ANNUAL REPORT FOR AWARD # 0205519

Scott Shenker ; *Internationl CompSci Inst*
ITR/ANI: Addressing Fundamental Issues for Robust Internet Performance

Participant Individuals:
CoPrincipal Investigator(s) : Vern Paxson; Sally Floyd
Senior personnel(s) : Mark Allman
Post-doc(s) : Alberto Medina; Morley Mao; Nicholas Weaver
Senior personnel(s) : Mark Handley; Richard Karp
Graduate student(s) : Aikaterini Argyraki; Cheng Tien Ee; Lakshmi Subramanian; Junwen Lai; Robin Sommer; Arun Venkataramani; Vinod Yegneswarin; Katerina Argyraki; Barath Raghavan
Senior personnel(s) : Nicholas Weaver

Participants' Detail

Partner Organizations:

Other collaborators:

As the list of publications indicates, we have a wide range of collaborators in this work, including former members of ICSI and researchers at the University of California at Berkeley. In addition, the IETF documents such as internet-drafts and RFCs all involved a wide range of collaboration and feedback.

Activities and findings:

Research and Education Activities:

The Internet is large, decentralized, and heterogeneous in its technology, administration and capacity. The core of the Internet's success arises from its adherence to a number of architectural principles, central to which is the notion that the network should try to achieve a robust, 'very often works pretty well' level of performance. One of the main techniques for achieving this across a wide range of conditions is making Internet protocols and mechanisms adaptive, so that they self-tune to work reasonably well in whatever circumstances they find themselves.

This approach has been extremely successful. However, some of the mechanisms, designed to be 'good enough' across a large range of conditions, must now or will soon operate in regimes beyond their effective dynamic range. For example, TCP congestion control mechanisms require revisiting for tomorrow's Internet, both the coming high speed paths, and the coming low speed paths (e.g., lossy wireless links). Similarly, the architecture's ability to gracefully tolerate failures does not extend to new forms of failures, such as misconfigured routing information or malfunctioning middle-boxes, nor to distributed stresses, such as flash crowds, rapidly-spreading worms, or denial-of-service (DoS) floods.

If we view robustness as the ability of the network to function well over a wide spectrum of conditions - particularly given a very large, ever-growing and ever-changing network - then we argue that the robustness of the future Internet is clearly at risk. In this project, we are taking a multifaceted approach to robustness, addressing the following areas:

* Robust performance in the presence of extreme environments such as very high speed and highly variable delay, which requires rethinking today's congestion control mechanisms.

* Robust performance in the presence of new forms of failure, both at the network layer, in terms of robust routing, and at the application layer, in terms of coping with the now widespread deployment of 'middleboxes' that have elbowed their way into the architecture.

* Robust performance in the presence of distributed stresses, both malicious (denial-of-service floods; congestion control cheaters; rapidly-spreading worms) and merely teeming (flash crowds). This will require an understanding of both the network's topology and the makeup of traffic aggregates, coupled with new control mechanisms deployed inside the network.

* Measurements to help understand and improve robustness. While part of our approach to these is to refine existing protocols and mechanisms, and investigate new ones, we also emphasize achieving robust performance by detecting incipient failures, on both short time scales (via distributed operational monitoring) and long time scales (via diagnostic probing of deployed protocol implementations).

While these activities do not address the full range of challenges facing the Internet architecture, they do address a few of the core issues in preserving and enhancing the Internet's robustness. In one sense, this research agenda is 'conservative,' in that we frame it in terms of working within the current Internet architecture, rather than advocating a ``clean sheet'' approach. However, we feel that this conservative approach makes for research that is more fundamental rather than less. The 'clean sheet' approach, while more tidy and much more conducive to easy exploration of basic principles, misses the crucial reality of how different mechanisms wind up interacting once integrated into a truly large-scale network.

Below is a list of activities, roughly grouped along the four broad categories listed above:

**Coping with Extreme Environments:

DCCP: Historically, the great majority of Internet unicast traffic has used congestion-controlled TCP, with UDP making up most of the remainder. UDP has mainly been used for short, request-response transfers, like DNS and SNMP, that wish to avoid TCP's three-way handshake, congestion control, retransmission, and/or stateful connections. Recent years have seen the growth of applications that use UDP in a different way. These applications, including RealAudio, Internet telephony, and on-line games such as Half Life,

Quake, and Starcraft, share a preference for timeliness over reliability. TCP can introduce arbitrary delay because of its reliability and in-order delivery requirements; thus, these delay-sensitive and loss- tolerant applications use UDP instead. This growth of long-lived non-congestion-controlled traffic, relative to congestion-controlled traffic, poses a real threat to the overall health of the Internet. We have designed a new protocol, Datagram Congestion Control Protocol (DCCP), that combines unreliable datagram delivery with built-in congestion control, intended for use by applications such as streaming media or on-line games.
DCCP is currently being standardized in the IETF.

High Speed TCP (2004):
HighSpeed TCP is a proposed modification to TCP's congestion control mechanisms for better operation in TCP connections with large congestion windows. The congestion control mechanisms of the current Standard TCP constrains the congestion windows that can be achieved by TCP in realistic environments. For example, for a Standard TCP connection with 1500-byte packets and a 100 ms round-trip time, achieving a steady-state throughput of 10 Gbps would require an average congestion window of 83,333 segments, and a packet drop rate of at most one congestion event every 5,000,000,000 packets (or equivalently, at most one congestion event every 1 2/3 hours). This is widely acknowledged as an unrealistic constraint. To address this limitation of TCP, we have proposed HighSpeed TCP, a modification to TCP's congestion control mechanisms, and standardized it as an Experimental protocol in the IETF. HighSpeed TCP is easy to deploy, only needing to be implemented at the transport data sender.

Transport's Robustness to Reordering:
TCP and SCTP's primary loss detection algorithm depends on packets arriving in roughly the order they were transmitted. If packet arrivals are significantly out of order TCP and SCTP will spuriously retransmit data. We have developed a scheme to detect such needless retransmissions, with the goal of allowing TCP and SCTP to undo mistakenly taken congestion control actions and adapt to the network conditions to prevent future spurious retransmissions. Such a mechanism enables increased robustness to packet reordering in networks. We are also working on an alternative approach that allows TCP and SCTP to avoid more spurious retransmits triggered by reordering, and therefore avoid the mistakes in the first place. The downside of this approach is that it makes the congestion control response more sluggish than the current standard allows. We are currently analyzing various methods to find the right balance to correctly distinguish between loss and reordering while also following proper congestion control procedures.

Transport's Robustness to Corruption:
TCP's congestion control algorithms assume that all packet loss is caused by contention for resources in a bottleneck router in the network (i.e., congestion). However, over many wireless networks loss happens due to link outages and packet corruption. TCP mistakenly assumes these losses are due to congestion and lowers the sending rate in an attempt to reduce the congestion in the network. We have developed a scheme, Cumulative Explicit Transport Error Notification (CETEN), whereby a transport protocol endpoint continuously requests information about corruption-based losses from routers along the network path. Using this aggregate information the sender can form a congestion response that is proportional to the amount of congestion observed, rather than the

amount of loss observed. CETEN, therefore, extends TCP's robustness and useful range into noisy networks with much non-congestion loss where current TCP is not a viable method to transfer data.

Transport's Robustness to Burstiness:
We have been studying the effects of naturally occurring bursts on TCP's performance. Over the years a number of researchers have suggested ways to mitigate these bursts. Our study is two-fold. First, we are comparing and contrasting these suggested burst mitigations with several of our own. In addition, we are assessing naturally occurring bursts found in passive network measurements, to determine the degree to which these bursts are a performance problem for TCP connections and a robustness issue for the network (in terms of creating congestion).

QuickStart:
Quick-Start is a proposed mechanism for faster start-up for TCP and other transport protocols in environments with significant available bandwidth. The current TCP protocol starts-up using an initial window of up to four packets, and then the TCP sender uses slow-start to double the sending rate each round-trip time until congestion is encountered. As a result, it may take a number of round-trip times in slow-start before the TCP connection begins to fully use the available bandwidth. Quick-Start is a proposal that would allow TCP connections to start-up faster, by asking the routers along the path to approve a higher initial sending rate. Quick-Start is designed to allow TCP connections to use high initial windows in circumstances when there is significant unused bandwidth along the path, and all of the routers along the path support the Quick-Start Request.

**Coping with New Forms of Failure:

Providing Packet Obituaries:
The Internet provides no information on the fate of transmitted packets. As a result, end-systems can neither verify whether their ISPs honor the corresponding service level agreements, nor intelligently adapt their routes to adverse network conditions. Current probing tools go a long way to help. Unfortunately they only give feedback on probe traffic and at router granularity, exposing the internals of service providers, who are slowly reasserting the opacity of their networks. To rectify this, we propose an accountability infrastructure of monitoring boxes on links between autonomous systems. These boxes, maintained by the providers themselves, supply the sources of traffic they observe with proactive feedback on per-flow path, loss, and delay at autonomous system granularity. For typical Internet traffic, our approach introduces less that 5% of bandwidth overhead with moderate requirements of memory and storage. Armed with the feedback supplied by our architecture, end systems can intelligently interact with proposed source routing or path selection architectures to gain control over their traffic without treading on providers' internals.

Evolving the Internet Architecture:
There is widespread (though not unanimous) agreement on the need for architectural change in the Internet, but very few believe that current ISPs will ever effect such changes. In this paper we ask what makes an architecture evolvable, by which we mean capable of gradual change led by the incumbent providers. This involves both technical and economic issues, since ISPs have to be able, and incented, to offer new architectures. Our analysis suggests that,

with very minor modifications, the current Internet architecture could be evolvable.

Lightweight Security Mechanisms for BGP (2004):
BGP, the current inter-domain routing protocol, assumes that the routing information propagated by routers is correct. A violation of this assumption leaves the current infrastructure vulnerable to misconfigurations and deliberate attacks that alter the behavior of the control and data planes. Deliberate attackers along a path can potentially render destinations unreachable, eavesdrop on data passing through them, impersonate a site, and take countermeasures against security measures. We developed a series of mechanisms of increasing complexity that deal with attacks of increasing sophistication called Listen and Whisper.

Robust Interdomain Routing (2004):
While BGP can be made more secure using mechanisms like Listen and Whisper, the basic protocol remains quite fragile. BGP has notoriously poor convergence, and routes are quite unstable as local faults often must be propagated to all routers. To address these problems, we embarked on a clean-sheet redesign of BGP. The resulting protocol, called HLP, was designed with robustness as a primary goal. It uses a hybrid of link-state and path-vector to achieve much better stability and convergence properties. A prototype has been implemented in XORP.

Coping with Middleboxes (2004):
Middleboxes, as currently deployed, violate IP semantics (by intercepting packets not addressed to them) and limit the flexibility of the Internet. Some have called for their elimination, but middleboxes do provide valuable functionality (or else they wouldn't exist). We developed a delegation-oriented architecture (DOA) that gracefully incorporates middleboxes into the architecture, so that they no longer violate IP semantics but can still provide the desired services. This requires no change to IP, or routers.

Persistence (2004):
One of the common ways that web applications fail is that the data they seek is no longer where they thought it was. This stems from the lack of persistent names in the Internet. Data is named in terms of URLs, which are based on a hostname; the data, however, might move to another host, thereby invalidating its name. IP addresses aren't persistent in the face of mobility and renumbering.

We have developed a layered naming architecture for the Internet that, like DOA, does not involve changes to IP, but does enable seamless movement of data and hosts. We have applied this idea to the web in the form of the Semantic-Free Referencing design.

Lightweight Fairness Mechanisms (2003):
The prevailing paradigm in congestion control is that end hosts implement control mechanisms and routers (either through packet drops or ECN bits) merely signal congestion. However, this paradigm is vulnerable to congestion control 'cheaters' who don't use the proscribed congestion control mechanism. One way to cope with congestion control cheaters is to have routers manage their bandwidth allocations; initial approaches to this have used per-flow state and per-flow queueing. More recent work, including RED-PD, CSFQ, and AFD, used much lighter-weight mechanisms to control bandwidth allocations. Our research has considered a variety of approaches to this problem, and tried to isolate the theoretical issues inherent

in such allocations.

Role-Based Architectures (2003):
Questioning whether layering is still an adequate foundation for robust networking architectures, this project investigates non-layered approaches to the design and implementation of network protocols. The goals are greater flexibility and control with fewer feature interaction problems, thereby leading to greater robustness. We propose a specific non-layered paradigm called role-based architecture.

Designing Robust Protocols (2003):
Robustness has long been a central design goal of the Internet. Much of the initial effort towards robustness focused on the 'fail-stop' model, where node failures are complete and easily detectable by other nodes. The Internet is quite robust against such failures, routinely surviving various catastrophes with only limited outages. This robustness is largely due to the widespread belief in a set of guidelines for critical design decisions such as where to initiate recovery and how to maintain state. However, the Internet remains extremely vulnerable to more arbitrary failures where, through either error or malice, a node issues syntactically correct responses that are not semantically correct. Such failures, some as simple as misconfigured routing state, can seriously undermine the functioning of the Internet. With the Internet playing such a central role in the global telecommunications infrastructure, this level of vulnerability is no longer acceptable. To make the Internet more robust to these kinds of arbitrary failures, we need to change the way we design network protocols. In this activity, we seek general guidelines or rules that will lead to more robust protocols.

**Coping with Distributed Stresses:

Hardware Support for Network Intrusion Detection:
With ever-increasing network speeds and traffic volumes, there is a growing need to support network intrusion detection using custom hardware. Yet designing such hardware in a fashion that is both robust and sufficiently flexible takes great care. Our recent efforts to address this need have been in two different areas, shunting and stream reassembly.

The shunting project aims to develop an architecture that combines the power of high speed network elements with the flexibility of highly programmable network intrusion detection systems (NIDSs). The core of the architecture is a network forwarding element (the ''shunt'') that works in conjunction with a NIDS by diverting a subset of the traffic stream through the NIDS. Because the NIDS receives the actual traffic itself rather than a copy, the architecture enables the NIDS to instantly block attack traffic (i.e., 'intrusion prevention'). The key insight leveraged by architecture is that in many environments, the vast majority of the high-volume traffic is confined to a small fraction of the connections. Furthermore, these high-volume connections are generally of little interest from an intrusion-detection perspective after they have been initially established. That is, it is important to analyze the connections' surrounding context (control session, initial authentication dialog, concurrent logins, etc.), but, once established, the connections themselves can be safely skipped. By basing the project's first prototype on 1 Gbps FPGA forwarding elements, we target enabling the architecture to scale to 10 Gbps in the near-term as next-generation FPGAs become available, and then to 40 Gbps based on striping the traffic across four 10 Gbps elements.

The stream reassembly project investigates the design of a hardware module to provide the basic operation of correctly reassembling any out-of-sequence packets delivered by an underlying unreliable network protocol such as IP, a task inherent to higher level analysis of traffic carried with a stateful transport protocol such as TCP. This module has applications to hardware support for a range of high-speed network devices performing packet processing at semantic levels above the network layer, such as layer-7 switches, content inspection and transformation systems, and network intrusion detection/prevention systems. This seemingly prosaic task of reassembling the byte stream becomes an order of magnitude more difficult to soundly execute when conducted in the presence of an adversary whose goal is to either subvert the higher-level analysis or impede the operation of legitimate traffic sharing the same network path.

Architecting Independent State for Network Intrusion Detection Systems:
Network intrusion detection systems (NIDS) rely on managing a great amount of state for tracking active connections and the specifics of behavior observed in the past. Often much of this state resides solely in the volatile processor memory accessible to a single user-level process on a single machine. In this ongoing work we developed with colleagues an architecture that facilitates independent state, i.e. internal fine-grained state that can be propagated from one instance of a NIDS to others running either concurrently or subsequently. Our unified architecture offers a wealth of possible applications that hold the potential to greatly enhance the power of a NIDS; we are exploring examples such as distributed processing, load parallelization, sharing attack information between different sites, controlling loss of state across restarts, dynamic reconfiguration, high-level policy maintenance, and support for profiling and debugging.

Building a Time Machine:
Insight into past network traffic can have enormous value, both for forensics when analyzing a problem detected belatedly, and to augment real-time decision-making, both to inform reactive measurement and to give additional pinpoint context to a network intrusion detection system (NIDS). This project aims to develop a network time machine, which works by passively bulk-recording as much network traffic as possible. The time machine maintains a ring buffer of recent network traffic that matches a given criteria. A major next step for this project is to integrate it with a real-time NIDS by providing an API by which the NIDS can query activity seen in the recent past for given connections or hosts. This coupling has the potential to greatly offload the NIDS, allowing it to process only lighter-weight request streams and not response streams, unless it sees a problematic request, in which case it can at that point ask the time machine for a copy of the reply to that particular request.

Integrating Traffic Sampling into Intrusion Detection:
Techniques to sample network traffic have seen a flurry of recent advancement in support of network measurement. On the other hand, historically sampling has been viewed as of little utility for network intrusion detection because attacks are generally a minor component of a traffic stream, and thus sampling that traffic stream is likely to diminish the available analysis signal rather than

augment it. This project investigates the application of sampling techniques to enhance intrusion detection. First, we are looking at ways to characterize different forms of ``large'' traffic flows. Some of these flows are of direct interest for detecting attacks---for example, rapidly discovering traffic floods can enable operators to take steps to ameliorate both damage to the victim and also excessive load on the intrusion detection system's analysis components. In addition, a general strategy we pursue in our network intrusion detection research is to find efficient mechanisms for detecting activity expressed in more abstract terms, whether benign or malicious. Such information can often augment the power of high-level security analysis by providing additional context.

Modeling Enterprise Traffic:
The characteristics of network traffic within an enterprise have gone unexamined in the literature for more than a decade. This project aims to develop such a characterization for modern Internet traffic, as recorded internal to the Lawrence Berkeley National Laboratory. Such basic questions as ``What are the dominant types of traffic'' and ``How do the traffic patterns differ form wide-area Internet traffic'' remain unanswered. Thus, this effort has the potential to yield many interesting and possibly surprising insights.

Minimizing Attacks via Tracking Behavior:
We have developed a high-level architecture for large-scale sharing of past behavioral patterns about network actors (e.g., hosts or email addresses) in an effort to inform policy decisions about how to treat future interactions. In our system, entities can submit reports of certain observed behavior (particularly attacks) to a distributed database. When decidng whether to provide services to a given actor, users can then consult the database to obtain a global history of the actor's past activity. Three key elements of our system are: (i) we do not require a hard-and-fast notion of identity, (ii) we presume that users make local decisions regarding the reputations developed by the contributors to the system as the basis of the trust to place in the information, and (iii) we envision enabling witnesses to attest that certain activity was observed without requiring the witnesses to agree asto the behavioral meaning of the activity.

Analyzing the Thread of Internet Worms (2004):
The ability of attackers to rapidly gain control of vast numbers of Internet hosts using automated 'worms' poses an immense risk to the overall security of the Internet. Over the past year, we have contributed substantially to the understanding of worms and the magnitude of the threat they post, and explored possible defenses. We have formulated and published a taxonomy of the general problem space. In order to assess the true potential magnitude of the threat, we have developed a defensible estimate of a 'worst-case' malicious worm attack, finding that such a worm could plausibly inflict $50B damage to the U.S. economy. We have developed a significantly improved 'containment' algorithm, able to efficiently detect scanning worms and quench their propagation within an enterprise. We have further enhanced containment through *cooperation*: using the early findings of distributed detectors to heighten the sensitivity of subsequent response. We have analyzed where defenses need to be placed, concluding that significant components of a worm defense will need to be embedded in the network element of enterprise LANs, creating 'Hard-LANs' designed to detect and respond to worm attacks. We have pursued development of 'honeyfarms' -- honeypots that respond to traffic sent to unused Internet addresses in order to detect new forms of malicious activity

-- and used a prototype deployment to perform the first study of Internet attack 'background radiation'. Finally, we have offered an increased understanding of 'scaledown' for simulating Internet-scale activity, and applied this to developing a deeper understanding of the propagation behavior of the Slammer worm.

Internet Background Radiation:
Monitoring any portion of the Internet address space reveals incessant activity. This holds even when monitoring traffic sent to unused addresses, which we term ``background radiation.'' Background radiation reflects fundamentally nonproductive traffic, either malicious (flooding backscatter, scans for vulnerabilities, worms) or benign (misconfigurations). In this collaborative project we developed the first systematic, broad characterization of the Internet's current background radiation. We based our characterizations on data collected from four unused networks in the Internet. Two key elements of our methodology were (i) the use of filtering to reduce load on the measurement system, and (ii) the use of active responders to elicit further activity from scanners in order to differentiate different types of background radiation.

Enhanced Network Telescope Analysis:
Network ``telescopes'', which record packets sent to unused blocks of Internet address space, have emerged as an important tool for observing Internet-scale events such as the spread of worms and the backscatter from flooding attacks that use spoofed source addresses. Previous telescope analyses have produced detailed tabulations of packet rates, victim population, and evolution over time. While such cataloging is a crucial first step in studying the telescope observations, incorporating an understanding of the underlying processes generating the observations allows us to construct detailed inferences about the broader ``universe'' in which the Internet-scale activity occurs, greatly enriching and deepening the analysis in the process.

Investigating Scaledown for Analyzing Large-Scale Internet Phenomena:
A major challenge when attempting to analyze and model large-scale Internet phenomena such as the dynamics of global worm propagation is finding appropriate abstractions that allow us to tractably grapple with size of the artifact while still capturing its most salient properties. In this project we investigated ``scaledown'' techniques for approximating global Internet worm dynamics by shrinking the effective size of the network under study, exploring scaledown with both simulation and analysis. We demonstrated the viability of scaledown, but also explore two important artifacts it introduces: heightened variability of results, and biasing the worm towards earlier propagation.

Lower Bounds on Speed of Worm Propagation:
Flash worms follow a precomputed spread tree using prior knowledge of all systems vulnerable to the worm's exploit. Our previous work suggested that a flash worm could saturate one million vulnerable hosts on the Internet in under 30 seconds. In this effort, we revisited this problem in the context of single-packet UDP worms (such as Slammer and Witty), using current Internet latency measurements for calibration. We examined the implications of flash worms for containment defenses, finding that such defenses must correlate information from multiple sites in order to detect the worm, but the speed of the worm will defeat this correlation unless a certain fraction of traffic is artificially delayed in case it later proves to be a worm.

Upper Bounds on Worm Damage:
In this collaborative effort, we explored the question of to what degree do worms potentially represent a substantial economic threat to the U.S. computing infrastructure. An estimate of how much damage might be caused can greatly aid in evaluating how much to spend on defenses. We constructed a parameterized worst-case analysis based on a simple damage model, combined with our understanding of what an attack could accomplish.

Detecting Triggers:
Many automated network attacks have the form in which an initial connection to a host triggers (upon success) a subsequent connection, either inbound from the attacker to test whether a ``back door'' has been established, or outbound from the victim to signal to the attacker that the exploit succeeded. This project aims to develop statistical anomaly detection techniques for detecting such attacks. We view the problem in abstract terms as attempting to identify pairs of connections that are causally related. By grouping connections related to the same application into sessions, we are able to formulate a model based on session-arrival statistics to determine when a new arrival that is separate from an existing session has occurred implausibly soon after a previous arrival. This forms the basis for detecting causalities other than those that arise naturally from expected application patterns. However, a major challenge with this work is determining the full set of such expected application patterns.

Research Coupled with Operational Intrusion Detection:
There is a world of difference between intrusion detection research as explored in a computer science department lab versus the real-world problems encountered with 24x7 intrusion detection operation at a busy site. This on-going project, in collaboration with the System and Network Security groups at the Lawrence Berkeley National Laboratory, the Technical University of Munich, and the University of California, Berkeley, centers on research and development in support of the 24x7 use of the Bro intrusion detection system as a primary component of site security at those institutions. As part of this effort, we aimed to address a dearth in the research literature of examinations of the practical difficulties of operating network intrusion detection systems (NIDSs) in large-scale environments, especially the extreme challenges with respect to traffic volume, traffic diversity, and resource management. Our results both helped us gauge the trade-offs of tuning a NIDS and led us to explore several novel ways of reducing resource requirements. These new techniques enabled us to improve the state management considerably, as well as balancing the processing load dynamically.

Bro Advanced Development:
The Bro intrusion detection system has served as the basis for ongoing intrusion detection research since the mid 1990s. One of the strengths of the development of Bro has been its ongoing operational use, first at the Lawrence Berkeley National Laboratory (LBNL), and now in addition at the Technical University of Munich and U.C. Berkeley. The ``Bro Lite'' project, in collaboration with LBNL, aims to develop a version of Bro amenable to broader, production use at sites that do not necessarily have any Bro expertise. Major elements of this effort include (i) enhanced documentation, (ii) turn-key operation, (iii) augmentation of Bro's signature matching facilities, a concept already familiar to many operators, (iv) development of a Graphical User Interface for exploring the logs Bro produces, (v) production-quality support for

Bro in terms of bug-tracking, web presence, and mailing list.

Scan Detection and Worm Containment:
Attackers routinely perform random ``portscans'' of Internet addresses to find vulnerable servers to compromise. Network intrusion detection systems (NIDS) attempt to detect such behavior and flag these portscanners as malicious. An important need in such systems is prompt response: the sooner a NIDS detects malice, the lower the resulting damage. At the same time, a NIDS should not falsely implicate benign remote hosts as malicious. Balancing the goals of promptness and accuracy in detecting malicious scanners is a delicate and difficult task. With colleagues we have developed a connection between this problem and the theory of sequential hypothesis testing and showed that one can model accesses to local Internet addresses as a random walk on one of two stochastic processes, corresponding respectively to the access patterns of benign remote hosts and malicious ones. The detection problem then becomes one of observing a particular trajectory and inferring from it the most likely classification for the remote host. We used this insight to develop Threshold Random Walk (TRW), a novel on-line detection algorithm that identifies malicious remote hosts. Using an analysis of traces from two qualitatively different sites, we showed that TRW requires a much smaller number of connection attempts (4 or 5 in practice) to detect malicious activity compared to previous schemes, while also providing theoretical bounds on the low (and configurable probabilities of missed detection and false alarms.

Effect of 9-11 (2003):
A study by the Computer Science and Telecommunications Board of the National Research Council on the effect on the Internet of the September 11, 2001, terrorist attacks, especially the destruction of the World Trade Center towers.

**Measurements for Robustness:

Sound Internet Measurement:
Conducting an Internet measurement study in a sound fashion can be much more difficult than it might first appear. In this effort we endeavored to formulate a number of strategies drawn from experiences for avoiding or overcoming some of the associated pitfalls. Some of these in particular included dealing with errors and inaccuracies, the importance of associating meta-data with measurements, the technique of calibrating measurements by examining outliers and testing for consistencies, difficulties that arise with large-scale measurements, the utility of developing a discipline for reliably reproducing analysis results, and issues with making datasets publicly available.

Reactive Measurement:
Reactive measurement (REM) is a measurement technique in which one measurement's results are used to decide what (if any) additional measurements are required to further understand some observed phenomenon. While reactive measurement has been used on occasion in measurement studies, what has been lacking is (i) an examination of its general power, and (ii) a generic framework for facilitating fluid use of this approach. This project undertakes to explore REM's power, and develop and implement an architecture for a system that provides general REM functionality to the network measurement community. We believe that by enabling the coupling of disparate measurement tools, REM holds great promise for assisting researchers

and operators in determining the root causes of network problems and enabling measurement targeted for specific conditions.
We are currently building a generic REM system that will be released to the research community and can be used to bind together disparate measurement tools as the basis for deeper understanding into the network's behavior.

NIMI/SAMI:
The NIMI (National Internet Measurement Infrastructure) project focuses on developing and deploying a system for facilitating coordinated measurement from a number of points around the Internet. The final efforts in this project have been towards SAMI (Secure Architecture for Measurement Infrastructures), a major revision of the NIMI architecture that aims to refine its authorization, security, and resource control mechanisms. SAMI is now close to final release.

Active Measurements of TCP (2004):
There are a range of TCP congestion control behaviors in deployed TCP implementations, include Tahoe, Reno, NewReno, and Sack TCP, which date from 1988, 1990, 1996, and 1996, respectively, along with other capabilities such as ECN (Explicit Congestion Notification), timestamps, and the like. In order to better understand the migration of new congestion control mechanisms to the public Internet, we developed a tool, called the TCP Behavior Inference Tool (TBIT), for exploring the TCP behaviors of web servers. Apart from testing for the congestion control ``flavor'', TBIT can test for correctness of SACK implementation, use of timestamps, ECN capabilities and many other TCP behaviors. TBIT has been significantly expanded, and a wide range of new tests have been run, characterizing the TCP behavior of more than 80,000 web servers. Many of the new tests explore the impact of middleboxes along the path of the end-to-end TCP connection, in terms of robustness in the presense of IP options or of ECN, or with the use of ICMP for path MTU discovery.

Findings:

**Coping with Extreme Environments:

DCCP:
The design of DCCP is fairly mature, and the specifications for DCCP had been approved for publication as RFCs for Proposed Standard. Documents are in progress for congestion control mechanisms specifically tailored to the needs of streaming audio and video. The next step will be to learn more from real-world experience with the protocol.

High Speed TCP (2004):
HighSpeed TCP has appeared as an Experimental RFC, and there are a number of implementations and reports of experimental and simulation results, many comparing HighSpeed TCP with other proposals for TCP in high-bandwidth environments. HighSpeed TCP shows that TCP can be easily adapted to cope with significantly higher speeds, even with default packet sizes.

Transport's Robustness to Reordering:
We have published an RFC that specifies the use of TCP's DSACK option and SCTP's Duplicate-TSN mechanism to detect spurious retransmits, as one tool to aid in robustness to reordering.

We have also published an Internet Draft outlining a mechanism for TCP connections to better distinguish between loss and reordering.

Transport’s Robustness to Corruption:
We have conducted a broad range of simulations on the applicability of CETEN mechanisms to aid in robustness in the presence of corruption. An initial outline of CETEN is given in a Computer Networks paper. A more in-depth treatment is found in a CCR paper from October 2004.

Transport’s Robustness to Burstiness:
We have published two papers on transport protocol’s robustness to burstiness. The first paper considers the impact of bursting in traffic passively observed in three different networks. The general finding is quite intuitive: that small bursts are omni-present but rarely cause problems, while large bursts are rare but nearly always cause problems. The second paper investigates mechanisms to cope with bursting and to increase TCP’s robustness in situations where bursting will hurt performance or impact the network.

QuickStart:
This design has been described in a series of Internet Drafts and discussed extensively in the IETF, and we expect it to be approved by the Working Group as Experimental shortly. We have also written a paper evaluating the costs and benefits of Quick-Start.

**Coping with New Forms of Failure:

Lightweight Security Mechanisms for BGP (2004):
We developed a series of mechanisms, called Listen and Whisper, of increasing complexity that deal with attacks of increasing sophistication. One mechanism (Listen) involves probing of data paths. The other mechanisms (Whisper) involve comparing route information along multiple paths, using redundancy and cryptographic one-way functions instead of shared key cryptography to establish the validity of a route advertisement. Although these mechanisms do not achieve perfect security, they do provide much better security than what exists today. They are easily deployable and do not require a key distribution infrastructure. However, even these measures are not sufficient against colluding attackers; here, we must augment our arsenal with proposed changes to acceptable BGP policies. This mechanism has been implemented and more throughly explored. We have also established a more formal basis for these methods, proving how it applies to other routing styles.

Robust Interdomain Routing (2004):
The HLP routing design has been implemented in XORP, an open-source routing platform (developed at ICSI). The HLP design has been tested, and its policies properties explored. We find that HLP can implement most of the policies in use today.

Coping with Middleboxes (2004):
A delegation-oriented architecture (DOA) has been implemented, and used to implement a ’network extender box’ and a ’network filtering box’, which are the architecturally clean versions of NATs and firewalls. These implementations show how DOA allows for graceful accommodation of middlebox functionality without violating IP semantics.

Persistence (2004):
The Semantic-Free Referencing design (SFR) has been implemented and

tested. We find that the scheme can achieve latencies not much worse than today's DNS-based web. Moreover, the SFR design has been incorporated into the broader LFN architecture.

Lightweight Fairness Mechanisms (2003):
The original Approximate Fair Dropping (AFD) proposal required substantial state. A new design, using only a flow table rather than a shadow buffer, matches the performance of the AFD proposal but requires far less state. Also, some of the theoretical underpinnings of fairness algorithms has been described in terms of 'iceberg' queries.

Role-Based Architectures (2003):
An initial design for a role-based architecture was presented at HotNets.

Designing Robust Protocols (2003):
We proposed a set of six design guidelines for improving the network protocol design. These guidelines emerged from a study of past examples of failures, and determining what could have been done to prevent the problem from occurring in the first place. The unifying theme behind the various guidelines is that we need to design protocols more defensively, expecting malicious attack, misimplementation, and misconfiguration at every turn.

**Coping with Distributed Stresses'

Hardware Support for Network Intrusion Detection:
The stream reassembly project designed a hardware module to reassemble out-of-sequence packets that is robust against attacks from adversaries.

Minimizing Attacks via Tracking Behavior:
We have presented a high-level architecture and started a dialog within the research community about a system for tracking the behavior of various actors within the Internet.

Analyzing the Thread of Internet Worms (2004):
The ability of attackers to rapidly gain control of vast numbers of Internet hosts using automated 'worms' poses an immense risk to the overall security of the Internet. We have contributed substantially to the understanding of worms and thesmagnitude of the threat they post, and explored possible defenses.

Internet Background Radiation:
We analyzed the components of background radiation by protocol, application, and specific exploits, assessed temporal patterns and correlated activity, and investigated variations across different networks and over time. We found a menagerie of activity, with probes from worms and ``autorooters'' heavily dominating.

Enhanced Network Telescope Analysis:
We applied our enhanced network telescope analysis to the propagation of the Witty worm, a malicious and well-engineered worm that when released in March 2004 infected more than 12,000 hosts worldwide in 75 minutes. We found that by carefully exploiting the structure of the worm, especially its pseudo-random number generation, from limited and imperfect telescope data we could with high fidelity: extract the individual rate at which each infectee injected packets into the network prior to loss; correct distortions in the telescope data due to the worm's volume overwhelming the monitor; reveal the

worm's inability to fully reach all of its potential victims; determine the number of disks attached to each infected machine; compute when each infectee was last booted, to sub-second accuracy; identify the effects of NAT on the observations; explore the ``who infected whom'' infection tree; uncover that the worm specifically targeted hosts at a US military base; and pinpoint Patient Zero, the initial point of infection, i.e., the IP address of the system the attacker used to unleash Witty.

Lower Bounds on Speed of Worm Propagation:
We found that with careful design, a flash version of Slammer that used a flash worm could saturate 95% of one million vulnerable hosts on the Internet in 510 milliseconds. We also investigated a similar design for a TCP-based worm, finding it could 95% saturate in 1.3 seconds.

Upper Bounds on Worm Damage:
Although our estimates are at best approximations, we argue that a plausible worst-case worm could cause $50 billion or more in direct economic damage by attacking widely-used services in Microsoft Windows and carrying a highly destructive payload.

Scan Detection and Worm Containment:
We developed worm containment algorithms, based on Threshold Random Walks, suitable for deployment in high-speed, low-cost network hardware. We generalized the insights from our investigations to argue that substantial anti-worm defenses will need to be embedded in the local area network, creating ``Hard-LANs'' designed to detect and respond to worm infections. When compared with conventional network-based intrusion detection systems, Hard-LAN devices will require two orders of magnitude better cost/performance, and at least two orders of magnitude better accuracy, resulting in substantial design challenges.

**Measurements for Robustness:

Active measurements of TCP:
The TBIT tests have been significantly extended, with new tests reporting on the failures in robustness when TCP packets encounter middleboxes that block TCP SYN packets with IP Options; TCP SYN packets attempting to negotiage ECN; and ICMP messages needed for path MTU discovery. A paper reporting on these results comments on the difficulties these results imply for new transport protocol mechanisms that propose to use IP options, ICMP messages, or any of the currently-reserved bits in the TCP header. As reported in a second paper, the current tests also explore many aspects of TCP implementations in web servers.

Training and Development:

It has been one of our goals for the project to contribute significantly to the research skills of the graduate students and post-docs working with the senior researchers on the various topics outlined in this report.

Outreach Activities:

ICSI researchers are highly active in the Internet research community. In addition to the normal academic duties of serving on program committees and editorial boards, ICSI researchers devote substantial time to more practical duties associated with the Internet Engineering Task Force (IETF), the Internet Research Task Force (IRTF), and other organizational activities within the research community. In particular, Vern Paxson was Chair of the IRTF from 2001 to 2005, Sally Floyd was a member of the IAB (the Internet Architecture Board) from 2001 to 2005, and Mark Allman is currently chair of the IRTF Internet Measurement Research Group and of the IETF TCP Maintenance and Minor Extensions Working Groups. More generally, instead of just publishing our work in academic journals and conferences, much of our work is disseminated via Internet Drafts and RFCs.

Our outreach activities also include the following:

2004-2005:

V. Paxson, invited speaker, National Science Foundation's Workshop on Disruptive Networking Research, 2004.

V. Paxson, co-teacher, Tutorial on cybersecurity for open scientific institutes, IEEE Supercomputing, November 2004.

2003-2004:

V. Paxson, invited keynote, 'Measuring Adversaries', ACM SIGMETRICS 2004.

V. Paxson, invited panelist, 'Some Thoughts on Where Worms are Going', Symposium on Recent Advances in Intrusion Detection (RAID 2003).

V. Paxson, invited speaker, 'Bro: an Intrusion Detection System Operating at Gigabit Speed', First International Grid Networking Workshop, March 2004.

V. Paxson, invited speaker, 'High-speed Intrusion Detection and Response', National Coordination Office for Information Technology Research and Development's JET Roadmap Workshop, April 2004.

V. Paxson, invited speaker, 'Attack Activity Substrate', Intel Research, Berkeley, 2003.

V. Paxson, invited speaker, 'Techniques and Issues for Anonymizing Network Traces', DHS workshop on Datasets for Network Security, 2004.

V. Paxson, invited speaker, 'Bolt-On Security: A Networking Perspective', DARPA ISAT study on Bolt-On Security, 2004.

V. Paxson, invited speaker, 'Experiences With Internet Traffic Measurement and Analysis', NTT Research, 2004.

V. Paxson, seminar speaker, 'The Threat of Internet Worms', U.C. Berkeley, 2003, and Stanford University, 2004.

M. Allman, seminar speaker, 'New Techniques for Making TCP Robust to Corruption-Based Loss', Case Western Reserve University, October 2003.

M. Allman, seminar speaker, 'Improving Performance of Internet

Protocols Over Wireless Networks', Kent State University, November 2003.

S. Floyd, invited speaker, 'Thoughts on the Evolution of TCP in the Internet', Second International Workshop on Protocols for Fast Long-Distance Networks (PFLDnet 2004), February 2004.

S. Floyd, invited panelist, 'High Speed Congestion Control', Second Workshop on Hot Topics in Networks (Hotnets-II), November 2003.

2002-2003:

S. Floyd, keynote talk on 'HighSpeed TCP and Quick-Start for Fast Long-Distance Networks', Workshop on Protocols for Fast Long-Distance Networks, Geneva, Switzerland, February 2003.

V. Paxson, lecture on 'Detecting Network Intruders in Real Time'. at University of California, Berkeley. April 2003

V. Paxson, invited speaker on the threat of large-scale worms, DARPA, Mar. 2003; Georgia Tech, Nov. 2002; Intel Research, Nov. 2002; and HP Labs, Feb. 2003.

V. Paxson, presentation on the threat of large-scale worms, IRTF End-to-End Task Force meeting, Feb. 2003.

V. Paxson, seminar on Internet measurement difficulties, HP Labs,

Journal Publications:

T. Anderson, S. Shenker, I. Stoica, and D. Wetherall, "Design Guidelines for Robust Internet Protocols", *ACM SIGCOMM Computer Communication Review First Workshop on Hot Topics in Networks (HotNets-I)*, vol. 33, (2003), p. 125. Published

B. Braden, T. Faber, and M. Handley, "From Protocol Stack to Protocol Heap: Role-based Architecture", *ACM SIGCOMM Computer Communication Review First Workshop on Hot Topics in Networks (HotNets-I)*, vol. 33, (2003), p. 1. Published

E. Kohler, J. Li, V. Paxson and S. Shenker, "Observed Structure of Addresses in IP Traffic", *Proceedings of the ACM SIGCOMM Workshop on Internet Measurement workshop (IMW 2002)*, vol. 2, (2002), p. 253. Published

R. Pan, L. Breslau, B. Prabhakar, and S. Shenker, "A Flow Table-Based Design to Approximate Fairness", *In Hot Interconnects: 10th Symposium on High Performance Interconnects*, vol. 10, (2002), p. 37. Published

S. Staniford, V. Paxson and N. Weaver, "How to 0wn the Internet in Your Spare Time", *Proceedings of the 11th USENIX Security Symposium*, vol. 11, (2002), p. 149. Published

J. M. Gonzalez and V. Paxson, "pktd: A Packet Capture and Injection Daemon", *Proc. Passive & Active Measurement: PAM-2003*, vol. (none), (2003), p. 87. Published

R. M. Karp, S. Shenker, and C. H. Papadimitriou, "A Simple Algorithm for Finding Frequent Elements in Streams and Bags", *Transactions on Database Systems*, vol. 28, (2003), p. 51. Published

U. Shankar and V. Paxson, "Active Mapping: Resisting NIDS Evasion Without Altering Traffic", *Proc. IEEE Symposium on Security and Privacy*, vol. (2003), (2003), p. 44. Published

I. Stoica, S. Shenker, and H. Zhang, "Core-Stateless Fair Queueing: a Scalable Architecture to

Approximate Fair Bandwidth Allocations in High-speed Networks", *IEEE/ACM Transactions on Networking*, vol. 11(1), (2003), p. 33. Published

E. Kohler, M. Handley, and S. Floyd, "Designing DCCP: Congestion Control Without Reliability", *Manuscript*, vol. (none), (2003), p. (none). URL "http://www.icir.org/kohler/dcp/"

T. Kelly, S. Floyd, and S. Shenker, "Patterns of Congestion Collapse", *IEEE ACM Transactions on Networking*, vol. , (2004), p. . Accepted

J. Feigenbaum, R. Sami, and S. Shenker, "Mechanism Design for Policy Routing", *Proceedings of the 23rd Symposium on Principles of Distributed Computing (PODC)*, vol. 23, (2004), p. 11. Published

M. Walfish, H. Balakrishnan, and S. Shenker, "Untangling the Web from DNS", *1st Symposium on Networked Systems Design and Implementation (NSDI)*, vol. 1, (2004), p. ?. Published

H. Balakrishnan, K. Lakshminarayanan, S. Ratnasamy, S. Shenker, I. Stoica, and M. Walfish., "A Layered Naming Architecture for the Internet", *ACM SIGCOMM 2004*, vol. 34, (2004), p. 343. Published

H. Chang, R. Govindan, S. Jamin, S. Shenker and W. Willinger, "Towards Capturing Representative AS-level Internet Topologies", *Computer Networks*, vol. 44, 6, (2004), p. 737. Published

R. Huebsch, J. Hellerstein, N. Lanham, B. Thau Loo, S. Shenker, and I. Stoica, "Querying the Internet with PIER", *Proceedings of VLDB*, vol. 19, (2003), p. ?. Published

L. Subramanian, V. Roth, I. Stoica, S. Shenker and R. H. Katz, "Listen and Whisper: Security Mechanisms for BGP", *First Symposium on Networked Systems Design and Implementation (NSDI'04)*, vol. 1, (2004), p. ?. Published

I. Stoica, D. Adkins, S. Zhaung, S. Shenker, and S. Surana, "Internet Indirection Infrastructure", *IEEE/ACM Transactions on Networking*, vol. 12, (2004), p. 205. Published

M. Walfish, J. Stribling, M. Krohn, H. Balakrishnan, R. Morris, and S. Shenker, "Middleboxes No Longer Considered Harmful :", *OSDI '04*, vol. (none), (2004), p. ?. Published

K. Argyraki, P. Maniatis, D. Cheriton, and S. Shenker, "Providing Packet Obituaries", *Hotnets III*, vol. III, (2004), p. ?. Published

K. Lakshminarayanan, I. Stoica, and S. Shenker, "Routing as a Service", *None*, vol. , (), p. . in progress

J. M. Hellerstein, V. Paxson, L. Peterson, T. Roscoe, S. Shenker, and D. Wetherall, "The Network Oracle", *None*, vol. , (), p. . in progress

R. Gummadi, N. Kothari, Y. J. Kim, R. Govindan, B. Karp, and S. Shenker, "Reduced State Routing in the Internet", *Hotnets III*, vol. III, (2004), p. ?. Published

L. Peterson, S. Shenker, and J. Turner, "Overcoming the Internet Impasse through Virtualization", *Hotnets III*, vol. III, (2004), p. ?. Published

L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica, "Towards a Next Generation Inter-domain Routing Protocol", *Hotnets III*, vol. III, (2004), p. ?. Published

W. Eddy, S. Ostermann, and M. Allman, "New Techniques for Making Transport Protocols Robust to Corruption-Based Loss", *ACM Computer Communication Review*, vol. 34(5), (2004), p. ?. Published

R. Krishnan, J. Sterbenz, W. Eddy, C. Partridge, and M. Allman, "Explicit Transport Error Notification (ETEN) for Error-Prone Wireless and Satellite Networks", *Computer Networks*, vol. 46(3), (2004), p. ?. Published

M. Allman and V. Paxson, "A Reactive Measurement Framework", *under submission*, vol. ?, (2004), p. ?. Submitted

N. Weaver, I. Hamadeh, G. Kesidis, and V. Paxson, "Preliminary Results Using Scale-Down to Explore Worm Dynamics", *Proc. ACM CCS WORM '04*, vol. ?, (2004), p. ?. Published

S. Staniford, D. Moore, V. Paxson and N. Weaver, "The Top Speed of Flash Worms", *Proc. ACM CCS WORM '04*, vol. ?, (2004), p. ?. Published

N. Weaver, D. Ellis, S. Staniford and V. Paxson, "Worms vs. Perimeters: The Case for Hard-LANs", *Proc. 12th Annual IEEE Symposium on High Performance Interconnects, 2004*, vol. 12, (2004), p. ?. Published

R. Pang, V. Yegneswaran, P. Barford, V. Paxson and L. Peterson, "Characteristics of Internet Background Radiation", *Proc. ACM SIGCOMM Internet Measurement Conference, 2004*, vol. 4, (2004), p. ?. Published

V. Paxson, "Strategies for Sound Internet Measurement", *Proc. ACM SIGCOMM Internet Measurement Conference, 2004.*, vol. 4, (2004), p. ?. Published

N. Weaver, S. Staniford and V. Paxson, "Very Fast Containment of Scanning Worms", *Proc. USENIX Security Symposium 2004*, vol. ?, (2004), p. ?. Published

N. Weaver and V. Paxson, "A Worst-Case Worm", *Proc. Third Annual Workshop on Economics and Information Security (WEIS04), May 2004*, vol. 3, (2004), p. ?. Published

R. Pang and V. Paxson, "A High-level Programming Environment for Packet Trace Anonymization and Transformation", *Proc. ACM SIGCOMM 2003*, vol. 33, (2003), p. 339. Published

N. Weaver, V. Paxson, S. Staniford and R. Cunningham, "A Taxonomy of Computer Worms", *Proc. ACM CCS Workshop on Rapid Malcode*, vol. 1, (2003), p. ?. Published

D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford and N. Weaver, "Inside the Slammer Worm", *Security and Privacy*, vol. 1, (2003), p. 33. Published

R.Karp, T. Nierhoff and T. Tantau, "Optimal Flow Distribution among Multiple Channels with Unknown Capacities", *Second Brazilian Symposium of Graphs, Algorithms, and Combinatorics (GRACO)*, vol. 2005, (2005), p. . Accepted

M. Zhang, B. Karp, S. Floyd, and L. Peterson, "RR-TCP: A Reordering-Robust TCP with DSACK", *IEEE International Conference on Network Protocols (ICNP'03)*, vol. 11, (2003), p. 95. Published

A. Medina, M. Allman, and S. Floyd, "Measuring Interactions between Transport Protocols and Middleboxes", *Internet Measurement Conference, October 2004*, vol. , (2004), p. ?. Published

M. Allman and E. Blanton, "Notes on Burst Mitigation for Transport Protocols", *ACM Computer Communication Review*, vol. 35(2), (2005), p. ?. Published

A. Medina, M. Allman, and S. Floyd, "Measuring the Evolution of Transport Protocols in the Internet", *ACM Computer Communication Review*, vol. 35(2), (2005), p. 37. Published

P. Sarolahti, M. Allman, and S. Floyd, "Evaluating Quick-Start for TCP", *IEEE Conference on Network Protocols*, vol. , (), p. . Submitted

B. T. Loo, J. Hellerstein, R. Huebsch, S. Shenker, and I. Stoica, "Enhancing P2P File-Sharing with an Internet-Scale Query Processor", *Proceedings of VLDB 2004*, vol. 30, (2004), p. ?. Published

L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica, "HLP: A Next-generation Interdomain Routing Protocol", *SIGCOMM 2005*, vol. 35, (2005), p. . Accepted

S. Ratnasamy, S. Shenker, and S. McCanne, "Towards an Evolvable Internet Architecture", *SIGCOMM 2005*, vol. 35, (2005), p. . Accepted

M. Allman, E. Blanton and V. Paxson, "An Architecture for Developing Behavioral History", *Proc. USENIX Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI)*, vol. 2005, (), p. . Accepted

A. Kumar, V. Paxson and N. Weaver, "Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event ", *ACM IMC 2005*, vol. 2005, (), p. . Submitted

S. Dharmapurikar and V. Paxson, "Robust TCP Stream Reassembly In the Presence of Adversaries", *Proc. USENIX Security Symposium 2005*, vol. 2005, (), p. . Accepted

J. Kannan, J. Jung, V. Paxson and C. E. Koksal, "Detecting Hidden Causality in Network Connections", *ACM CCS 2005*, vol. 2005, (), p. . Submitted

Book(s) of other one-time publications(s):

C. Partridge, P. Barford, D. Clark, S. Donelan, V. Paxson J. Rexford, M. Vernon, J. Eisenberg, M. Blumenthal, D. Padgham, K. Batch, D. Drake, and J. Briscoe., "The Internet Under Crisis Conditions: Learning from September 11" , bibl. National Academy Press, Washington D.C., (2002). *Book* Published
of Collection: Computer Science and Telecommunications Board, National Research Council, ""

S. Floyd and E. Kohler, "Profile for DCCP Congestion Control ID 2: TCP-like Congestion Control" , bibl. Internet-draft draft-ietf-dccp-ccid2-10.txt,, (2005). *Internet Draft* Approved for publication as Proposed Standard

S. Floyd, M. Handley, and E. Kohler., "Problem Statement for DCCP" , bibl. Internet draft draft-ietf-dccp-problem-00.txt, (2002). *Internet Draft* internet-draft

E. Kohler, M. Handley, S. Floyd, and J. Padhye., "Datagram Congestion Control Protocol (DCCP)" , bibl. Internet draft draft-ietf-dccp-spec-11.txt, (2005). *Internet draft* Approved for publication as Proposed Standard

S. Floyd, E. Kohler, and J. Padhye., "Profile for DCCP Congestion Control ID 3: TFRC Congestion Control" , bibl. Internet-draft draft-ietf-dccp-ccid3-11.txt, (2005). *Internet draft* Approved for publication as Proposed Standard

M. Allman, S. Floyd, and C. Partridge., "RFC 3390: Increasing TCP's Initial Window." , bibl. Experimental. Obsoletes RFC 2414., (2002). *RFC* Published

N. Weaver, V. Paxson, S. Staniford, and R. Cunningham, "Large Scale Malicious Code: A Research Agenda." , bibl. DARPA Report, (2003). *Project Report* Published

R. Atkinson and S. Floyd, editors, "RFC 3869: IAB Concerns & Recommendations Regarding Internet Research & Evolution" , bibl. RFC 3869, (2004). *RFC* Published

S. Floyd, "RFC 3649: HighSpeed TCP for Large Congestion Windows" , bibl. Experimental., (2003). *RFC* Published

A. Jain, S. Floyd, M. Allman, and P. Sarolahti, "Quick-Start for TCP and IP" , bibl. internet-draft draft-amit-quick-start-04.txt, work in progress, (2005). *internet draft* internet-draft

M. Handley, S. Floyd, J. Padhye, and J. Widmer., "RFC 3448: TCP Friendly Rate Control (TFRC): Protocol Specification. " , bibl. Proposed Standard, (2003). *RFC* Published

W. Willinger, V. Paxson, R. H. Riedi and M. S. Taqqu, "Long-range Dependence and Data Network Traffic" , bibl. Birkhauser, (2002). *Book* Published
of Collection: P. Doukhan, G. Oppenheim and M. S. Taqqu, eds, "Long-range Dependence: Theory and Applications"

D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford and N. Weaver, "The Spread of the Sapphire/Slammer Worm" , bibl. CAIDA www.caida.org/outreach/papers/ 2003/sapphire/sapphire.html, (2003). *technical report* Published

J. Feigenbaum and S. Shenker, "Distributed Algorithmic Mechanism Design: Recent Results and Future Directions" , bibl. ACM Press, New York, 2002, pp. 1-13., (2002). *Book* Published of Collection: , "Proceedings of the 6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communication (DIALM)"

E. Blanton and M. Allman, "RFC 3708: Using TCP Duplicate Selective Acknowledgement (DSACKs) and Stream Control Transmission Protocol (SCTP) Duplicate Transmission Sequence Numbers (TSNs) to Detect Spurious Retransmissions" , bibl. Experimental, (2004). *RFC* Published

M. Allman, K. Avrachenkov, U. Ayesta and J. Blanton, "Early Retransmit for TCP and SCTP" , bibl. Internet-Draft. draft-allman-tcp-early-rexmt-03.txt, (2003). *internet-draft* internet-draft

K. Chandrayana, Y. Zhang, M. Roughan, S. Sen and R. Karp, "Search Game in Inter-domain Traffic Engineering" , bibl. Manuscript, (2004). *Manuscript*

J. Scott and R.Karp, "Achieving Fairness Through Selective Early Dropping" , bibl. Manuscript, (2004). *Manuscript*

S. Floyd, "RFC 3742: Limited Slow-Start for TCP with Large Congestion Windows" , bibl. Experimental, (2004). *RFC* Published

S. Floyd and J. Kempf, editors, "RFC 3714: IAB Concerns Regarding Congestion Control for Voice Traffic in the Internet" , bibl. Informational, (2004). *RFC* Published

S. Floyd, T. Henderson, and A. Gurtov, "RFC 3782: The NewReno Modification to TCP's Fast Recovery Algorithm" , bibl. Proposed Standard., (2004). *RFC* Published

S. Floyd and E. Kohler, "TCP Friendly Rate Control (TFRC) for Voice: VoIP Variant and Faster Restart" , bibl. internet-draft draft-amit-quick-start-04.txt, work in progress, (2005). *internet-draft* internet-draft

Other Specific Products:

Software (or netware)

```
The TBIT code for testing the TCP behavior in web servers has been
extensively revised.

This is publically available on the TBIT web page at
"http://www.icir.org/tbit/".
```

Software (or netware)

```
tgat 2.0 - released July 2004.
The tgat (traffic generation and tracing) utilities are used for
sourcing and sinking TCP traffic, in addition to capturing the
packets that constitute the transfer.

Available from "http://www.icir.org/mallman/software/tgat/".
```

Internet Dissemination:

"http://www.icir.org/floyd/hstcp.html", "http://www.icir.org/kohler/dccp/", "http://www.icir.org/floyd/quickstart.html", "http://www.icir.org/tbit/"

Web sites for DCCP, HighSpeed TCP, TBIT, and QuickStart.

Contributions:

Contributions within Discipline:

Contributions within Discipline:

Robustness is an oft-spoken goal but an elusive target. Most networking papers focus on increased performance along one dimension or another, because the challenge can be easily articulated and the results precisely quantified. Robustness is a more amorphous goal, seeking not a narrow performance improvement in any particular setting but instead an ability to function 'pretty well' across a wide range of settings. Our research is contributing to both a renewed focus on robustness as a research topic and also specific advances relevant to the Internet.

Contributions to Other Disciplines:

Contributions beyond Science and Engineering:

The central goal of our work is to make the Internet more robust. Any such improvement in robustness will redound to society as a whole. More specifically, our goals are to make the Internet more capable of (1) functioning in extreme environments, such as high speeds and lossy wireless links, (2) coping with new forms of failures, such as misconfigurations and middleboxes, and (3) dealing with distributed stresses both malicious and benign. Society will benefit greatly if we make progress on any of these fronts.

Contributions Beyond Science and Engineering:

Contributions beyond Science and Engineering:

The central goal of our work is to make the Internet more robust. Any such improvement in robustness will redound to society as a whole. More specifically, our goals are to make the Internet more capable of (1) functioning in extreme environments, such as high speeds and lossy wireless links, (2) coping with new forms of failures, such as misconfigurations and middleboxes, and (3) dealing with distributed stresses both malicious and benign. Society will benefit greatly if we make progress on any of these fronts.

Special Requirements for Annual Project Report:

Unobligated funds: less than 20 percent of current funds

Categories for which nothing is reported:
Participants: Partner organizations
Contributions to Education and Human Resources
Contributions to Resources for Science and Technology
Special Reporting Requirements
Animal, Human Subjects, Biohazards