NEBULA - A Future Internet That Supports Trustworthy Cloud Computing

Tom Anderson, Ken Birman, Robert Broberg, Matthew Caesar, Douglas Comer, Chase Cotton, Michael Freedman, Andreas Haeberlen, Zack Ives, Arvind Krishnamurthy, William Lehr, Boon Thau Loo, David Mazières, Antonio Nicolosi, Jonathan Smith, Ion Stoica, Robbert van Renesse, Michael Walfish, Hakim Weatherspoon and Christopher Yoo

Abstract

NEBULA is a future Internet architecture that is intrinsically more secure and addresses threats to the emerging *computer utility* capabilities called cloud computing while meeting the challenges of flexibility, extensibility and economic viability. NEBULA's architecture surrounds a highly-available and extensible core network interconnecting data centers with new trustworthy transit and access networks that enable many new forms of distributed communication and computing. NEBULA mobile users will have quick, secure, 24x7 access to services such as financial transactions and electronic medical services at any location. Local device software systems will evolve to select from a continuum of distributed computing and storage services provided by data centers accessible via NEBULA. A major technical concern for such an architectural vision is trustworthiness, e.g., that each user's data is kept private and that communication is always available. NEBULA addresses the security properties of confidentiality, integrity and availability with a systems approach.

NEBULA has three interrelated parts: (1) the NEBULA Data Plane (NDP) that establishes policy-compliant paths and provides both flexible access control and defense against availability attacks, e.g., DoS; (2) NEBULA Virtual and Extensible Networking Techniques (NVENT), a control plane for NEBULA, that provides access to application-selectable service and network abstractions such as redundancy, consistency, and policy routing; and (3) the NEBULA Core (NCore) that redundantly interconnects enterprise data centers containing replicated data with ultra-high availability next-generation core routers developed in collaboration with Cisco. NVENT provides new control plane security with policy-selectable network abstractions, including multipath routing and use of new networks as they become available (and thus complements many other networking projects). NDP employs a novel provenance approach to network path establishment, exploiting cryptographic mechanisms to establish policy-controlled trustworthy paths among NEBULA routers.

1. Vision

We are, at last, on the verge of realizing the *computer utility* vision [54,63,75]. Its name today is *cloud computing* [18]. In 1965, two of the Multics architects, Corbato and Vyssotsky, stated [54]:

"*Such systems must run continuously and reliably 7 days a week, 24 hours a day in a way similar to telephone or power systems, and must be capable of meeting wide service demands…Such information processing and communication systems are believed to be essential for the future growth of computer use in business, in industry, in government and in scientific laboratories which would be otherwise undone. Because the system must ultimately be comprehensive and able to adapt to unknown future requirements, its framework must be general, and capable of evolving with time.*"

The computer utility now emerging [18] differs from Multics in that it is a *global-scale distributed computing infrastructure* composed of multiple data centers intended to support hundreds of millions of users. Forms of cloud computing such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) have emerged to provide demand-driven allocation of virtually arbitrary amounts of computing and data storage. Data centers provide dynamically-allocated storage facilities that hold business data, videos, logs, blogs and backups. This computing model offers significant economic advantages of scale for purchasing, operational advantages of scale from automated operations, and energy usage advantages since the number of idle machines can be reduced by sharing hardware among multiple applications. Cloud services may also offer a potential advantage for security and economics because administration (especially software upgrades) can be performed promptly, and the cost of elite security professionals can be amortized over many machines.

The *missing piece is a network architecture*, which must both interconnect data centers and connect users to their data. To facilitate low latency, we envision many data centers. To support mobility, we envision that data centers will be coordinated, and will migrate a user's data from one data center to another as the user moves. The core of the new Internet must be engineered to provide high availability and increased security for traffic that flows among data centers; it will become a trusted part of societal infrastructure that must be immune to attack and upon which we can depend for services during times of crisis.

To set the stage for the NEBULA (Latin for cloud) architecture, imagine a future healthcare application that might run on a future Internet. A diabetic wears both an insulin pump and a continuous blood glucose monitor. Measurements from the blood glucose monitor are sent to a NEBULA data center every 5 minutes. These measurements are recorded and analyzed against historical data from the individual and anonymously correlated with masses of data from other data sources, including records from other diabetics. The analysis includes data mining algorithms to estimate appropriate micro-dosages of insulin to be delivered by the pump, and detects anomalies. The anomalies are forwarded to alert human experts who can ensure that no medical problem has occurred (for example, side-effects from a concurrent therapy). Dosage numbers are fed from the cloud into the patient's insulin delivery system for infusion.

The challenges of devising a future Internet to support such applications are substantial. First, the architecture must provide high availability because the network may be part of a life-critical medical feedback loop with timeliness constraints. Second, the network must provide access wherever the user or users are located. Third, the architecture must provide a network path between the access and data center over which data can be guaranteed confidentiality and integrity. The data must remain correct and safe in the face of malicious acts as well as equipment errors. Solving *all* the challenges of the insulin pump application involves details of medical data representation, government regulations and machine learning that go far beyond network architecture. Still, the proposed NEBULA architecture offers a core with high availability and the flexibility to add technologies and protocols; the fundamental security properties of confidentiality, integrity and availability; and is technically and economically viable.

2. The NEBULA Future Internet Architecture

A future Internet architecture must address three fundamental challenges: (1) it must intrinsically (by architectural choice) be more secure, against both threats that that have arisen since the original Internet architectural principles were laid out and threats that are yet unknown; (2) it must provide flexibility and extensibility to support further evolution of applications; and (3) it must provide a viable path for migration and deployment that is conscious of technical feasibility, economics, and regulation. Attempts to reinvent the Internet that ignore any of the three are doomed.

NEBULA is an architecture that addresses all three challenges with new research that can enable many new classes of applications that are technically advanced, economically attractive and trustworthy. To support a cloud-oriented model of computing, our architecture is based on three key insights, each of which is discussed in more depth in later sections:

- Any future global scale Internet, like the current Internet, is likely to involve many organizationally distinct network service providers. *It is impossible to predict* what policies service providers may need in the future, so our approach is to provide a data plane that is efficient yet policy neutral, permitting industry to evolve policies that reflect business needs, government regulations, and user demand for control over route selection and resource allocation. Our data plane model is "deny by default": all parties, including the end user, internet service provider, cloud computing operator, and the application provider, must consent to the path and its behavior for the path to be used. (However, if an entity wishes to delegate or abdicate its power to grant consent, it may do so.) Further, all parties can verify that their requirements have been met. In our view, this flexible, verifiable data plane is a strict requirement of the future Internet. Many of the security, reliability, and performance problems of the current Internet are due to the

inflexibility and inherently unverifiable behavior of its policy enforcement mechanisms. This is the focus of the NEBULA data plane (NDP) effort, described in Section 4.

- Trust requires that every component of the network have externally verifiable behavior, so that failing devices, software implementations, and even service providers, can have their impact identified, isolated and contained. By providing strict behavioral characterizations of network elements, we can allow *policy* to be set *declaratively*, that is, with precise, predictable impact. In today's Internet, configuration errors are rife, because of the inherent complexity of low level, operational semantics for the Internet's control knobs and because Internet administrators are intimately involved in performance optimization. Improving the reliability and security of the future Internet requires us to take a higher level approach: put simply, we need to get humans out of the details. Industry is already moving in this direction, but realizing the potential of this approach requires considerable research. This is the focus of the NEBULA Virtual and Extensible Networking (NVENT) effort, described in Section 5.
- Increasingly, *routers themselves will be built in the same way as data center computing and storage is today*: out of modular components that can be assembled into any scale system that is needed to support the desired workload. While this might appear cosmetic — ISP Points of Presence have long been built as stylized networks of individual routers — the difference is that the collection of hardware that forms a router can now be managed as a single system akin to how a data center is managed: with fault-tolerance techniques to ensure that it is always available, with atomic hot upgrade at every level, and with the ability to redirect slices of traffic to new versions of hardware and software for rolling out new protocols and services. In fact, continuous router operation demands these techniques. Thus, in addition to the interdomain and intradomain aspects discussed above, we have a third focus*: intrarouter*. Again, industry is already moving in this direction, but realizing the potential of this approach requires considerable research; this is the focus of the NEBULA Core (NCore) effort, described in Section 6.

Our research has specific goals, which we outline next; subsequent sections describe how we will achieve these goals.

3. Goals

Security and trustworthiness. A new Internet must go beyond availability and robustness to assure users that their data will be kept safe and confidential. The Internet will also need ways to ensure that the *network path that data traverses is trustworthy*, that data arrives unchanged, and that data is confidential during all steps of communication. We address these concerns in the NEBULA Data Plane (Section 4). Communications over long distance are of particular interest because as a user's data migrates from one data center to another, it may pass across networks that are owned and operated by independent groups that do not all follow the same routing policies. Joint control over the path taken is a requirement to ensure that policy and legal constraints are followed (Section 5). Isolation of computation and storage within a data center is beyond the scope of this proposal, but we assume that virtual operating systems will provide guarantees, and will look for ways to unify authentication and authorization mechanisms used by the operating systems and the network. We must also re-architect routers themselves to be based on formal methods to achieve trustworthiness and reliability of the underlying software infrastructure (Section 6). *The future will demand a diversity of security and trust models, with a diversity of implementations that will often demand specialized communication support. Our architecture incorporates the needed flexibility and offers the key mechanisms to guarantee trustworthy operation across a federation of independent subunits.*

Highly available and reliable services with non-disruptive upgrades. Before they will trust cloud providers with data and computational services, users must be assured that both data storage and access are guaranteed. Thus, *networks used to facilitate cloud services must be reliable and highly available*. In fact, next-generation networking equipment must be designed to operate continuously with *no scheduled down-time for routine maintenance or periodic reboot*. The supporting software infrastructure will need

to exploit state-of-the art software robustness mechanisms. Because we must anticipate an increasingly hostile operating environment, the systems must tolerate outright attack, in addition to the usual notions of reliable hardware and software achieved through redundancy, hot spares, and rapid recovery schemes. To allow a provider to change services or deploy new services without removing old services [84,162], the infrastructure must support a form of virtualization. Just as cloud vendors support multiple copies of application code running simultaneously, network equipment vendors are aware that future routers will need to support multiple copies of routing protocols running side-by-side without interference, with one version in production and a new version being tested before being deployed. Moreover, this reliability must persist in the presence of attacks, and the mechanisms used in the network must be tightly integrated with the reliability and security mechanisms used in the cloud data centers that host services. *A key element of our proposal is that we assert that this set of problems can be solved, and plan to prove our claim by building a working system.*

Integration of data centers and routers. Because a modern core router is a large distributed system comprising multiple racks, a key part of our research will focus on integration between the cluster of computers in a data center and a core router. Multiple physical connections [177] will be used to achieve both reliability and high throughput. Because parallel forwarding paths will exist, new addressing and routing problems arise, and new routing protocols will be needed to balance traffic and make optimum use of the interconnect between a data center and the Internet core routing system. *We will break the barrier between the data center and the Internet.*

Evolve with technology. Industry does not stand still, and to be adoptable, our research must be shown to work with the highest end equipment available. In addition to tracking router performance, low latency will become a key requirement as more users engage in real-time collaboration, High capacity transport service will be especially important in the Internet core because migration of data among data centers implies that large volumes of data (including virtual machine images of several gigabytes) may move when a user changes location. Furthermore, video traffic will continue to increase, meaning that in addition to accommodating additional users, the new design must accommodate a higher per-user traffic demand. Our collaboration with Cisco allows unprecedented early access to next-generation core routing systems, which include a complete bottom-to-top rethinking of the architecture of the router control subsystem. *As participants in the process, we will be able to influence all aspects of the design.*

Economic and regulatory viability. Because they are operated by major service providers, core networks are subject to many telecommunication regulations. For any architecture to be economically viable, the design must take into account the regulations and guidelines imposed on the industry. We will study regulatory constraints and ensure that our innovations are within these boundaries. *No solution can succeed unless the operators of the network and cloud see the approach as both mutually advantageous and viable within the regulatory constraints. Because expertise in our group crosses disciplinary boundaries, we will be able to ensure that our solution is viable.*

4.0 Research Agenda: NEBULA Data Plane (NDP)

In this section, we consider the requirements on the NEBULA data plane. How can we accommodate the broad variety of potential policy requirements by the various stakeholders in the future Internet, with an architecture that can be formally verified and efficiently implemented? A blizzard of proposals has been made by network researchers, including many just from those researchers participating in this proposal, but also from outside. These projects, in one way or another, have been to grant increased rights to various participants in the communication, to constrain what other participants can do, to require specific in-band processing (or to prevent it), to improve performance and reliability, and/or to address specific known security flaws. Although one could argue for or against any of those particular projects as a specific design point, we believe that the future should ultimately decide – we explicitly argue for separating the next Internet's mechanisms from the tussle space [48] of its participants.

Of course, from early work on software routers to the more recent efforts, such as GENI [59,146], a long-held goal of Internet research has been to develop a network architecture that is flexible. An unfortunate stumbling block in these efforts has been that flexibility is often at odds with efficiency and

security, another long-held goal of Internet research. For example, several schemes seek the flexibility of forwarding packets along multiple paths to improve performance and reliability, while schemes such as network capabilities seek to constrain packets to a single "approved" network path. We will investigate ways to provide security in the face of unknown policy requirements of the future, while still achieving performance and reliability.

The key aspect of our *policy architecture* approach is to build mechanisms that one can compose to express all reasonable transit policies. Two important questions arise: (1) What are all reasonable transit policies? and (2) What mechanisms would enforce them?

With regard to question (1), we hypothesize, based on preliminary investigations, that the following three factors form a "minimal spanning set" of all current transit policy projects. All the factors involve an entity along the path of a communication (sender, provider, middlebox, edge network, receiver) deciding whether a flow is authorized: (a) How the entity would dispose of the packet internally (e.g., what priority would it receive, what local middleboxes would it travel through, what traffic type is it, etc.); (b) Which other entities are along the path of the communication; and (c) What other information, not to do with the characteristics of the flow, is available at flow set-up time. To test whether our hypothesis is correct, we need to develop a precise formalism for describing policy proposals. With regard to question (2), we need a mechanism that stays fixed even while the *policy* function evolves. That is, we need a way to run an *arbitrary* control plane and receive guarantees from a *fixed* data plane. The required research here is to develop a data plane interface. One approach is for a packet to travel with the equivalent of explicit MPLS labels. When a packet arrives at an intermediate entity, the entity can check whether the control plane authorized the label, and then map the label to a required internal action. A label mechanism will allow substantial new functionality, from allowing users to push "turbo" buttons on Web sites to request better service in the core (for a fee), to assigning a set of end-hosts an isolated sub-network within a given provider.

A comprehensive policy architecture in which policies can be *enforced* creates a foundation for security, a point we expand on (in Section 4.2), after proposing NDP.

4.1 What is NDP?

NDP is a network protocol in which packets will contain the following four elements per administrative domain in the packet's path:

(1) a domain identifier;
(2) a proof, called a PoC, that the administrative domain has authorized the path;
(3) a proof, called a PoP, that the packet has followed that path; and
(4) an MPLS-style token.

This token serves as a hook with which to bind approved communications to policy-dependent data-plane functions. This token can map to RBF-style [150] rules, providing, at one fell swoop, all of the flexibility and functions of the RBF project. This token can also express queuing priority, restrict intra-domain routing, mandate middleboxes or traffic shaping, or be used to trigger unanticipated future data plane features.

Though this is somewhat surprising, it turns out, per our preliminary investigations, that these four elements are sufficient not only for networked entities to express their policies about packet carriage but also for those entities to *enforce* those policies. The core reason is that, when a packet arrives at an administrative domain, that domain has all of the information that it needs to decide whether to devote its internal resources to the packet, namely: was the packet authorized? (check the PoC.) What internal resources would the packet consume, and which middleboxes should it travel through? (check the token.) Did the packet actually take the authorized path? (check the PoP.)

Preliminary experiments and prototypes [136,137,138,150,161] have demonstrated that this architecture is feasible, both in terms of packet space and data plane processing cost, e.g., by efficient representation in the packet, by aggressively caching at connection setup, and by leveraging the increasing computational power of specialized and general-purpose processing on router line cards.

Despite this feasibility, the architecture's flexibility does carry a penalty relative to the status quo. This

penalty is principally restricted to the data plane. (In the control plane, the architecture pays only in proportion to the control that is exercised. For example, if all of the entities abdicate their fine-grained control, then the control plane reduces to the status quo.)

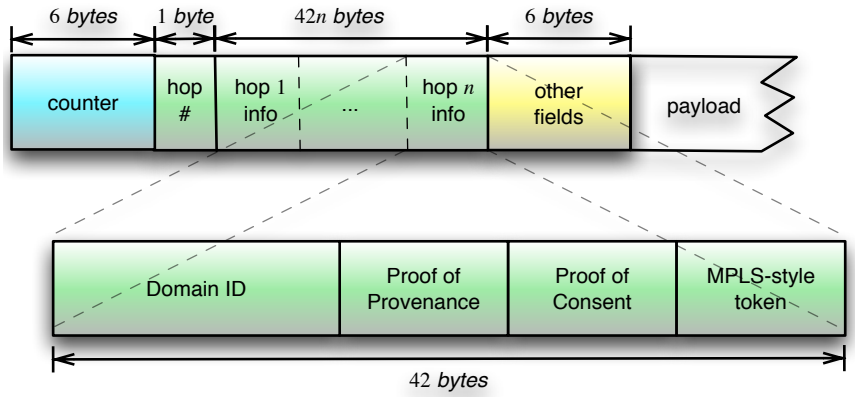

Figure 1: NDP packet format

The data plane penalty is, we believe, the necessary price of moving to an architecture that upholds all stakeholders' legitimate interests: stakeholders whose interests are left out of the architecture will fight the adoption of the architecture, rather than fight within it. To quantify the penalty, we estimate that NDP packets as illustrated in Figure 1 would, on average, have 20% larger packets than in the status quo and would require 50% more logic area in routers [136]. Part of our research is to reduce these numbers, which we already have experience doing: the current estimated overhead is an order of magnitude lower than what a naive design would cost.

4.2 What properties does NDP uphold?

We posit that the above building blocks, when composed in various ways, can subsume the policy goals of a very large number of other projects. To see why, note that it captures the functions of both ICING [161], and RBF [150], and as argued separately in those papers, each of those mechanisms individually subsumes several dozen other projects, as well as enabling completely new network functions. In Tables 1-4, we categorize the functionality enabled by various projects in terms of security policy, path selection, middlebox processing, attack resilience, and control/data plane alignment. In each case, NDP provides a superset of the union of the features provided by other projects. Although we cannot yet show that NDP is universal, since NDP's four primitives capture 50-60 other projects, we argue that we have a promising candidate for a set of *fundamental* primitives. Moreover, in contrast to much of this prior work, NDP can actually *enforce* its policy goals -- even under very strong threat models.

Specifically, NDP provides the following properties, which we argue are required of any network architecture aiming to be secure:

Assured paths: as mentioned above, for communication to happen, all of the entities along the path must approve of the entire path (if they wish; an entity can delegate or abdicate its control, though showing how requires more detail than we are able to present here). This property generalizes the point properties of prior work, such as a receiver approving of a sender, a sender controlling the downstream path, or a provider controlling its prior hops and downstream paths.

Controlled access: the converse of the property above is that if a path is *not* approved, packets will not flow. Since the path includes the destination and potentially a service identifier (namely the destination's token), the architecture neatly implements access control, whether it's which clients should access which Web services, or which geographic areas should have access to which remote data centers. This function is sometimes called "pushing firewalls into the network".

Approach	dest can constrain sender	resource attribution	provider policy granularity			src can constrain routes	MB can constrain routes
			prefix	Suffix	subsequence		
BGP				X			
Capabilities [190,193]	x						
Filters [24,46,57,82,88, 116,125,186,192]	x						
Intserv, RSVP [33,34]	x	x					
Visas [61]		x					
Platypus [154]		x				x	
LSRR [14]	x					x	
Policy routing, Nimrod [42,50]					x	x	
Pathlets [71]				X	x	x	
Wiser [124]			x				
MIRO [188]				X			
Src. routing [60,77,99,194,196]						x	
Byzantine routing [144,145]						x	
NUTSS [76]							x
i3, DOA [171,179]							
DONA [107]							
Active Networks [175,176]							
NDP	**x**	**x**	**x**	**X**	**x**	**x**	**x**

Table 1: *Security related policy controls available in many, but not all, network-layer projects (MB is middlebox). Each of the listed controls (columns) can be viewed as an entity constraining some portion of the path of the communication. Each control represents a legitimate policy interest of some stakeholder; only NDP's framework makes available all of the controls in the columns.*

Availability: because path selection takes place outside of the data plane (in contrast to the status quo, where paths are revealed hop-by-hop), there is ample opportunity for end-points to negotiate multiple paths between them and, if a path fails, to use a backup path. We estimate that this process would

be far faster than the time it takes BGP to calculate new routes [109].

Autonomous control of resources: No entity is forced to dispose of its resources in a manner it disapproves of. This is a fundamental building block for security; it ensures that no entities' transit policies are ever violated.

Privacy-enhanced communication: for our purposes here, privacy consists, first, of keeping secret the *content* of a communication, and second, keeping secret the *fact* of that communication. The former is a concern of the layers above the network layer. The latter is squarely a consideration of the network layer, and NDP supports it by giving two communicating entities control over how their communication travels. They can route the communication through providers they trust (just as businesses in the analog world choose their couriers for important documents). More exotic options here are for the endpoints to specify an onion routing system or to specify that their communication take place along an isolated, always utilized channel so that no other communications can infer the existence of the private one.

We note that without the four primitives mentioned earlier, the architecture would not be able to provide the above properties (it would be able to provide subsets and point solutions, but not all of them, together). As an example, without packet provenance, Internet2 cannot enforce a policy such as, "All traffic we carry must originate and terminate at a university". Or, a provider may wish to have a policy like, "All traffic I carry has been vetted by this off-site scrubbing service". On the other hand, these primitives, particularly packet provenance, require careful design to work correctly. In fact, our experience has been that unless an architecture is designed from the ground up to achieve these functions, it will be unable to provide them robustly if later on that proves to be essential.

At the same time, more work here is required. While we have a proposed design that achieves a number of novel properties (such as allowing a networked entity to verify that the packet has taken the path that the packet claims to have taken), we also want to address a number of other issues, which requires research. Two of these questions are: (1) How can the source of a communication prevent a given carrier from transparently subcontracting (e.g., handing the packet off) to a another provider? (As proposed, NDP can enforce that an ISP authenticated and approved a packet but cannot ensure the ISP's failure independence or prevent the ISP from disclosing the communication to others.)

Approach	src. can invoke MB	rcvr. can invoke MB	provider can invoke MB	src. or rcvr. mobility	src./rcvr. can use router state in forwarding	rcvr. anycast	src./rcvr. can invoke router extensions	rcvr. can record router state
Active Networks [175,176]	X				x	x	x	x
ESP [41]					x		x	x
i3, DOA [171,179]	x	x		x		x		
Platypus, SNAPP [142,154]	x							
NUTSS [76]		x	x					
Src routing [60,77,99, 194,196]	x							
DONA [107]			x			x		

NDP	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**

Table 2: *Flexibility-related policy controls available in many network-layer projects (MB is middlebox). These controls (generally for end-hosts) provide flexibility in path selection, use of in-network functionality, mobility and recording path information.*

(2) How can a networked entity along the path of a communication verify that the other entities along the communication are giving contracted service to the communication? (Note: although much research has focused on failure localization, our context poses new problems.)

Proposed defense	DDoS	source spoofing	forged routing advts	router thwarts src routing	router thwarts hop-by-hop routing
Self-certifying addresses [9,195]	x	X	x		
Capabilities, filters [15,24,46,57,82,88,116,125, 186,189,190,192,193,199]	x				
Charging resources [66,143,170,182,183]	x				
Source authentication [35,93,112,115,191]		X			
Probes, secure traceroute, auditing [16,17,21,23,25,72,139,159,185,198,200]					x
BGP security [1,87,101,172]			x		
Authenticated routing [58,85,86,140,141,197]		X	x		
Byzantine routing [19,20,132,144,145]		X		x	
Zodiac [45]	x	X	x		
RBF [150]	x				
NDP	**x**	**X**	**x**	**x**	**x**

Table 3: *Some attacks addressed by some prior network-layer work. While this table is incomplete at the network layer (e.g., it leaves out firewalls), a key point is that many of the listed works cannot be implemented together; NDP aims to address all of these attacks.*

Mechanism	all participants can deny based on path	comm. held to described path	malicious behavior tolerated	decentralized	fixed and feasible data plane
IP+BGP (the status quo)				x	x
Ethane		x	x		x
Auditing [198]			x		x
MPLS, virtual circuits, resource reservation [22,34,157]	x			x	x
Capabilities, Platypus [154,190,193]				x	x
Passport [115]			x	x	x
Byzantine routing [144,145]		x	x		
Secure routing [19,20,132]		x	x	see caption	x
Secure policy routing [62]	x		x	see caption	
PoMo Architecture [27,40]	x	x	x		
NDP	**x**	**x**	**x**	**x**	**x**

Table 4: *Prior approaches to aligning control and data planes, in terms of requirements. For MPLS, two entities can collude to skip a third, and it lacks cryptographic assurance to provide proof that a packet is following its approved path. Secure routing and secure policy routing don't require a PKI, but do require prior coordination and pre-configuration among the hops, thus not fully meeting our decentralized requirement.*

5.0 Research Agenda: NEBULA Virtual and Extensible Networking Techniques (NVENT)

The existing Internet is predominantly *enterprise-centric*, with an assumption that various organizations each run servers, and communication occurs between individual computers that serve as endpoints [55,158]. In contrast, the cloud is *service-centric* and *data-centric*: computational and data services can be provided redundantly across multiple data centers [69] with duplication selected to increase reliability or performance. The cloud allows many evolutions, such as those focused on content [92]. NEBULA is a network architecture with evolutionary advantages: it is easy to extend NEBULA (at the edge) while providing a new core (NCore) within which highly available services (or portions of highly available services) can be located. This locus for data and availability addresses the availability challenges for the computer utility [18] while preserving the ability to innovate at the edge. Flexibility at both edge and core are preserved through the interface presented by NVENT, which both provides a locus at which edge systems can discover paths they require and discover new network services with query-able attributes. NVENT discovers new services as they are made available on routers; this facility can be used to evolve network services as they are developed [165].

5.1 Distributed Services

One aspect of our research [69] is better network support for mobile users and distributed services, by (1) moving from human-readable host names to machine-readable service identifiers, (2) moving from individual packets to flows, and (3) moving from unicast communication to anycast. Our approach to mobility hides network addresses from applications to enable dynamic remapping as end-points change, (e.g., due to virtual-machine migration, failover, or device mobility); directs traffic based on successively refined identifiers [69] to scale routing and limit churn; and more tightly integrates service end-points and network elements for better scalability and responsiveness to change. Although this implies a new service architecture, its benefits can be realized through an incremental deployment.

Moreover, a service instance may be hosted across multiple machines (sometimes referred to as "shards"). Highly reliable intra-domain and inter-domain routing protocol are required; these protocols must reflect real-world commercial constraints while ensuring that traffic is delivered whenever there is a policy compliant route from source to destination. Since a route is useless without resources to back it up, we further need to change the nature of Internet resource discovery and resource allocation to ensure packets are delivered even when adversaries are attempting to block access through denial of service or route hijacking. While NDP specifies the mechanism for the data plane, and NVENT specifies the policy framework for the control plane, we also need consistent distributed state management with rapid failure recovery at the interdomain and intradomain level to achieve trustworthy and reliable operation. As a general rule, cloud applications [70] seek to provide the *appropriate level of consistency required by the application* because doing so allows for higher scalability, reliability in the presence of network partition, and increased performance across the global Internet.

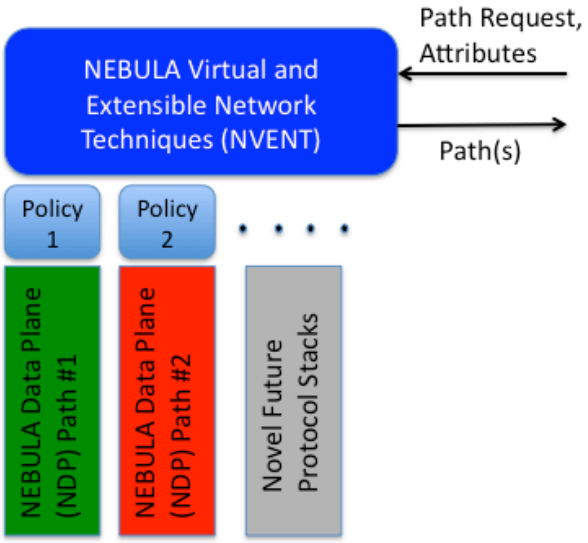

Figure 2: NVENT path selection is used for policy and the interface for extensions

The NVENT service interface would allow an application or access provider to request a service and specify the level of availability required. For example, an access ISP that offers emergency services can request high availability that is provided with multi-path interdomain routing. The key is flexibility: no single version of service properties needs to be chosen because each service can request transit that is appropriate (and presumably pay a corresponding fee).

We envision a service interface that uses a distributed resolution service to supply information about each service, the method(s) used to access the service, and the properties of the service. The design of this global-scale distributed resolution service is part of future research, as we might want scalability, flexibility, and dynamism beyond what the DNS or BGP systems offer today. The NVENT resolution service might be populated by information advertised by data centers at the core of the network, and will be interrogated by ISPs (who will use it to request communication services) and indirectly by users (who will request application services). For example, a medical doctor might use the resolution service to find a named service that provides access to patient records, and an access provider might use the service

system to find/create a HIPAA-compliant encrypted path to the nearest data center that offered the requested service using NDP paths.

5.2 What is NVENT's interface to NDP?

NVENT's job is to determine appropriate values for the packet elements described in Section 4. Specifically, NVENT is responsible for determining packet paths, gathering the approval of all intermediate domains (the PoCs); and learning which tokens should be in the path. In the general case, prospective senders would query NVENT servers to gather this information and place it in packets. In normal operation, however, senders would continue sending packets as they do today, and proxies and gateways would transform their legacy traffic into NVENT queries followed by NDP packets. Now, when NDP packets enter the network, the domains along the path would have the needed information to perform the checks above.

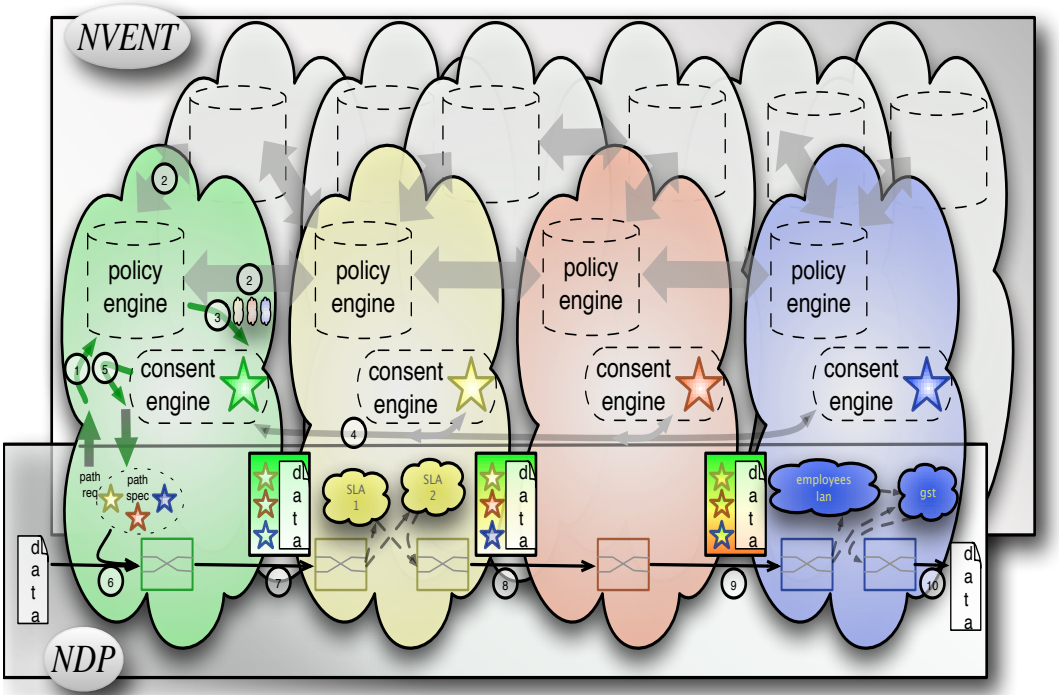

Figure 3: Overview of flow establishment in NVENT and data forwarding in NDP

Figure 3 provides considerably more detail on the interactions:

1. NDP requests a path from NVENT. The desired attributes of the path (destination, preferred transit policy) are conveyed via a connection descriptor (not shown).

2. NVENT performs path discovery based on a pluggable policy engine, e.g., a BGP-like link-state protocol that propagates topology and transit policy information asynchronously. (Background communication is denoted by the wide gray arrows.) The result of this step is a list of domains through which to route (green, yellow, red, and blue, in the example), along with associated MPLS-style tokens (not shown) to evoke dataplane-specific functions during data forwarding (cf. steps 8-10 below).

3-4. Once found, a suitable path is processed via NVENT's consent engine to obtain an assured path, that is, a path that is amenable to enforcement by the dataplane. This process amounts to gathering and/or deriving cryptographic proofs of consent to carry traffic on the part of the domains listed in the path. (Proofs of consent, or PoCs, are denoted by color-coded stars.)

5. The assured path is returned to NDP.

6-7. The originating NDP router wraps the data into NDP packets for the assured path that was obtained from NVENT. PoCs are fused into the packet to thwart stealing. (Depicted as background

green color in 7.)

8-10. NDP routers at each independent domain check the cryptographic values in the packet, and process the packet according to the token included for the local domain. In the example, the second (yellow) domain offers two different levels of service: incoming packets are forwarded either to the "SLA1", or to the "SLA2" subnet, or to both (dispersity routing) depending on the token they carry. Similarly, the fourth (blue) domain (a company, say) might provide access to part of a local network to its employees, but not to guests. Besides checking the cryptographic PoCs and honoring the tokens, intermediary routers contribute to the enforcement of the assured path by stamping packets with Proofs of Provenances (PoPs, depicted as color bands in the packet's background).

At a high level, the data plane (NDP) exposes a narrow interface (just domain IDs, PoCs, tokens, and PoPs), and pushes the policy and routing complexity to NVENT. Being implemented on general-purpose commodity servers, NVENT can rapidly evolve, while the specialized data-plane hardware, and the interface to it, remains constant.

5.3 Accountability

We will investigate *accountability* as a way to increase the resilience of the new architecture against faults and misbehavior. Such faults can occur for a variety of reasons, ranging from accidental mis-configurations to rational manipulation and even deliberate attacks. Accountability can ensure that a large class of faults and misbehaviors can be detected. This enables network administrators and service operators to quickly respond to faults, even in cases when the system is unable to prevent them or to mask their effects. Accountability can also produce evidence that irrefutably links each fault to a specific component or a specific domain. This enables domains to hold each other responsible for faults, and thus creates an additional incentive for each domain to make its infrastructure as reliable as possible.

Prior work [78,79,80,81] has developed techniques that can enforce accountability for distributed systems, but these techniques focus on faults that occur on the nodes and assume that the network itself does not fail. For NEBULA, we will develop new techniques that can apply accountability to primitives provided directly by the network. We will also investigate ways to combine accountability with confidentiality: ideally, each domain should be able to release enough information to enable fault detection without compromising any sensitive information, such as its routing policies or its internal topology.

5.4 NVENT Control Strategy

To simplify user control of this functionality, we will investigate use of *declarative networking* as a configuration framework for NVENT. Declarative networking is a programming method that enables developers to concisely specify network protocols and services, which are directly compiled to a data framework that executes the specifications. We plan to build upon the *Network Datalog* (NDLog) declarative networking language to develop: (1) a language to allow users to efficiently describe and construct flexible network services and NDP packet rules; and (2) an efficient compiler that translates this language into low-level instructions for the network (e.g., configurations in OpenFlow switches), and that coordinates actions of the network and server infrastructure to achieve a common unified goal. As an initial proof of concept MOSAIC [127] was developed as a declarative platform for composing new overlay networks from existing ones by specifying high level functionalities to be composed. Extending MOSAIC's composition capabilities to more complex network services, support virtualization of the network layer, and leveraging NDP, e.g., hooking in to be a rule, is an interesting avenue of research that we plan to explore.

5.5 An NVENT Prototype Implementation

To enable an extensible policy engine for NDP, we plan to leverage DS2 (Declarative Secure Distributed Systems) [56], a unified declarative platform for specifying, implementing, and analyzing secure extensible distributed systems. DS2 will be used for specifying and analyzing NDP security policies at Internet-scale. DS2 unifies declarative networking and security specifications into a new language called

Secure Network Datalog (SeNDlog). We have used DS2 as a platform for implementing a variety of secure network routing protocols [201], extensible anonymity [163], and secure distributed data management applicable to cloud computing environments [129].

One prototype of NVENT and NDP will be implemented using the RapidNet declarative networking system [134,135,155]. One of the interesting opportunities presented by integrating RapidNet with NDP is the opportunity to perform a variety of analysis and verification security policies at runtime and prior to deployment. For example, the dataflow framework used in declarative networking captures information flow as distributed queries. Hence, it is natural to utilize *data provenance* to explain the existence of any network state, which is analogous to the use of proof trees in security audits. This leads to the notion of *network provenance* [200,202], for which runtime analysis and debugging of network protocols, network forensics, and the enforcement of complex trust management policies have been developed in DS2. An interesting possibility for NEBULA is applying the Formally Verifiable Networking described in Section 6.1 to SeNDlog, as a means of verifying security properties of network protocols.

6.0 Research Agenda: NEBULA Core (NCore)

The NEBULA Core will be built on a future generation of core routers that can support the highest transport speeds at any given time, while providing always-on availability. This latter requirement demands viewing future generation router control plane software as a fault-tolerant distributed system [38]. Geography, latency [167] and federation stand in the way of viewing a collection of these systems as a single logical router, but interestingly, many of the same reliability issues affect both the provision of network services by NVENT and the internals of a router; thus, new reliability algorithms from both NVENT and NCore routers may be deployable in both contexts with appropriate modifications.

6.1 High Availability Core Routers

Because a single CPU is incapable of forwarding data at rates sufficient for tier-1 ISPs, the next generation of high-end routers will use a distributed approach. In the future, a router will consist of multiple chassis (Cisco plans to scale routers to include up to 48 chassis in the near term), each of which has multiple line cards, multiple processors for forwarding, and multiple control processors. Pieces of the router are tied together with a high-speed switching fabric, which means a single core router functions internally as a large distributed system.

It is important to recognize that the industry push towards scalable routers built out of smaller components is not only a technical or manufacturing issue (e.g., how to achieve better scalability at lower cost through higher volume components and parallel internal links). More importantly, it is being driven by the security and reliability demands of future cloud applications. Existing Internet protocols, such as BGP or OSPF, for managing the interactions between routers are *too weak* to accomplish the precise semantics, fast failover, hot software and hardware upgrade, continuous operation, and multi-version support that we see for applications in the rest of the data center. In our view, the abstraction presented by the router should be an ultra-reliable, ultra-secure, scalable, self-managing device that can be extended to meet any practical workload asked of it.

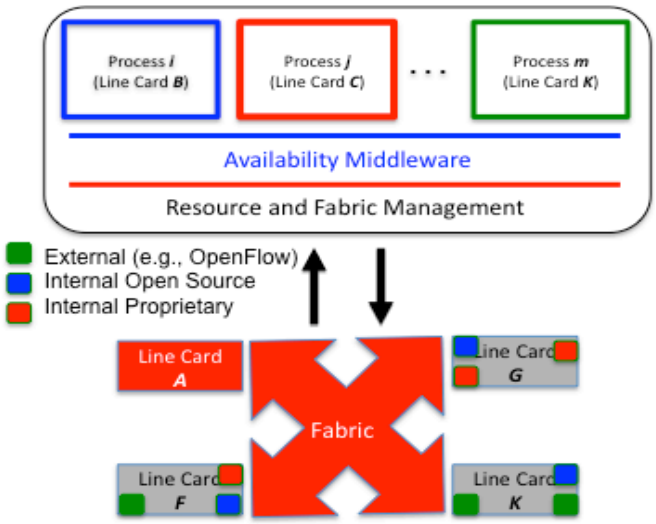

Figure 4: Hardware/software architecture of a future core router shows distributed control and how multiple services can be supported concurrently

Network services demand high availability and consistent response to events, including routing updates, management commands, and requests for services. To support a distributed security architecture and to enable a trustworthy core, the entire router must make atomic updates and insure that only authorized principals can access protected data. We define *consistency* to encompass such properties, and assert that one of the key challenges for a future Internet will lie in providing such consistency in the face of failures and dynamic reprovisioning while continuing to forward traffic at full data rate. For example, how can distributed systems techniques, such as redundancy and voting, improve the availability of the overall system? How can we guarantee consistency of the forwarding information base across all line cards? How can we guarantee that routing converges to a valid state on all line cards? Can an automated monitoring system be constructed that detects anomalous behavior in such a distributed system?

We will re-architect the internal architecture of core routers to reflect their scalable hardware components. That is, we will make routers that fail because of (data center) power outages, and not because of hardware or software upgrades, or software crashes. To build such a distributed, redundant, core routing system, we must solve three key problems:

1. Building router software with strong consistency properties: The additional functionality offloaded into the network coupled with ever-increasing demands on uptime and scalability of routers requires a new approach to designing and building router software. The new router control software must solve three problems: it must operate across a single router that is itself distributed, it must enable live upgrades, and reliability and security must compose across a set of routers distributed in the wide area, all while retaining strong consistency properties on its operation. We will address these challenges as follows.

To operate within a single distributed router reliably, we will design and implement a new router software stack based on the new Dynamic Reconfigurable Service (DRS) model [32]. DRS unifies the virtual synchrony model with the Paxos/State Machine Replication model in wide use within cloud computing systems. This will form a new and scalable foundation for network security services and consistent replication mechanisms. DRS could run at extremely high speeds, yet would also be formally verifiable using formal theorem provers such as NuPRL to reason about protocols and to support end-user application development. A new "tools layer" will support higher levels of the network stack and end user applications, running across core network routers and in the data center, providing a range of distributed systems consistency models. Examples include the new DRS model, weaker convergence properties for applications that can tolerate relaxed consistency, stronger Byzantine properties for applications with third party plug-ins, etc. The lesson from cloud computing is that engineering large scale distributed

systems is tractable only with simple layering, with precise semantics, tuned to the demands of the higher levels of the software stack. To our knowledge, this has not been done before at the level of a router.

To handle live updates/patches reliably, we will investigate software version transition and validation: because traffic never stops at the core of the Internet, core routers must operate continuously – the router cannot be taken offline during route changes and the router cannot be powered down. An important question arises about the control plane: how can new versions of control software be installed without jeopardizing the continuous and correct operation of the router? In particular, can a new version be tested under load before it is used in production? The question is further complicated because a large core router may need to run multiple versions of control software in production at the same time (because, for example, an ISP may communicate with each of its neighbors using a different version of the interdomain control protocol).

2. Dealing with implementation errors in router software. Some of the most complex aspects of Internet technology arise in the software running on routers and servers. The software is highly complex, with both modern router and server implementations comprising millions of lines of code. Introducing additional functionality in the network, coupled with the additional flexibility provided by our network (which may enable third parties to dynamically download new code) introduces potential for vulnerabilities, software errors, and mis-configurations. We will attempt to eliminate most errors by developing infrastructures for *verifiable* network software. To build verifiable network software, we will investigate *Formally Verifiable Networking* (FVN) [181], a formal methodology towards verifying the properties of network protocols deployed on NEBULA. FVN is a novel approach towards unifying the design, specification, implementation, and verification of networking protocols with a logic-based framework. In FVN, formal logical statements are used to specify the behavior, and the properties of the protocol. FVN then uses *declarative networking* [117,118,119] to move from high-level logical specifications of the network model to low-level properties of network protocols. A theorem prover [180] is used to statically verify the specified properties of the declarative network protocols. Moreover, a property preserving translation exists for generating declarative networking implementations from verified formal specifications. For instance, using meta-routing [74] as our driving example, we demonstrate the possibility of using FVN to design and specify network models in a systematic and compositional way with correctness guarantees.

3. Increasing reliability through monitoring. Designing high-performance control-plane router software is a highly challenging task. We will investigate *increasing reliability through monitoring*: self-checking systems that take a global view of a distributed router and measure whether operations and performance are within bounds. The work will focus on outcomes, not on the operation of individual elements. To pursue our approach, we assume that additional control processors are available (an assumption derived from discussions with our Cisco partner), and use the additional processors to monitor the router. The self-checking system will have a global view of both hardware and software. It will be able to inspect routing packets as well as internal data structures, such as forwarding tables. The monitor will be able to exercise both control and data paths. For example, a monitoring processor on one line will be able to insert a time-stamped packet that a monitoring processor on another line card can receive. Therefore, the monitoring system will be able to check both forwarding paths and latency across the interior of the router.

6.2 New Local and Wide Area Interconnects

We will create a hardware and software architecture that interconnects a data center directly to a large core router. We will work with our corporate partners, Cisco and Intel, to explore parallel connections [177] between data center switches and core routers that can provide both high speed and reliability. This will address the mismatch between highly connected meshes of storage and computing within the data center and the WAN links in NEBULA's NCore. We will investigate addressing and routing for such an interconnect. More important, we will investigate ways that multi-path routing [26,130] and fast fail-over can be employ to guarantee virtually uninterrupted, load-balanced service despite the failure of one or

more of the redundant paths.

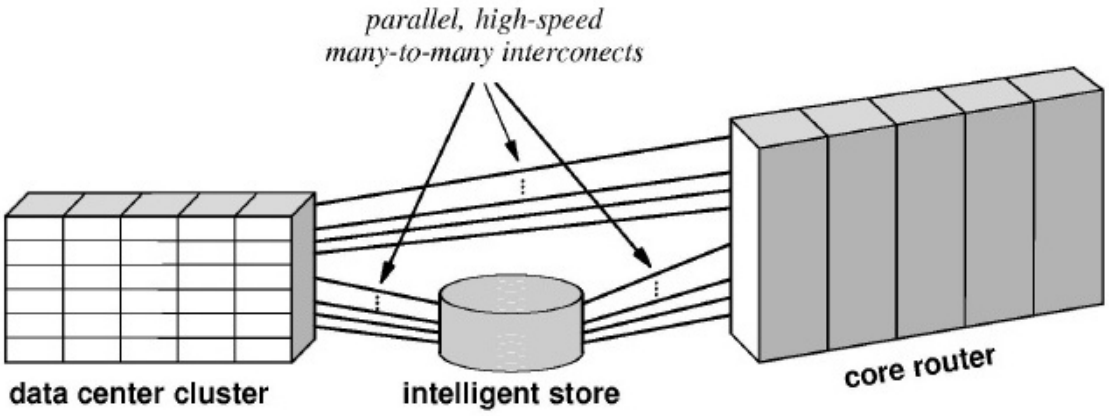

Figure 5: Parallel Many-to-Many Interconnects Among Clusters and Core Routers

To improve performance of high-bandwidth transport, we will work on the simultaneous use of multiple diverse paths between sources and destinations. Path diversity [8,26,130] allows for graceful degradation in the face of link and router disruptions and provides the security property that the adversary must monitor or disrupt all paths to capture or interfere with end-to-end communication. The approach in NEBULA will be to construct a new path diversity routing protocol to be coupled with the NDP mechanisms and policy engines. A set of diverse paths would be constructed through network nodes that accept participation in a path. The diverse paths would then transport packets according to their agreed capacities, giving an aggregate capacity for the set of paths.

7.0 Research Agenda: Economically Viable Path to Deployment

The proposed research will focus on implementing a high speed, trustworthy[29] architecture for core "cloud computing" infrastructure. Successful deployment of a new architecture raises several economic questions (for example, about industry/market structure and business models) and policy issues (for example, regulatory management). While we intend to design an architecture that is flexible with respect to market/policy assumptions, we also intend to design an architecture that we believe represents a plausible trajectory for deployment. We expect that the most likely vector for commercialization of this architecture would be deployment by backbone ISPs using the highly-reliable routing infrastructure, collocated with data center computation and storage facilities. To be economically viable, the architecture will need to accommodate a future with multiple cloud providers that are interconnected and offering services to a multiplicity of access ISPs and their end-users (which may be less trustworthy and reliable than the cloud resources).

From a capabilities/constraints perspective, we anticipate that the new architecture will support a significant improvement in speed and reliability of core routing and data center access, with corollary constraints on power density and system costs. One role for economics will be to clearly articulate the gap between current commercial capabilities and desired performance goals, and help design a roadmap for how the gap may be closed. The analysis will help set the stage for evaluating the architecture's viability. We will ask: are the reliability/performance improvements commercially achievable and, if so, over what time frame? What level of investment is required? Where are the key gaps or biggest changes from today?

Given a clear articulation of how the proposed architecture will change current capabilities and industry economics, we will consider the impact on, and the incentives of, key stakeholders. We will also consider benchmark issues, such as:

- *Optimal firm/market structure for architecture adoption*: we anticipate that the need to meet reliability, security, and performance goals will entail extending the management of core ISP "cloud" capabilities into access ISPs. It may call for new types of third-party entities and may call for further

evolution of ISP vertical and horizontal business organization/relationships. We will examine the implications for locating aspects of the core functionality and control (decision-making) points for firm (ISP) and market boundaries, and the implications on market structure and regulatory policy.

- *Implications for ISP interconnection*: The architecture suggests two potential types of interconnection between ISPs: peering between cloud service providers and transit between cloud resources and access networks. A third form of interconnection relates to how resources within a single cloud communicate with each other (and addresses challenges of reliability when a distributed set of data centers are designed to act as a single, unified center). We will map the requirements of our proposed interconnection architecture to existing interconnection practices, and will consider the implications for competition and the regulation of interconnection (open access policies) within our three-tiered model.
- *Risk management*: the enhanced security/reliability model we wish to support poses new challenges for industry structure and policy. Our ideal is for users of cloud resources to act as if these are effectively 100% reliable, when in fact we recognize that that goal is only asymptotically achievable technically. We propose to analyze our architecture to assess its robustness and compatibility with non-technical contractual and liability management (insurance) mechanisms. This will include special analysis of catastrophic failure scenarios.

The economic and policy analysis required does not call for the development of new theory or techniques, but rather the careful application of established tools of institutional and industrial economics, especially as they have been applied to regulated and networked industries.

8.0 Summary

NEBULA is a new Internet architecture based on a high-performance highly-available core network, a novel data plane protocol that incorporates fundamental primitives required for access control, and a new distributed control plane architecture which provides an interface with which network resources can be allocated. We have outlined the research required to bring NEBULA to fruition, including technologies ranging from parallel interconnects to high availability software control planes for core routers, and also including economic and regulatory expertise to ensure viability of the architecture.

References

[1] William Aiello, John Ioannidis, and Patrick McDaniel. “Origin authentication in interdomain routing,” In *Proc. ACM Conference on Computer and Communications Security (CCS)*, October 2003.

[2] D. Scott Alexander, Marianne Shaw, Scott M. Nettles, and Jonathan M. Smith, ‘‘Active Bridging,’’ in *Proceedings, ACM SIGCOMM Conference*, Cannes, FR (October 1997), pp. 101-111.

[3] D. S. Alexander, W. A. Arbaugh, A. D. Keromytis, and J. M. Smith, ‘‘A Secure Active Network Environment Architecture: Realization in SwitchWare,’’ *IEEE Network Magazine, special issue on Active and Programmable Networks* **12**(3), pp. 37-45 (May/June 1998).

[4] D. S. Alexander, W. A. Arbaugh, M. W. Hicks, P. Kakkar, A. D. Keromytis, J. Moore, C. A. Gunter, S. M. Nettles, and J. M. Smith, ‘‘The SwitchWare Active Network Architecture,’’ *IEEE Network Magazine, special issue on Active and Programmable Networks* **12**(3), pp. 29-36 (May/June 1998).

[5] D. S. Alexander, W. A. Arbaugh, A. D. Keromytis, and J. M. Smith, ‘‘Safety and Security of Programmable Network Infrastructures,’’ *IEEE Communications Magazine* **36**(10), pp. 84-92 (October 1998).

[6] D. Scott Alexander, Paul B. Menage, Angelos D. Keromytis, William A. Arbaugh, Kostas G. Anagnostakis, and Jonathan M. Smith, ‘‘The Price of Safety in an Active Network,’’ *Journal of*

Communications and Networks **3**(1), pp. 5-18 (March 2001).

[7] Scott Alexander, Steve Bellovin, Alice Cheng, Brian Coan, Andrei Ghetie, Vikram Kaul, Nicholas F. Maxemchuk, Henning Schulzrinne, Steven Schwab, Bruce Siegell, Angelos Stavrou, and Jonathan M. Smith, ''The Dynamic Community of Interest and its Realization in ZODIAC,'' *IEEE Communications Magazine (Special Issue on Military Communications)* **47**(10), pp. 40-47 (October 2009).

[8] David G. Andersen, Hari Balakrishnan, M. Frans Kaashoek and Robert Morris, "Resilient Overlay Networks", in *Proc. SOSP*, 2001.

[9] David Andersen, Hari Balakrishnan, Nick Feamster, Teemu Koponen, Daekyeong Moon, and Scott Shenker. Accountable Internet protocol. In Proc. ACM SIGCOMM, August 2008.

[10] T. Anderson, A. Collins, A. Krishnamurthy and J. Zahorjan, "PCP: Efficient Endpoint Congestion Control", in *Proc. of Networked Systems Design and Implementation*, 2006.

[11] William A. Arbaugh, David J. Farber, and Jonathan M. Smith, ''A Secure and Reliable Bootstrap Architecture,'' in *IEEE Security and Privacy Conference*, Oakland, CA (May, 1997), pp. 65-71.

[12] W. A. Arbaugh, A. D. Keromytis, D. J. Farber, and J. M. Smith, ''Automated Recovery in a Secure Bootstrap Process,'' in *Internet Society 1998 Symposium on Network and Distributed System Security*, San Diego, CA (1998), pp. 155-167.

[13] William A. Arbaugh, James R. Davin, David J. Farber, and Jonathan M. Smith, ''Security for Virtual Private Intranets,'' *IEEE Computer (Special Issue on Broadband Networking Security)* **31**(9), pp. 48-55 (September 1998).

[14] K. Argyraki and D. R. Cheriton. Loose source routing as a mechanism for traffic policies. In *Proc. SIGCOMM Workshop on Future Directions in Network Architecture*, September 2004.

[15] K. Argyraki and D.R. Cheriton. Active internet traffic filtering: Real-time response to denial-of-service attacks. In *USENIX Technical Conference*, April 2005.

[16] K. Argyraki, P. Maniatis, D. Cheriton, and S. Shenker. "Providing packet obituaries", In *Proc. ACM Workshop on Hot Topics in Networks (HotNets),* October 2004.

[17] Katerina Argyraki, Petros Maniatis, Olga Irzak, Subramanian Ashish, and Scott Shenker. Loss and delay accountability for the Internet. In *Proc. IEEE International Conference on Network Protocols (ICNP)*, October 2007.

[18] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica and Matei Zaharia, "Above the Clouds: A Berkeley View of Cloud Computing", *Technical Report No. UCB/EECS-2009-28*, Electrical Engineering and Computer Sciences, University of California at Berkeley, February 10, 2009.

[19] I. Avramopoulos, H. Kobayashi, R. Wang, and A. Krishnamurthy. Amendment to: Highly secure and efficient routing. February 2004. http://www.cs.washington.edu/homes/arvind/papers/amendment.pdf.

[20] I. Avramopoulos, H. Kobayashi, R. Wang, and A. Krishnamurthy. Highly secure and efficient routing. In *INFOCOM*, March 2004.

[21] Ioannis Avramopoulos and Jennifer Rexford. Efficient data-plane security for IP routing. In *USENIX Technical Conference*, June 2006.

[22] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE: Extensions to RSVP for LSP tunnels. RFC 3209, Network Working Group, December 2001.

[23] Baruch Awerbuch, David Holmer, Cristina Nita-Rotaru, and Herbert Rubens. An on-demand secure routing protocol resilient to byzantine failures. In *Proc. ACM Workshop on Wireless Security (WiSE)*, September 2002.

[24] Hitesh Ballani, Yatin Chawathe, Sylvia Ratnasamy, Timothy Roscoe, and Scott Shenker. "Off by default!," In *Proc. ACM Workshop on Hot Topics in Networks (HotNets),* November 2005.

[25] Boaz Barak, Sharon Goldberg, and David Xiao. Protocols and lower bounds for failure localization in the Internet. In *Proc. EUROCRYPT*, April 2008.

[26] Paul Baran, "On Distributed Communications: IX. Security, Secrecy and Tamper-Free Considerations", RAND Corporation, Santa Monica, CA, *Memorandum RM-3765-PR*, August 1964.

[27] Bobby Bhattacharjee, Ken Calvert, Jim Griffioen, Neil Spring, and James Sterbenz. FIND proposal. Postmodern internetwork architecture. http://www.nets-find.net/Funded/PostModern.php.

[28] Matt Blaze, Sampath Kannan, Angelos D. Keromytis, Insup Lee, Wenke Lee, Oleg Sokolsky, and Jonathan M. Smith, ''Dynamic Trust Management,'' *IEEE Computer (Special Issue on Trust Management)*, pp. 44-52 (February 2009).

[29] Dirk Bergemann, Joan Feigenbaum, Scott Shenker and Jonathan M. Smith, ''Towards An Economic Analysis of Trusted Systems,'' in *Proceedings, 3rd Annual Workshop on Economics and Information Security* (May 13-14, 2004).

[30] Ken Birman, "The League of SuperNets", *IEEE Internet Computing*, **7**(5) 2003, pp. 92-96.

[31] Kenneth P. Birman, "Reliable Distributed Systems: Technologies, Web Services, and Applications", Springer, 2005, ISBN: 0-387-21509-3.

[32] Ken Birman, Dahlia Malkhi, Robert Van Renesse, "Virtually Synchronous Methodology for Building Dynamic Reliable Services", Submitted to *ACM PODC 2010*, January 2010.

[33] R. Braden, D. Clark, and S. Shenker. "Integrated Services in the Internet architecture: an overview," RFC 1633, Network Working Group, June 1994.

[34] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin. "Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification," RFC 2205, Network Working Group, September 1997.

[35] A. Bremler-Barr and H. Levy. "Spoofing prevention method", In *INFOCOM*, March 2005.

[36] Matthew Caesar, Tyson Condie, Jayanthkumar Kannan, Karthik Lakshminarayanan, Ion Stoica, Scott Shenker, "ROFL: Routing on Flat Labels," *ACM SIGCOMM*, September 2006.

[37] Matthew Caesar, Miguel Castro, Edmund Nightingale, Greg O' Shea, Antony Rowstron, Virtual

Ring Routing: Network routing inspired by DHTs, ACM SIGCOMM, September 2006.

[38] Matthew Caesar, Donald Caldwell, Nick Feamster, Jennifer Rexford, Aman Shaikh, Kobus van der Merwe, "Design and Implementation of a Routing Control Platform," *Second Symposium on Networked Systems Design and Implementation (NSDI'05)*, April 2005.

[39] M. Caesar, L. Subramanian and R. H. Katz, "Root cause analysis of Internet routing dynamics", Technical Report, Univ. of California, Berkeley, 2003.

[40] K. Calvert, J. Griffioen, and L. Poutievski. Separating routing and forwarding: A clean-slate network layer design. In *Proc. IEEE Broadnets*, September 2007.

[41] Kenneth L. Calvert, James Griffioen, and Su Wen. "Lightweight Network Support for Scalable End-to-End Services". In *ACM SIGCOMM*, August 2002.

[42] I. Castineyra, N. Chiappa, and M. Steenstrup. The Nimrod routing architecture. RFC 1992, August 1996.

[43] V. Cerf and R. Kahn, "A Protocol for Packet Network Intercommunication", *IEEE Transactions on Communications*, Vol. **COM-22**, No. 5, pp 637-648, May 1974.

[44] F. Chang, J. Dean, S. Ghemawat, W. Hsieh, D. Wallach, M. Burrows, T. Chandra, A. Fikes and R. Gruber, "Bigtable: A distributed storage system for structured data", in *Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation (OSDI'06)* (2006).

[45] Yuu-Heng Cheng, Mariana Raykova, Alex Poylisher, Scott Alexander, Martin Eiger, and Steve M. Bellovin. The Zodiac policy subsystem: a policy-based management system for a high-security MANET. In *IEEE Policy*, July 2009.

[46] J. Chou, B. Lin, S. Sen, and O. Spatscheck. Proactive surge protection: a defense mechanism for bandwidth-based attacks. In *USENIX SECURITY*, July 2008.

[47] D. D. Clark, "The design philosophy of the DARPA internet protocols". Proc. SIGCOMM 1988, pp. 106-114.

[48] David D. Clark, John Wroclawski, Karen R. Sollins and Robert Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet", in Proc. ACM SIGCOMM 2002.

[49] David D. Clark, Bruce S. Davie, David J. Farber, Inder S. Gopal, Bharath K. Kadaba, W. David Sincoskie, Jonathan M. Smith, and David L. Tennenhouse, ''The AURORA Gigabit Testbed,'' *Computer Networks and ISDN Systems* **25**(6), pp. 599-621, North-Holland (January1993).

[50] D. Clark. "Policy routing in internet protocols", RFC 1102, May 1989.

[51] Douglas Comer, "Internetworking with TCP/IP: Volume 1: Principles, Protocols and Architecture, 5th Edition", Prentice-Hall 2006.

[52] Douglas Comer, "Computer Networks and Internets, 5th Edition", Prentice-Hall, 2009.

[53] Brian F. Cooper, Raghu Ramakrishnan, Utkarsh Srivastava, Adam Silberstein, Philip Bohannon, Hans-Arno Jacobsen, Nick Puz, Daniel Weaver, Ramana Yerneni, "PNUTS: Yahoo!'s hosted data serving platform", in PVLDB 1(2): 1277-1288 (2008)

[54] F. J. Corbato and V. A. Vyssotsky, "Introduction and Overview of the Multics System", in Proc. Fall Joint Computer Conference, 1965, v27.

[55] Y. Dalal and C. Sunshine, "Connection Management in Transport Protocols", *Computer Networks*, Vol. 2, No. 6, pp. 454-473, December 1978.

[56] Declarative Secure Distributed Systems, http://netdb.cis.upenn.edu/ds2/.

[57] C. Dixon, T. Anderson and A. Krishnamurthy, "Phalanx: Withstanding multimillion-node botnets", in *Proc. of Networked Systems Design and Implementation*, 2008

[58] Z. Duan, X. Yuan, and J. Chandrashekar. Constructing inter-domain packet filters to control IP spoofing based on BGP updates. In INFOCOM, April 2006.

[59] C. Elliott and A. Falk, "An update on the GENI project", in *SIGCOMM Comput. Commun. Rev.* 39, 3 (Jun. 2009), pp. 28-34.

[60] D. Estrin, T. Li, Y. Rekhter, K. Varadhan, and D. Zappala. Source demand routing: Packet format and forwarding specification (version 1). RFC 1940, May 1996.

[61] D. Estrin, J. Mogul, and G. Tsudik. "VISA protocols for controlling inter-organizational datagram flow". *IEEE JSAC*, **7**(4), May 1989.

[62] D. Estrin and G. Tsudik. Security issues in policy routing. In *Proc. IEEE Symposium on Security and Privacy*, May 1989.

[63] R. M. Fano, "The MAC System: The Computer Utility Approach," *IEEE Spectrum*, vol. 2, pp. 56-64 (Jan. 1965).

[64] Anja Feldmann, Olaf Maennel, Z. Morley Mao, Arthur Berger and Bruce Maggs, "Locating Internet routing instabilities", in *Proc. SIGCOMM*, 2004.

[65] D. C. Feldmeier, A. J. McAuley, J. M. Smith, D. S. Bakin, W. S. Marcus, and T. M. Raleigh, ''Protocol Boosters,'' *IEEE Journal on Selected Areas in Communication (Special Issue on Protocol Architectures for 21st Century Applications)* **16**(3), pp. 437-444 (April 1998).

[66] W. Feng. "The case for TCP/IP puzzles," *In Proc. SIGCOMM Workshop on Future Directions in Network Architecture*, August 2003.

[67] S. Floyd, "TCP and explicit congestion notification", in *SIGCOMM Computer Communications Review*, **24**(5), 1994 pp. 8-23.

[68] Formally Verifiable Networking, http://netdb.cis.upenn.edu/fvn/.

[69] Michael J. Freedman, Prem Gopalan, Steven Y. Ko, Jennifer Rexford, and David Shue. "A SCAFFOLD for Service-Centric Networking", Technical Report, Princeton University, Department of Computer Science. 2010.

[70] S. Ghemawat, H. Gobioff, and S.-T. Leung, “The Google file system”, in *Proc. 19th ACM Symposium on Operating Systems Principles* (New York, NY, USA, 2003), ACM, pp. 29–43.

[71] P. B. Godfrey, I. Ganichev, S. Shenker, and I. Stoica. “Pathlet routing”. In *Proc. ACM SIGCOMM*, August 2009.

[72] Sharon Goldberg, David Xiao, Eran Tromer, Boaz Barak, and Jennifer Rexford. “Path-quality monitoring in the presence of adversaries,” In *SIGMETRICS*, June 2008.

[73] A. Greenhalgh, M. Handley, and F. Huici. “Using routing and tunneling to combat DoS attacks,” In *Proc. USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI),* July 2005.

[74] Timothy G. Griffin and João Luís Sobrinho, “Metarouting”, SIGCOMM 2005.

[75] F. Gruenberger, “Computers and Communications; Toward a Computer Utility”, Prentice-Hall, 1968.

[76] S. Guha and P. Francis. “An end-middle-end approach to connection establishment,” In *Proc. ACM SIGCOMM*, August 2007.

[77] Krishna P. Gummadi, Harsha V. Madhyastha, Steven D. Gribble, Henry M. Levy, and David Wetherall. “Improving the reliability of Internet paths with one-hop source routing”, In *Proc. USENIX Symposium on Operating Systems Design and Implementation (OSDI),* December 2004.

[78] Andreas Haeberlen, Petr Kuznetsov, "The Fault Detection Problem", *13th International Conference on Principles of Distributed Systems (OPODIS '09)*, Nîmes, France, December 2009.

[79] Andreas Haeberlen, "A Case for the Accountable Cloud", *3rd ACM SIGOPS International Workshop on Large-Scale Distributed Systems and Middleware (LADIS '09)*, Big Sky, MT, October 2009.

[80] Andreas Haeberlen, Ioannis Avramopoulos, Jennifer Rexford, Peter Druschel, "NetReview: Detecting when interdomain routing goes wrong", *Proceedings of the 6th Symposium on Networked Systems Design and Implementation (NSDI '09)*, Boston, MA, April 2009.

[81] Andreas Haeberlen, Petr Kuznetsov, Peter Druschel, "PeerReview: Practical Accountability for Distributed Systems", *Proceedings of the 21st ACM Symposium on Operating Systems Principles (SOSP '07)*, Stevenson, WA, October 2007.

[82] Mark Handley and Adam Greenhalgh. Steps towards a DoS-resistant Internet architecture. In Proc. SIGCOMM Workshop on Future Directions in Network Architecture, August 2004.

[83] M. W. Hicks, J. T. Moore, D. S. Alexander, C. A. Gunter and S. M. Nettles, “PLANet: An Active Internetwork”, Proc. IEEE Infocom, 1999.

[84] M. Hicks, M. and S. Nettles, “Dynamic software updating”, in *ACM Trans. Program. Lang. Syst.* 27, 6 (Nov. 2005), pp. 1049-1096

[85] Y.C. Hu, A. Perrig, and D.B. Johnson. Ariadne: A secure on-demand routing protocol for ad hoc networks. In *MOBICOM*, September 2002.

[86] Yih-Chun Hu, Adrian Perrig, and Dave Johnson. Efficient security mechanisms for routing protocols. In *NDSS*, February 2003.

[87] Yih-Chun Hu, Adrian Perrig, and Marvin Sirbu. "SPV: Secure path vector routing for securing BGP," In *Proc. ACM SIGCOMM*, September 2004.

[88] J. Ioannidis and S. M. Bellovin. "Implementing pushback: Router-based defense against DDoS Attacks," In *Proc NDSS*, 2002.

[89] Sotiris Ioannidis, Angelos D. Keromytis, Steve Bellovin, and Jonathan M. Smith, ''Implementing a Distributed Firewall,'' in *Proceedings, ACM Conference on Computer and Communications Security*, Athens, GREECE (November 2000), pp. 190-199.

[90] Zachary G. Ives, Todd J. Green, Grigoris Karvounarakis, Nicholas E. Taylor, Val Tannen, Partha Pratim Talukdar, Marie Jacob, Fernando Pereira, "The Orchestra Collaborative Data Sharing System", in ACM SIGMOD Record, September 2008.

[91] V. Jacobson, "Congestion Avoidance and Control", in *Proc. SIGCOMM 1988*, Stanford, CA., pp. 314-329.

[92] V. Jacobson, Diana K. Smetters, James D. Thornton, Michael F. Plass, Nicholas H. Briggs, Rebecca L. Braynard, "Networking named content", in *Proc. CoNEXT 2009*.

[93] C. Jin, H. Wang, and K.G. Shin. Hop-count filtering: an effective defense against spoofed DDoS traffic. In Proc. ACM Conference on Computer and Communications Security (CCS), October 2003.

[94] J. John, E. Katz-Bassett, A. Krishnamurthy, T. Anderson and A. Venkataramani, "Consensus routing: the Internet as a distributed system", in Proc. of *Networked Systems Design and Implementation*, 2008.

[95] J. John, A. Moshchuk, S. Gribble and A. Krishnamurthy, "Studying Spamming Botnets using Botlab", in *Proc. of Networked Systems Design and Implementation*, 2009.

[96] Ethan Katz-Bassett, John P. John, Arvind Krishnamurthy, David Wetherall and Thomas Anderson and Yatin Chawathe, "Towards IP Geolocation using Delay and Topology Measurements", in *Proc. of Internet Measurements Conference*, 2006.

[97] E. Katz-Bassett, H. Madhyastha, J. John, A. Krishnamurthy, D. Wetherall and T. Anderson, "Studying blackholes in the Internet with Hubble", in *Proc. of Networked Systems Design and Implementation*, 2008.

[98] E. Katz-Bassett, H. Madhyastha, V. Adhikari, C. Scott, J. Sherry, P. van Wessep, T. Anderson and A. Krishnamurthy, "Reverse Traceroute", in *Proc. of Networked Systems Design and Implementation*, 2010.

[99] Hema Tahilramani Kaur, Andreas Weiss, Shifalika Kanwar, Shivkumar Kalyanaraman, and Ayesha Gandhi. "BANANAS: An evolutionary framework for explicit and multipath routing in the internet," In *Proc. SIGCOMM Workshop on Future Directions in Network Architecture*, August 2004.

[100] Eric Keller, Minlan Yu, Matthew Caesar, Jennifer Rexford, "Virtually Eliminating Router Bugs",

ACM CoNEXT, December 2009.

[101] S. Kent, C. Lynn, and K. Seo. Secure border gateway protocol (S-BGP). IEEE JSAC, 18(4), April 2000.

[102] Angelos Keromytis, John Ioannidis, and Jonathan M. Smith, ''Implementing IPSec,'' in *Proceedings, IEEE GlobeCom Conference*, Phoenix, AZ (November, 1997), pp. 1948-1952.

[103] Angelos D. Keromytis and Jonathan M. Smith, ''Requirements for Scalable Access Control and Security Management Architectures,'' *ACM Transactions on Internet Technology* **7**(4) (November 2007).

[104] Changhoon Kim, Matthew Caesar, Alex Gerber, Jennifer Rexford, "Revisiting Route Caching: The World Should Be Flat, Passive and Active," *Measurement Conference*, April 2009.

[105] Changhoon Kim, Matthew Caesar, Jennifer Rexford, "Floodless in SEATTLE: A Scalable Ethernet Architecture for Large Enterprises", *ACM SIGCOMM*, August 2008.

[106] Firat Kiyak, Brent Mochizuki, Eric Keller, Matthew Caesar, "Better by a HAIR -- Hardware Amenable Internet Routing", *IEEE ICNP*, October 2009.

[107] Teemu Koponen, Mohit Chawla, Byung-Gon Chun, Andrey Ermolinskiy, Kye Hyun Kim, Scott Shenker, and Ion Stoica. "A data-oriented (and beyond) network architecture", In *Proc. ACM SIGCOMM*, August 2007.

[108] R. Krishnan, H. Madhyastha, S. Srinivasan, S. Jain, A. Krishnamurthy, T. Anderson and J. Gao, "Moving Beyond End-to-End Path Information to Optimize CDN Performance, in *Proc. of Internet Measurement Conference*, 2009.

[109] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. "Delayed Internet routing Convergence", in *ACM/IEEE Trans. on Networking*, 9(3):293-306, June 2001.

[110] Karthik Kalambur Lakshminarayanan, Matthew Chapman Caesar, Murali Rangan, Thomas Anderson, Scott Shenker and Ion Stoica, "Achieving Convergence-Free Routing using Failure-Carrying Packets", in *Proc. of ACM SIGCOMM*, 2007.

[111] B. Lampson, "Protection," in *Proc. 5th Princeton Symp. Information Science and Systems* (Mar. 1971), pp. 437-443. (Reprinted in ACM Operating Syst. Rev., vol. 8, pp. 18-24, Jan. 1974.)

[112] J. Li, J. Mirkovic, M. Wang, P. Reiher, and L. Zhang. "SAVE: Source address validity enforcement protocol", In *INFOCOM*, June 2002.

[113] Chia-Chi Lin, Matthew Caesar, Jacobus van der Merwe, "Towards Interactive Debugging for ISP Networks", *HotNets-VIII*, October 2009.

[114] Mengmeng Liu, Nicholas E. Taylor, Wenchao Zhou, Zachary Ives, and Boon Thau Loo, "Recursive Computation of Regions and Connectivity in Networks". In Proc. ICDE 2009.

[115] X. Liu, A. Li, X. Yang, and D. Wetherall. "Passport: Secure and adoptable source authentication," In *Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI),* April 2008.

[116] Xin Liu, Xiaowei Yang, and Yanbin Lu. "To filter or to authorize: Network-layer DoS defense against multimillion-node botnets", In *Proc. ACM SIGCOMM*, August 2008.

[117] Boon Thau Loo, Joseph M. Hellerstein, Ion Stoica, and Raghu Ramakrishnan, "Declarative Routing: Extensible Routing with Declarative Queries", ACM SIGCOMM Conference on Data Communication, Philadelphia, PA, Aug 2005.

[118] Boon Thau Loo, Tyson Condie, Joseph M. Hellerstein, Petros Maniatis, Timothy Roscoe, and Ion Stoica, "Implementing Declarative Overlays", 20th ACM Symposium on Operating Systems Principles (SOSP), Brighton, UK, October 2005.

[119] Boon Thau Loo, Tyson Condie, Minos Garofalakis, David E. Gay, Joseph M. Hellerstein, Petros Maniatis, Raghu Ramakrishnan, Timothy Roscoe, and Ion Stoica, "Declarative Networking: Language, Execution and Optimization", ACM SIGMOD International Conference on Management of Data, Chicago, June 2006.

[120] Harsha V. Madhyastha, Tomas Isdal, Michael Piatek, Colin Dixon, Thomas Anderson, Arvind Krishnamurthy and Arun Venkataramani, "iPlane: An Information Plane for Distributed Services", in *Proc. of Operating System Design and Implementation*, 2006.

[121] H. Madhyastha, E. Katz-Bassett, T. Anderson, A. Krishnamurthy and A. Venkataramani, "iPlane Nano: Path Prediction for Peer-to-Peer Applications", in *Proc. of Networked Systems Design and Implementation*, 2009.

[122] Ratul Mahajan, David Wetherall and Tom Anderson, "Understanding BGP misconfiguration", in *Proc. of ACM SIGCOMM*, 2002, pp. 3-16.

[123] Ratul Mahajan, David Wetherall and Thomas E. Anderson, "Negotiation-Based Routing Between Neighboring ISPs", in *Proc. NSDI*, 2005.

[124] Ratul Mahajan, David Wetherall and Thomas E. Anderson, "Mutually Controlled Routing with Independent ISPs", in *Proc. NSDI*, 2007.

[125] R. Mahajan, Steven M. Bellovin, Sally Floyd, John Ioannidis, Vern Paxson, and Scott Shenker. Controlling high bandwidth aggregates in the network. ACM CCR, 32(3), July 2002.

[126] Yun Mao, Boon Thau Loo, Zachary Ives, and Jonathan M. Smith, "The Case for a Unified Extensible Data-centric Mobility Infrastructure", in *Proc. 2nd ACM International Workshop on Mobility in the Evolving Internet Architecture (MobiArch),* August 2007.

[127] Yun Mao, Boon Thau Loo, Zachary Ives, and Jonathan M. Smith, "MOSAIC: Unified Declarative Platform for Dynamic Overlay Composition", in 4th Conference on emerging Networking EXperiments and Technologies (ACM CoNEXT), Madrid, Spain, Dec 2008.

[128] W. Marcus, I. Hadzic, T. McAuley, and J. Smith, ''Protocol Boosters: Applying Programmability to Network Infrastructures,'' *IEEE Communications* **36**(10), pp. 79-83 (October 1998).

[129] William R. Marczak, Shan Shan Huang, Martin Bravenboer, Micah Sherr, Boon Thau Loo, and Molham Aref, "SecureBlox: Customizable Secure Distributed Data Processing", ACM SIGMOD International Conference on Management of Data (SIGMOD), June 2010.

[130] N. Maxemchuk, "Dispersity Routing", Ph.D. Thesis, University of Pennsylvania, 1975.

[131] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: Enabling innovation in campus networks" in *ACM SIGCOMM Computer Communication Review* **38**(2) (April 2008).

[132] A. T. Mizrak, Yu-Chung Cheng, K. Marzullo, and S. Savage. Fatih: Detecting and isolating malicious routers. In IEEE DSN, June 2005.

[133] Morris, R., Kohler, E., Jannotti, J., and Kaashoek, M. F., "The Click modular router" in *SIGOPS Oper. Syst. Rev.* 33, 5 (Dec. 1999), pp. 217-231

[134] Shivkumar C. Muthukumar, Xiaozhou Li, Changbin Liu, Joseph B. Kopena, Mihai Oprea, Richardo Correa, Boon Thau Loo, and Prithwish Basu, "RapidMesh: Declarative Toolkit for Rapid Experimentation of Wireless Mesh Networks", in 4th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WiNTECH 2009), in conjunction with ACM MobiCom, Beijing, China, Sept, 2009.

[135] Shivkumar C. Muthukumar, Xiaozhou Li, Changbin Liu, Joseph B. Kopena, Mihai Oprea, and Boon Thau Loo, "Declarative Toolkit for Rapid Network Protocol Simulation and Experimentation", ACM SIGCOMM Conference on Data Communication (demo), Barcelona, Spain, Aug 2009.

[136] Jad Naous, Arun Seehra, Michael Walfish, David Mazières, Antonio Nicolosi and Scott Shenker, "Defining and enforcing transit policies in a future Internet," *Technical Report TR-10-07*, Department of Computer Science, The University of Texas at Austin, February 2010.

[137] Jad Naous, Arun Seehra, Michael Walfish, David Mazières, Antonio Nicolosi, and Scott Shenker, "The design and implementation of a policy framework for the future Internet," *Technical Report TR-09-28*, Department of Computer Sciences, The University of Texas at Austin, September 2009.

[138] Jad Naous, Michael Walfish, David Mazières, Antonio Nicolosi, and Arun Seehra, "Network Security Via Explicit Consent," *Technical Report TR-09-12*, Department of Computer Sciences, The University of Texas at Austin, March 2009.

[139] Venkata N. Padmanabhan and Daniel R. Simon. "Secure traceroute to detect faulty or malicious routing", *Proc. ACM SIGCOMM*, 33(1):77–82, 2003.

[140] P. Papadimitratos and Z.J. Haas. Secure routing for mobile ad hoc networks. In Proc. SCS Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS), January 2002.

[141] P. Papadimitratos and Z.J. Haas. Secure link state routing for mobile ad hoc networks. In Proc. IEEE Workshop on Security and Assurance in Ad hoc Networks, January 2003.

[142] Bryan Parno, Adrian Perrig, and David G. Andersen. SNAPP: Stateless Network-Authenticated Path Pinning. In ACM ASIACCS, 2008.

[143] Bryan Parno, Dan Wendlandt, Elaine Shi, Adrian Perrig, Bruce Maggs, and Yih-Chun Hu. Portcullis: Protecting connection setup from denial-of-capability attacks. In Proc. ACM SIGCOMM, August 2007.

[144] R. Perlman. Network layer protocols with Byzantine robustness. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1988.

[145] R. Perlman. Routing with Byzantine robustness. Technical Report TR-2005-146, Sun Microsystems, August 2005.

[146] L. Peterson, T. Anderson, D. Culler, and T. Roscoe, "A Blueprint for Introducing Disruptive Technology into the Internet", in *Proc. HotNets-I '02*, Oct. 2002

[147] Michael Piatek, Tomas Isdal, Arvind Krishnamurthy and Thomas Anderson, "Do incentives build robustness in BitTorrent?", in *Proc. of Networked Systems Design and Implementation*, 2007.

[148] M. Piatek, T. Isdal, A. Krishnamurthy and T. Anderson, "One hop Reputations for Peer-to-Peer File Sharing Workloads", in *Proc. of Networked Systems Design and Implementation*, 2008.

[149] M. Piatek, A. Krishnamurthy, A. Venkataramani, R. Yang and D. Zhang, "Contracts: Practical Contribution Incentives for P2P Live Streaming", in *Proc. of Networked Systems Design and Implementation*, 2010.

[150] Lucian Popa, Ion Stoica, and Sylvia Ratnasamy. "Rule-based Forwarding (RBF): improving the Internet's flexibility and security", In *Proc. ACM Workshop on Hot Topics in Networks (HotNets),* October 2009.

[151] Postel, J. (ed.), "User Datagram Protocol", RFC 768, USC/Information Sciences Institute, August 1980.

[152] Postel, J. (ed.), "Internet Protocol - DARPA Internet Program Protocol Specification", RFC 791, USC/Information Sciences Institute, September 1981.

[153] Postel, J. "Transmission Control Protocol - DARPA Internet Program Protocol Specification", RFC793, USC/Information Sciences Institute, September 1981

[154] Barath Raghavan and Alex C. Snoeren. A system for authenticated policy-compliant routing. In Proc. ACM SIGCOMM, September 2004.

[155] RapidNet Declarative Networking System, http://netdb.cis.upenn.edu/rapidnet/.

[156] T. Ristenpart, G. Maganis, A. Krishnamurthy and T. Kohno, "Privacy-Preserving Location Tracking of Lost or Stolen Devices: Cryptographic Techniques and Replacing Trusted Third Parties with DHTs", in *Proceedings of Usenix Security*, 2008.

[157] E. Rosen, A. Viswanathan, and R. Callon. "Multiprotocol label switching", RFC 3031, Network Working Group, January 2001.

[158] J. H. Saltzer, D. P. Reed and D. D. Clark, "End-to-End arguments in system design," in *ACM Transactions on Computer Systems*, **2**(4), November 1984, pp. 277-288.

[159] K. Sanzgiri, B. Dahill, B.N. Levine, C. Shields, and E.M. Belding-Royer. "A secure routing protocol for ad hoc networks". In *Proc. IEEE Conference on Network Protocols*, November 2002.

[160] S. Savage, D. Wetherall, A. Karlin, and T. Anderson. "Network support for IP traceback," *ACM/IEEE Transactions on Networking*, 9(3), June 2001.

[161] Arun Seehra, Jad Naous, Michael Walfish, David Mazières, Antonio Nicolosi, and Scott Shenker, "A policy framework for the future Internet", in *ACM Workshop on Hot Topics in Networks (HotNets)*, New York, NY, October 2009.

[162] Segal, M. E. and Frieder, O. "On-the-Fly Program Modification: Systems for Dynamic Updating". *IEEE Softw*. 10, 2 (Mar. 1993), pp. 53-65.

[163] Micah Sherr, Andrew Mao, William R. Marczak, Wenchao Zhou, Boon Thau Loo, and Matt Blaze. A3: An Extensible Platform for Application-Aware Anonymity. 17th Annual Network & Distributed System Security Symposium (NDSS), Feb 2010.

[164] W. D. Sincoskie and C. J. Cotton, "Extended Bridge Algorithms for Large Networks", *IEEE Network* **2**(1), January 1988, pp. 16-24.

[165] Jonathan M. Smith, Kenneth L. Calvert, Sandra L. Murphy, Hilarie K. Orman, and Larry L. Peterson, ''Activating Networks: A Progress Report,'' *IEEE Computer* **32**(4), pp. 32-41 (April 1999).

[166] Jonathan M. Smith, ''Application-Private Networks,'' in *Computer Systems: Theory, Technology and Applications: A Tribute to Roger Needham*, ed. Andrew Herbert and Karen Sparck Jones, Springer-Verlag (2004), pp. 273-278.

[167] Jonathan M. Smith, ''Fighting Physics: A Tough Battle,'' *ACM Queue Magazine*, pp. 20-26, (February/March 2009) (extended version appeared in CACM July 2009).

[168] Neil Spring, Ratul Mahajan and Thomas Anderson, "The causes of path inflation", in *Proc. of ACM SIGCOMM*, 2003.

[169] Neil Spring, Ratul Mahajan, David Wetherall and Tom Anderson, "Measuring ISP Topologies with Rocketfuel", in *IEEE/ACM Transactions on Networking*, 2004.

[170] Angelos Stavrou, John Ioannidis, Angelos D. Keromytis, Vishal Misra, and Dan Rubenstein. "A pay-per-use DoS protection mechanism for the Web", In *Proc. International Conference on Applied Cryptography and Network Security*, June 2004.

[171] Ion Stoica, Daniel Adkins, Shelley Zhuang, Scott Shenker, and Sonesh Surana. "Internet Indirection Infrastructure," In *ACM SIGCOMM*, Pittsburgh, PA, August 2002.

[172] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. H. Katz. "Listen and whisper: Security mechanisms for BGP," In *Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, March 2004.

[173] Lakshminarayanan Subramanian, Matthew Caesar, Cheng Tien Ee, Mark Handley, Morley Mao, Scott Shenker, Ion Stoica, "HLP: A Next-generation Interdomain Routing Protocol", *ACM SIGCOMM*, August 2005.

[174] Nicholas E. Taylor and Zachary G. Ives. "Reliable Storage and Querying for Collaborative Data Sharing Systems", to appear in Proc. ICDE 2010.

[175] D. L. Tennenhouse, J. M. Smith, W. D. Sincoskie, D. J. Wetherall, and G. J. Minden, ‘‘A Survey of Active Network Research,’’ *IEEE Communications* **35**(1), pp. 80-86 (January, 1997).

[176] David L. Tennenhouse and David J. Wetherall. Towards an Active Network Architecture. SIGCOMM Comput. Commun. Rev., 37(5), 2007.

[177] C. Brendan S. Traw and Jonathan M. Smith, ‘‘Striping within the Network Subsystem,’’ *IEEE Network*, pp. 22-32 (July/August 1995).

[178] Ymir Vigfusson, Hussam Abu-Libdeh, Mahesh Balakrishnan, Ken Birman, Robert Burgess, Haoyuan Li, Gregory Chockler, Yoav Tock, “Dr. Multicast: Rx for Data Center Communication Scalability,” in *Proc. Eurosys*, April 2010 (Paris, France).

[179] Michael Walfish, Jeremy Stribling, Maxwell Krohn, Hari Balakrishnan, Robert Morris, and Scott Shenker. “Middleboxes no longer considered harmful,” In *Proc. USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, December 2004.

[180] Anduo Wang, Prithwish Basu, Boon Thau Loo, and Oleg Sokolsky, “Declarative Network Verification”, 11th International Symposium on Practical Aspects of Declarative Languages (PADL), in conjunction with POPL, Savannah, Georgia, Jan 2009.

[181] Anduo Wang, Limin Jia, Changbin Liu, Boon Thau Loo, Oleg Sokolsky, and Prithwish Basu, “Formally Verifiable Networking”, 8th Workshop on Hot Topics in Networks (ACM SIGCOMM HotNets-VIII), New York, Oct 2009.

[182] XiaoFeng Wang and Michael K. Reiter. A multi-layer framework for puzzle-based denial-of-service defense. International Journal of Information Security, 2007.
Forthcoming and published online, http://dx.doi.org/10.1007/s10207-007-0042-x.

[183] Brent Waters, Ari Juels, J. Alex Halderman, and Edward W. Felten. “New client puzzle outsourcing techniques for DoS resistance”. In *Proc. ACM Conference on Computer and Communications Security (CCS)*, October 2004.

[184] Andrew G. West, Adam J. Aviv, Jian Chang, Vinayak S. Prabhu, Matt Blaze, Sampath Kannan, Insup Lee, Jonathan M. Smith, and Oleg Sokolsky, ‘‘QuanTM: A Quantitative Trust Management System,’’ in *EUROSEC 2009: Proceedings of the Second European Workshop on Systems Security* (2009), pp. 28-35.

[185] Edmund L. Wong, Praveen Balasubramanian, Lorenzo Alvisi, Mohamed G. Gouda, and Vitaly Shmatikov. “Truth in advertising: lightweight verification of route integrity,” In *PODC*, August 2007.

[186] Geoffrey Xie, Jibin Zhan, Dave Maltz, Hui Zhang, Albert Greenberg, Gisli Hjalmtysson, and Jennifer Rexford. “On static reachability analysis of IP networks”, In *INFOCOM*, March 2005.

[187] H. Xie, R. Yang, A. Krishnamurthy, Y. Liu and A. Silberschatz, “P4P: Provider Portal for (P2P) Applications”, in *Proceedings of SIGCOMM 2008*.

[188] Wen Xu and Jennifer Rexford. “MIRO: Multi-path interdomain routing”, In *Proc. ACM SIGCOMM*, September 2006.

[189] A. Yaar, A. Perrig, and D. Song. “Pi: a path identification mechanism to defend against DDoS attacks,” In *Proc. IEEE Symposium on Security and Privacy*, May 2003.

[190] A. Yaar, A. Perrig, and D. Song. “SIFF: A stateless Internet flow filter to mitigate DDoS flooding attacks,” In *Proc. IEEE Symposium on Security and Privacy*, May 2004.

[191] A. Yaar, A. Perrig, and D. Song. “FIT: fast Internet traceback,” In *INFOCOM*, March 2005.

[192] A. Yaar, A. Perrig, and Dawn Song. “StackPi: New packet marking and filtering mechanisms for DDoS and IP spoofing defense”, *IEEE JSAC*, **24**(10):1853–1863, October 2006.

[193] Xiaowei Yang, David Wetherall and Thomas Anderson, “A DoS-limiting Network Architecture”, in *Proc. of ACM SIGCOMM*, 2005.

[194] Xiaowei Yang, David Clark, and Arthur W. Berger. NIRA: A new inter-domain routing architecture. *ACM/IEEE Transactions on Networking*, **15**(4), August 2007.

[195] Xiaowei Yang and Xin Liu. “Internet Protocol made accountable,” In *Proc. ACM Workshop on Hot Topics in Networks (HotNets)*, October 2009.

[196] Xiaowei Yang and David Wetherall. “Source selectable path diversity via routing deflections,” In *Proc. ACM SIGCOMM*, September 2006.

[197] Manel Guerrero Zapata and N. Asokan. Securing ad hoc routing protocols. In *Proc. ACM Workshop on Wireless Security (WiSE),* September 2002.

[198] Xin Zhang, Abishek Jain, and Adrian Perrig. “Packet-dropping adversary identification for data plane security”, In *Proc. ACM Conference on emerging Networking EXperiments and Technologies (CoNEXT),* December 2008.

[199] Hang Zhao, Maritza Johnson, Chi-Kin Chau, and Steven M. Bellovin. “Source prefix filtering in ROFL,” *Technical Report CUCS-033-09*, Department of Computer Science, Columbia University, July 2009.

[200] Wenchao Zhou, Eric Cronin and Boon Thau Loo, “Provenance-aware Secure Networks”, 4th International Workshop on Networking meets Databases (NetDB), in conjunction with ICDE, Cancun, Mexico, Apr 2008.

[201] Wenchao Zhou, Yun Mao, Boon Thau Loo, and Martín Abadi, “Unified Declarative Platform for Secure Networked Information Systems”, 25th International Conference on Data Engineering (ICDE), Shanghai, China, Apr 2009.

[202] Wenchao Zhou, Micah Sherr, Tao Tao, Xiaozhou Li, Boon Thau Loo, and Yun Mao, “Efficient Querying and Maintenance of Network Provenance at Internet-Scale”, ACM SIGMOD International Conference on Management of Data (SIGMOD), June 2010.