANANAS", an evolutionary framework for explicit and multipath routing. BANANAS is not an acronym, but an analogy, adapted from the comedy "Herbie goes Bananas". Kalyanaraman observed that Internet paths have much multiplicity, and proposed to use a single mechanism to exploit various forms of multiplicity (such as intra-domain multiplicity and inter-domain multiplicity). Although a lot of work has been done in multipath routing, BANANAS addressed the question of whether we can do multipath and explicit routing without signaling, without variable and large per-packet overhead, being backward compatible with OSPF and BGP, and allowing incremental network upgrades.

The key idea of BANANAS is to compute global path identifiers from well-known global variables such as routers' IP addresses and Autonomous System Numbers in order to avoid explicit signaling for path setup. Kalyanaraman explained that ATM and MPLS encode paths using local labels, thus requiring a signaling protocol to map global identifiers to local labels. The canonical method

to avoid the signaling cost in the BANANAS approach is to compute MD5 hashing of the sequence of node identifiers followed by a CRC-32 checksum to get a 32-bit hash value as the global path identifier. Kalyanaraman showed that simple extensions can be added to OSPF and BGP to implement BANANAS. He also showed examples for partial deployment.

The discussion after the talk focused on the scalability of the framework. It was observed that BGP only maintains a single path but is already very complicated, and a router already has too many routing entries. Kalyanaraman was asked how BANANAS scaled and is their per-flow state in the routers? He responded that the framework provides tools and mechanisms to control how much state a router maintains.

Towards a Logic for Wide-Area Internet Routing

Presented by Nick Feamster (MIT LCS)

In this talk, Feamster argued that a framework for reasoning about wide-area Internet routing protocols is required. Protocol designers and network operators need a way to describe and reason about protocol behavior. He described his framework as a routing logic but emphasized that his focus is on practical improvements for routing protocols.

According to Feamster, a routing logic includes two components: properties and rules. First, high-level properties describe the behavior of a routing protocol. Second, these properties are defined in terms of rules (i.e., sufficient conditions) that are easier to reason about. Feamster demonstrated the practical use of a routing logic using BGP, a deceptively simple protocol. Feamster listed the five key properties BGP should have: validity, it advertises valid routes; visibility, every valid path has a corresponding route; safety, it will converge to a unique stable answer given a set of choices; determinism, the converged state is not affected by the order of messages or the set of available routes; and information flow control, it should not expose more information than intended. Feamster showed examples of rules that define the validity property and the information flow control property.

Feamster briefly addressed open problems at the end of his talk. First, protocol timing related issues are not covered in the logic yet. Second, configuration validation is sometimes about verifying intent of an operator (e.g. aggregation), which makes some things difficult to address with a routing logic. Third, more work is required to figure out how to apply the routing logic framework to routing protocols other than BGP.

The discussion session focused initially on the capabilities of the routing logic. It was observed that the related work is all about reasoning about negative properties. Similarly the audience member asked, it seems like this may be only able to prove that things are going to go bad; can one prove that things are going to go right? Feamster said that constructive results are shown in the paper. However, he cannot promise that it is possible to say the configuration is guaranteed correct.

Feamster next was asked whether the logic is accessible for a regular network operator. Feamster said they are trying to move towards that direction. Ideally, one can take a configuration and ask if it is going to go on a router without causing a problem using the routing logic.

Feamster was then asked to contrast and compare his work on papers presented the day before (i.e., one on a BGP algebra, and one on BGP policy languages). Feamster said both papers focused on convergence aspects, but his work does not touch on convergence at all. He explained his is more about a formal reasoning framework and protocol design.

NIRA: A New Internet Routing Architecture

Presented by Xiaowei Yang (MIT LCS)

Yang presented a talk on the design of a new Internet routing architecture (NIRA). The author observed that in today's Internet, users can pick their own ISPs, but once the packets have entered the network, users have no control over the routes their packets take. The author argued that it would be a better alternative to let users pick their domain-level routes because user choice fosters competition. The author used the telephone network as an example. In the telephone system, users are able to choose long distance providers separate from their local providers. This user choice has created a competitive market for long distance providers and has significantly reduced the cost for long distance phone calls. The author hypothesized that the stagnation in introducing new services, e.g. QoS, in today's Internet was a signal of the lack of competition in the form of user selected routes.

The design of NIRA focuses on the domain-level choices because choices at this level encourages ISP competition. It leaves router-level choices unspecified. The design is optimized for the special structure of the Internet. In the Internet, business relationships determines transit policies between domains. Common transit policies state that providers would provide transit service for customers, but not vice versa; peers only provider transit service for each other's customers. According to these policies, domain-level routes are said to be "valley-free". There exists a densely connected "core" of the Internet.

Based on this structure, the author proposed to use

a fixed length and provider-rooted hierarchical addressing scheme to optimize route representation. Top-level providers obtain unique address prefixes and recursively allocate subdivisions of the addresses to their customers. As a result, the route segment from a domain to a top-level provider can be uniquely identified by an address prefix. Two addresses can be used to represent a valley-free route. Source routing header is used for representing non-valley-free routes. The assumption is that valley-free route is the common case so that this route representation scheme is efficient.

The design of NIRA separates the task of route discovery into two halves. Each user only needs to know his part of the network as it grows. The infrastructure service, Topology Information Propagation Protocol (TIPP), tells users topology information of his providers, and the corresponding address allocation information. TIPP is a policy-based, link-state like protocol and is less likely to suffer from the slow convergence problem of BGP.

The user looks up the destination's addresses, and optionally the topology information of the providers of the destination , using another infrastructure service, Name-to-Route Resolution Service (NRRS). Similar to DNS, NRRS uses a hierarchical namespace and hard-coded addresses for bootstrapping. However, because addresses in NIRA are topology-sensitive, a fundamental tradeoff is that topology changes would cause address changes. The author argued that only static topology changes, such as providers changing peering agreements, would cause address changes. Recent Internet measurement results shows that the static Internet topology changes at a relative low rate, and therefore NRRS server updates are likely to be manageable.

Route failure handling in NIRA is a combination of proactive notification and reactive discovery. TIPP proactively notifies a user of route failures regarding to his part of the network. When a packet traverses across the destination part of the network, a router finds the route specified in a packet header is unavailable, the router will try its best to send a control message to the sender of the packet, notifying him of the route failure. In cases such router feedbacks are unavailable, users shall always use timeout to discover route failures.

The author also briefly discussed provider compensation problem. Similar to today's Internet, NIRA's provider compensation model is based on contractual agreements between providers and customers. As users can specify arbitrary routes in their packet headers, providers shall install policy filters to prevent illegitimate route usage. In cases where business relationships exist only in directly connected entities, policy filter checking for valley-free routes becomes verifying that the source address in a packet header matches the interface the packet comes from. For more complicated routes, policy filter checking may require matching source routing header against policy filters. For indirect business relationships, policy checking becomes verifying the identify of the sender of a packet at forwarding time. The author pointed out that it was an open but general problem. Overlay providers, or second-hop providers that want to sell QoS to end customers face the same problem. Any solution to this general problem can be plugged into NIRA.

Yang was asked during the discussion why projects such as NIMROD (mentioned in her related work) had failed. Yang disagreed that these routing architectures were a failure. She commented that as a research project, NIMROD has influenced the design of a number of protocols. She noted that research is much like fashion design; people take ideas from research projects and made on-street versions.

Yang was questioned next about the motivation for providers to implement NIRA. Providers lose control over routes, and the market would become more competitive for them. Yang explained that providers that are already in monopoly position probably do not want competition. But for smaller providers, if they want to enter the market, they probably would welcome user choices. Yang explained that employing user choices to discipline the ISP market is a hypothesis; only real experiment could justify it. She believed it was a worthy experiment and this work is the first step towards it.

The final question dealt with whether NIRA would work well in situations where peering agreements become more dynamic. Yang said she expected the peering agreements would change at a manageable time-scale in the real Internet.

Session 4: New Techniques and Approaches

The final session of the day contained a set of papers presenting new techniques and approaches for addressing various networking problems. The presenters considered new principles for organizing addressing and the ensuing implications for the Internet. An argument was presented for why denial of service attacks should be addressed at the TCP/IP level of the end users network stack. Finally, a case was made for radically changing the service model of the Internet to expose network storage primitives to enable programmable networking.

FARA: Reorganizing the Addressing Architecture

Presented by Robert Braden (USC ISI)

Braden's talk presented a new addressing architecture for networks. It was a product of the NewArch Project that explored whether abstract architectural reasoning can help in creating a better technical design for the Internet to meet today's and future requirements. This talk presented one exercise in exploring whether or not that premise was actually true. Braden emphasized that this is not a "rerun" of the design effort of the original Internet. The original design effort was largely bottom up, finding one approach that met the apparent requirements guided by some abstract thinking about protocol modularity. In contrast, this research was a top down approach.

The authors went through a three stage process. First, they defined an abstract architectural model for addressing that encompassed an interesting part of the design space but left many of the details unconstrained. They then defined an architecture that instantiates this addressing model, and finally they built a prototype system.

The abstract model for addressing they developed rested on the core principle of cleanly decoupling end-system identity from network layer forwarding functions. The motivation for this split was the traditional location identity split. The implications of this in their work were the need for a remodularization of function into entities and associations. Braden explained each of the concepts in detail in the talk.

To demonstrate the consistency of the addressing model they developed, Braden then presented an instance of it, M-FARA. This represented one point in the spectrum of choices laid out by their addressing model. He also discussed the relationship between their addressing architecture and the IPv4 architecture.

During the discussion that followed people primarily challenging the model that Braden had described of entities and associations. One person for instance pushed for the need for distributed entities (anycast). Braden responded that this was the wrong model if entities were that distributed. Another person questioned whether Braden thought it would be important to be able to change entities that were involved in an association. Braden responded that this would be a fundamentally different association.

The Case for TCP Puzzles

Presented by Wu-chang Feng (Oregon Graduate Institute)

Distributed denial-of-service (DDos) attacks via worms and viruses continually disrupt the Internet. Proposals for how to deal with such attacks abound. One mitigation approach is to require that senders do a small amount of work for each network operation that they initiate. This work often involves solving some puzzle that is designed to take a small amount of time. Malicious sources that initiate large amounts of attack traffic would be slowed while "regular" senders would be only minimally delayed.

Feng argued during his talk that puzzles must be placed within the TCP/IP protocol stack layer in order to provide protection. The key reason for this is that denial-of-service activity can happen at any layer and only needs to break one link in the end-to-end chain in order to be successful. He argued that there is a corollary to the end-to-end principles that one must put security functions in a common waistline layer if the security property is otherwise destroyed unless implemented universally across a higher and or lower layer.

Feng discussed implementing the network layer puzzles in the form of a "push-back" firewall that required senders to do some amount of work before allowing a connection through the firewall to be opened. The difficulty of the puzzles presented to senders is dynamically adjustable. Feng discussed the research challenges presented by his approach. These included how to make the puzzle system itself resistant to denial of service attacks, how to ensure that the mechanism is tamper resistant to replay and spoofing attacks, and how to design control algorithms and ensure that hosts with different computational power are fairly treated.

During the discussion Feng was asked to clarify various points about his system. It was noted that he must be assuming a streaming model where puzzles were valid for some number of packets. Feng agreed and proceeded to describe his various approaches for where puzzles were issued and how long the results were valid once a puzzle was solved.

There were a variety of questions about the need to do the puzzle solving at the IP layer. One participant asserted that the right layer for the puzzles was the layer that the attack was occurring. Feng elaborated on why he felt it was only appropriate to implement the puzzles at the IP layer. He argued we did not need multiple puzzle mechanisms at each layer of the network stack.

An End-to-End Approach to Global Scalable Programmable Networking

Presented by Micah Beck (University of Tennessee)

Beck's talk was a presentation on a new architecture for a global network that exposes storage and computation primitives. The goal of the new architecture is to support globally scalable programmable networking by exposing network primitives that are generically useful to advanced

applications.

Beck appealed to the end-to-end principles in designing the network compute services. He emphasized that scalability can only be achieved by adhering to the principles that have guided network architecture design. His appeal rested on the premise that the arguments dictate not where functionality should be placed but rather the scalable nature of that functionality.

He claimed that the most important consequence of requiring scalability is that the semantics of the services offered in the network must be simple and weak. If the semantics are too complex, the services will fail the requirement that services at intermediate nodes in the network be generic. If the services are too strong then they will not compose with a scalable network without breaking.

Building on this philosophy his research group has built Logistical Networking (LoN), which exposes network storage primitives. In his talk, Beck reviewed the LoN research and discussed how to extended it to supply an abstraction of the execution layer resources. He presented an abstraction for time-sliced operation system services and discussed how they could be employed to create network services.

Beck was questioned on how his system would change the legal status of network operators. Network operators are now considered common carriers because they do not actually store or manipulate the data they transport. If they actually stored and operated on bits, then they would potentially expose themselves to legal liability. Beck responded by citing a number of cases that he claimed established precedence that the storage of bits did not create any legal liability for network operators.

Beck was then questioned about the relationship between his work and active networking. He emphasized that while the goals of active networking are similar, his approach is dramatically different. Activate networking, he claimed, does not fundamentally concern itself with scalability, while scalability is his primary concern.

Acknowledgements

The workshop organizers are tremendously grateful to the scribes, Steve Bauer and Xiaoewi Yang, for their critical role in capturing and reporting on a full day of rapid, nonstop presentations and discussion.

The workshop itself could not have materialized without the (unanticipatedly large) efforts of the program committee - Andrew Campbell, Ted Faber, Mark Handley, John Heidemann, Larry Peterson, James Sterbenz, and John Wroclawski. Thanks go to the ACM SIGCOMM2003 Workshop organizers (and in particular Craig Partridge) for providing advice, assistance and logistical support for workshop advertising, website and registrations.

Finally, and most importantly, we thank the presenters and participants in the workshop for their interesting and thoughtful discussions. The organizers are most grateful for their time, interest, and thoughtful energy.