Report on DIMACS* Workshop on Large-scale Internet Attacks

Date of workshop: September 23 – 24, 2003

Workshop Organizers:

Vern Paxson, ICSI Center for Internet Research
Steve Bellovin, AT&T Labs-Research
Stuart Staniford, Silicon Defense
Stefan Savage, UC San Diego

Report Author:

Xuhui Ao, Dept. of Computer Science, Rutgers University
ao@cs.rutgers.edu
Date of report: November 30, 2003

1 Workshop Focus

With the increasing size of the Internet, we have seen an increasing number of attacks that take advantage of the network's large scale. These kind of large-scale Internet attacks are usually difficult to counter because of the difficulties in tracing them back or deploying widespread defensive measures. This workshop explored four general types of large-scale attacks and the possible countermeasures:

(1) Distributed Denial of Service (DDoS), in which collections of hundreds or thousands of compromised machines are coordinated to simultaneously send floods of bogus traffic towards a target, completely overwhelming the target's resources, or those of the target's network;

*DIMACS was founded as a National Science Foundation Science and Technology Center. It is a joint project of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Laboratories America, and Telcordia Technologies, with affiliated partners Avaya Labs, IBM Research, Microsoft Research, and HP Labs.

(2) Self-propagating Malicious Code, or Worms, which have in recent years compromised hundreds of thousands of Internet hosts in a matter of hours (with recent work arguing that future worms will likely be even more rapid, and/or much harder to detect);

(3) Infrastructure Attacks, which attempt to subvert the key components of the Internet's underlying infrastructure (domain name system, routing);

(4) Attacks on Large-scale Services, which take advantage of the fact that the Internet's growth has seen the rise of some very large, publicly accessible services (such as portals, search engines, and auctions), which gain their utility by their very scale, but generally do so by making access to the service extremely cheap and thus open to a new class of sophisticated, highly automated attacks.

2 Summary of the Presentations

2.1 A Large-scale View of Large-scale Attacks, S. Donalen, SBC Internet Services

Dr. Donalen talked about experience with attacks from the point of view of an Internet service provider. He began his talk with an observation that one third of his customers have the experience of being affected by worms and viruses. Then he discussed how to measure a large-scale attack, which can be based on the number of the attacks, number of the targets or victims, the damage that the attack causes, or the effort required to repair that damage. Several types of attack were also discussed, such as the attack on the users, on the shared infrastructure or even the network infrastructure itself, which includes the routing attack, naming attack and timing attack.

Dr. Donalen further observed that the number of reported Internet security incidents (attacks) maintains a constant percentage with regard to the total number of the machines in the Internet, which implies that our security measures haven't been improved much compared with what we had before. Another interesting observation Dr. Donalen made is that the percentage of problems in the Internet caused by software bugs is approximately the same as that caused by attacks.

Finally, Dr. Donalen discussed his experience with how to push customers to fix the problems caused by the attacks. In general, it seems to be difficult to push the customers to fix the problems. Approximately 40 percent of the customer will install the patches after the pushing. While 95 percent of the customers will fix the problem eventually, there is always 5 percent remaining having the problem. Dr. Donalen concluded his talk with

the question of how to divide the responsibility or obligation between the network service providers and their customers with regard to the response to the attacks, and which kind of business service model is needed here.

2.2 Attacks on Content Delivery Network, Name and Affiliation of the Speaker Withheld by Request of His Company

The speaker discussed the experience of the large-scale attacks as the content delivery network provider, which supports thousands of networks. The major traffic in that network is HTTP, HTTPs, DNS and Stream content. He described two kinds of attacks: the first one is to the customers of the content network, which are the victims of most denial of service attacks. The second one is to the content delivery network infrastructure itself, which include the DNS attack, BGP routing attack and other attacks to the content network nodes. He finally talked about various approaches to defend against those attacks.

2.3 Experience with DDoS, Name and Affiliation of the Speaker Withheld by Request of His Company

Our next speaker talked about his experience with the distributed denial of service (DDoS) attacks at a large Internet service. He at first introduced various DDoS attacks they experienced, including the ICMP ping, SYN and the application level attacks. Much ping traffic is caused by the users of the service trying to test the availability of the service and that kind of traffic will not cause big problems because of the Juniper rate limiting mechanism. As to the SYN attack, smaller scale attacks (in 10 Mbps) happen once or twice on average per month, larger scale attacks (in 100 Mbps) happen about three times in the most recent year and very large attacks (in excess of 1 Gbps) haven't been seen yet. The service also experienced application level DDoS attacks, with occasional coordinated attacks.

The speaker then discussed how to deal with those DDoS attacks. Technically, they defended against those attacks by the large scale of the service. He also pointed out that currently the cooperation from the ISPs is low due to the variety and the complicated international nature of the ISPs. There is the need to have a formal collaboration between different sites, ISPs and various agencies, which can automate the coordination among related parties in the case of being attacked, otherwise we may see worse case scenarios in the future.

2.4 Infrastructure Attack Trends, Craig Labovitz, Arbor Networks

Dr. Labovitz discussed the attack trends on the network infrastructure as a network provider providing network availability to the service providers. The main task of such a provider is to protect both its customers and the network infrastructure from various attacks, such as DDoS and Worm attacks. It deployed the distributed monitoring probes among the backbone routers and can monitor various attacks by analyzing the collected information. Dr. Labovitz discussed several observations about the attacks. First, the "Botnet" attacks became the mainstream of the attacks, which consist of thousands of coordinated and compromised hosts. Second, there were increasing attacks against the routing ports and the routers. Third, there was clear evidence that there were a significant number of compromised edge routers in the Internet. Finally, it was also observed that most attacks were short, but some of them were heavily-tailed (lasting hundreds of minutes).

Dr. Labovitz concluded his talk with the discussion of the infrastructure protection challenges. The first one is how to detect such attacks. Sometimes, it is difficult to distinguish the normal activities from the attacks. The second one is how to collect and analyze the huge amount of information. The third one is how to do all of the above in real-time.

2.5 DDoS Overview, John Ioannidis, AT&T Labs - Research

In his talk, Dr. Ioannidis gave an overview of the distributed denial of service attack. The target of the DDoS attacks can either be the end node or the network link. In the case of the end node, the attack is trying to consume the node resources as much as possible, such as the CPU cycles by causing unnecessary processing (application level attack) or the memory by memory exhaustion. For the network link, the purpose of the attack is to make the targeted link severely congested. The difficulty of dealing with DDoS is that the victims usually can do nothing to protect themselves without other's help.

After describing some characteristics of current DDoS attacks, Dr. Ioannidis discussed various defense approaches to the attack. The first part of the defense is to detect the attack, which can be done by traffic monitoring, link-interfacing monitoring or traffic marking. The second part of the defense is how to respond to the attack after it happens. There are several approaches: The first one is to do the traffic management via traffic rate limiting, filtering and redirection. The second one is to distribute the re-

sponse either to the specific points (e.g., border routers) or along the attack path (e.g., pushback). Then Dr. Ioannidis discussed more detail about the blackholing and pushback approaches. In the blackholing approach, once detecting that a service is under a DDoS attack, one can shutdown some outside path through which the malicious traffic comes and keep at least the targeted service available to the local clients. In the pushback approach, once detecting the flash crowds or the flooding-style DDoS attacks, a router can ask upstream routers to control that attack by controlling the aggregate upstream. Finally, Dr. Ioannidis concluded his talk by pointing out some of the limitations of current DDoS defense approaches and the questions that needed to be answered in the future to defend against the DDoS attack.

2.6 Countering DDoS without Changing the Internet, Angelos Keromytis, Columbia University

Dr. Keromytis presented a mechanism called WebSoS for countering DDoS attacks against Web servers. The WebSoS doesn't require much support from the ISP which provides the networking service to an attacked site. It combines the overlay network, content-based routing, packet filtering and client authentication. The WebSoS mechanism separates the good traffic from the bad or unknown traffic by allowing only the authenticated clients' traffic to enter the overlay network and then routing that authenticated good traffic over the overlay nodes, which act as the Web proxies. The overlay nodes will finally route the good traffic to the protected Web servers. The routers immediately surrounding the protected Web servers will aggressively filter and block any packets from all hosts except those allowed overlay nodes, which are kept secret from the attackers. Those secrets can be changed dynamically if they are believed to be disclosed. Dr. Keromytis concluded his talk by pointing out a remaining issue about the presented mechanism–the requirement that the clients need to be known to the Web servers, and its possible solution. More detail about the talk can be found in "http://www.cs.columbia.edu/~angelos/Papers/websos.pdf".

2.7 Source Address Filtering, Name and Affiliation of the Speaker Withheld by Request of His Company

The speaker discussed source address filtering as a technique to deal with the DDoS attack. The attacker of the DDoS attack usually uses the spoofed source address to prevent the tracing of the attack, so if all the edge routers can implement source address filtering, then it will prevent some of the

DDoS and make the attack tracing much easier. But there are still some problems needed to be solved to make this approach feasible. One of such problems is the multi-home problem and the other more practical problem is how to motivate the different ISPs to add this function into their routers. The latter problem triggered a lot of discussion among the workshop participants, including whether governmental regulation or the commercial service contract can be part of the motivations.

2.8 Telescopes, David Moore, UCSD

Dr. Moore gave an overview of the telescope technique and how to use it to detect and monitor the DDoS and Worm attacks. The telescope is a chunk of globally routed IP address space, which has little or no legitimate traffic. The unexpected traffic through that chunk of IP address space usually means some unexpected events, such as attacks, are happening. So the telescope can provide an effective way to observe and detect various Internet attacks. The larger the size of the telescope, the more accurate its detection result will be.

Dr. Moore then described various types of telescopes. The telescope can be distributed, which uses non-continuous blocks of address space to increase its size. The advantage of such a telescope is that it can reduce the dependence on the reachability of a single address block and the traffic load can be spread over multiple sites. But there are also some constraints in the distributed telescopes: the data analysis may be trickier than the centralized peer because different address pieces have different reachability at different times and one needs to solve the problem of time synchronization and data distribution. The telescope can also be a transit one to be placed in the middle of the network instead of on the edge. Another technique of telescope is to respond to the attack actively instead of just monitoring passively, as used in the Honeyfarms.

2.9 Introspection to Worms, George Varghese, UCSD

In his talk, Dr. Varghese discussed how to use the introspective technique to detect the Internet worms. He at first presented various basic patterns and algorithms that can be used to detect the worms, which include the Heavy-hitters and Many Flows. The Heavy-hitters pattern is to monitor the percentage of intended packet within some amount of time window, and the Many Flows pattern is to detect whether the number of flows exceeds a threshold. Both patterns can be implemented via the multi-resolution

bitmap counting algorithm.

In the second part of the talk, Dr. Varghese discussed how to combine those basic patterns to detect worms, especially how to automatically extract the signature of a specific worm by automatically detecting an abstract worm. The abstract worm, as seen by the router, usually has the following characteristics: First is the content repetition. The packet payload of a worm can be seen frequently and can be detected by the Heavy-hitters pattern. The second one is the increasing infection level which can be detected by the Many Flows pattern. Finally, worms also have random IP address probing behavior and code fragment payload in their packets. Dr. Varghese concluded his talk by discussing about how to deal with the polymorphism of the worms, including the syntactic and semantic polymorphism, which require further research efforts.

2.10 P2P Systems for Worm Detection, Joel Sandin, Stanford University

Dr. Sandin focused on how to use the P2P systems to detect the worms in his talk. He discussed how to make such a P2P-based worm detection system to be intrusion tolerant, or in other words, the P2P worm detection system should allow the malicious participants in the P2P network. There are two kinds of attack to the detectors: one is the false positive, say due to the compromised detectors; and the other is the attack to P2P infrastructure itself. In order to make a P2P-based worm detection system defend against such kinds of attack, one needs to solve the problem of maintaining consistence among those P2P detectors and to build a fault-tolerant infrastructure from the P2P network.

2.11 Honeynets, Dave Dittrich, University of Washington

Mr. Dittrich gave an overview of the Pacific North West Honeynet project in his first part of the talk. The main research areas of the project include prototyping a distributed Honeynet using GenII Honeywall technologies, building the databases for the clean/compromised system images, isolating mal-ware artifacts for reverse engineering and studying cross-sector activity and trends. This project involves several universities and provides both network diversity and Honeypot diversity. Mr. Dittrich also discussed the structure and various components of the Honeynet, such as the Honeypot, Honeywall, and the data control. The second part of Mr. Dittrich's talk was about a cyber attack simulation experiment. The main observation in that

experiment is that under a cyber attack, the players were not coordinated enough with each other and most of them responded independantly with a narrow focus on just their own networks. Similar to most sites under DDoS attack, very little emphasis was placed on capture and analysis of attack traffic, identification of malware network "fingerprint", or tracing attacks back. Mr. Dittrich concluded with an interesting question: what should be an ISP's responsibility to help its customers or other ISPs in the event of attacks?

2.12 Worms Overview, Stuart Staniford, Silicon Defense

Dr. Staniford gave an overview of current Internet worms in his presentation, including various strategies of the worms to propagate (scan) and the possible defense measures–the worm containment. Dr. Staniford began with the observation that the existence of the worm can be explained by the vulnerabilities and the limitations of software development and testing. Then he described a theory to model the random scanning worms and how the worms can propagate within an enterprise environment.

After that, Dr. Staniford introduced some worm scanning strategies other than the pure random scanning one. The first is called "subnet scanning", which can differentially choose a destination address near the source address. One example is the Code Red II worm, which chooses a random scanning address from class B with probability of 3/8, class A with probability of 1/2 and Internet with probability of 1/8. By using this kind of differential subnet scanning strategy, the worm can exploit pieces of network it finds and propagate much faster than the pure random scanning one. The more powerful and theoretical worm propagation strategy is called Flash worm. The Flash worm is different from other scanning strategies in that the worm is using the propagation map which has been built before it is launched. In the Flash worm case, the attacker will at first scan all the vulnerable nodes and then build a map of worm spread, which can be optimized for routing. When the worm is launched, it will carry the built address map with it and propagate as planned. This kind of worm propagation is only limited by the available bandwidth and can saturate the Internet in tens of seconds and the internal network in hundreds of milliseconds. Another kind of similar worm is called Topological worm, which uses the network topological information it gets from the host instead of the pre-computed map. Both Flash Worms and Topological Worms are not reliably containable currently due to their propagation strategies.

Finally, Dr. Staniford presented the possible approaches to worm con-

tainment. For the scanning worm, anything that will block scans will do in principle and if one can ensure that an average scan will see less than one vulnerable machine, then the worm can be contained. Because the worm usually spreads faster than a human being can respond, one can not detect it by the signature, instead the detection will depend on correlating multiple worm-like anomalies. Dr. Staniford also pointed out that the containment measures need complete deployment so as to not only slow the worm propagation but also contain it completely.

2.13 Diverse Axes of Scaling, Dan Ellis, MITRE

Mr. Ellis discussed a possible future attack—the targeted attack in his talk. Most current attacks in the Internet are not so discriminating, in the sense that the attack is random, such as the randomized denial-of-service attack, and basically tries to compromise any possible targets. The targeted attack aims to compromise a specific target instead of all possible targets. To satisfy the attacker's objective, the attack will do no more and no less than what is necessary. This kind of targeted attack, e.g., the targeted worms, will be much less likely to be detected using current technology which mostly deals with the large-scale attack.

2.14 Modeling and Detecting the Spread of Active Worms, Lixin Gao, University of Massachusetts

Dr. Gao discussed various issues about how to model and detect the spread of active worms in her talk. Dr. Gao first discussed what needs to be monitored in order to model and detect the worms, which include both inbound and outbound traffic in the network, and various monitoring strategies, such as selectively monitoring and adaptive monitoring. After presenting various issues that needed to be addressed when modeling and detecting the worm spread, such as spoofed IP and multi-vector worm, Dr. Gao introduced a mathematical model, called "Analytical Active Worm Propagation (AAWP) Model", which characterizes the propagation of worms that employ random scanning. The AAWP model can also be extended to model and monitor the local subnet scan and other more malicious scans, such as selective scan, routable scan, divide-conquer scan and the hybrid scan which combines the previous simple scan methods.

2.15 Wormholes and a Honeyfarm, Nick Weaver, UCB

Mr. Weaver presented an approach to automatically detect novel worms via Wormholes and Honeyfarm. He at first pointed out that deploying the Honeypots across the Internet to detect novel worms requires a lot of cost and the trust in all Honeypots, which is problematic. Mr. Weaver then presented their approach of splitting the network endpoints from the Honeypots. The main idea is to use the Wormholes as the network endpoints to be plugged into the network, and let the Honeyfarm host the virtual machine Honeypots which can be created dynamically on demand. The Wormholes are the traffic tunnels and connect with the Honeyfarm by forwarding all traffic to and from it. The Honeyfarm receives the traffic, which may be the worms, from the Wormholes and can create the Honeypots as Virtual Machine images and forward that traffic to the Honeypots. By monitoring and analyzing the traffic and its effect on the Honeypots, one can automatically find whether a new worm is in the Internet, what are the configurations vulnerable to the worm, what is the malicious behavior of the worm and its possible attack signatures. Finally, Mr. Weaver also analyzed the trust relationship among the Wormhole deployers, Honeyfarm operators and the responding systems receiving the alerts.

2.16 Router Attacks, Name and Affiliation of the Speaker Withheld by Request of His Company

The speaker discussed various possible attacks to the router and its routing functions (e.g., BGP). One such attack is the flow cache CPU attack, which causes the unnecessary computation and updating of the BGP routing table, the IP routing table and the forwarding table. Other attacks include configuration attacks, routing and session hijacking, memory exhaustion, session dropping, and possible DoS attacks. Other problems can also be caused by the incompatibility of BGP implementation among the routers, router software bugs and the poor human interface for the router operator commands.

2.17 Link-cutting Attack, Steve Bellovin, AT&T Labs - Research

Dr. Bellovin discussed a kind of network infrastructure attack called "link-cutting attack". The classic routing attack, which propagates false routing information among the routers, can be defended against by deploying secure routing protocols. But if the attacker can control some links or nodes and

has a map of the routing topology, then it is computationally feasible for the attacker to calculate what links to cut to force the traffic to pass the controlled points. Dr. Bellovin also presented some experimental results, showing that in hundreds of trials on intra and inter-ISP topologies, one can have an attack success rate of 80-90 percentage and each link calculation takes at most a few seconds, even on very large topologies.

2.18 Auto-patching, Angelos Keromytis, Columbia University

In this talk, Dr. Keromytis discussed a reaction mechanism that seeks to automatically patch vulnerable software. The motivation for this work is the observation that there always exist many software flaws, which can be exploited by various attacks, and people always forget to update the patches. Furthermore. the worms can spread at rates that are much faster than any human's manual reaction. Both call for an automatic reaction mechanism to the attacks. The mechanism described by Dr. Keromytis employed a collection of sensors that can detect and capture potential worm infection vectors and analyze those suspicious traffic in an isolated environment. After that, it can automatically identify the exploited software weakness and generate the corresponding patches, which will be updated to the vulnerable software. More detail about the talk can be found in "http://www.cs.columbia.edu/~angelos/Papers/endpointpatching.pdf".

3 Future Research Challenges

Several future research challenges were discussed in the workshop. Here are some of them:

- How to accurately detect the large-scale attacks in the Internet. Sometimes, it is difficult to distinguish the normal activities from the attacks.
- How to collect and analyze the huge amount of attack monitoring information and do it in real time.
- How to divide the responsibility and obligation between the network service providers and their customers with regard to the response to the attacks, and which kind of business service model is needed here.

- How to build a formal collaboration between different sites, ISPs and various agencies, which can automate the coordination among related parties in the case of being attacked.
- How to motivate the ISPs to deploy the defensive measures, e.g., source address filtering to defend against the DDoS attacks.
- How to protect the network infrastructure itself from the possible attacks.
- How to deal with the polymorphism of the worms, including the syntactic and semantic polymorphism.
- How to build a secure worm detection system in the distributed systems.
- Given the possible high spreading speed of the future worms, how to deploy the worm containment mechanism to contain them.

4 Acknowledgements

The author of this report wants to thank Dr. Brenda Latka, Associate Director of DIMACS, and Dr. Fred S. Roberts, Director of DIMACS, for their valuable comments. And also thank the organizers and the speakers of this workshop. The author and the DIMACS Center acknowledge the support of the National Science Foundation under grant number CCR 03-14161 to Rutgers University.