OASIS Project Validation Report

for effort titled

Randomized Failover Intrusion Tolerant System (RFITS)

Ranga S. Ramanujan
Architecture Technology Corporation

December 1, 2001

1. Technology Description and Survivability Problem Addressed

The Global Information Grid (GIG), envisioned by the DoD, represents a widely distributed networked information system consisting of a wide variety of sub-systems interconnected by a shared IP-based network infrastructure (or the GIG backbone). Notionally, the GIG is a federated system whose constituent subsystems may be owned and operated by autonomous entities. That is, there is no notion of a central authority that can exercise control over the entire system. While the ubiquitous networking capabilities supported by the GIG enables the implementation of the next generation of network centric applications for the military of tomorrow, it also exposes mission critical applications implemented on the GIG to network borne denial of service (DoS) attacks. Such DoS attacks can be launched from anywhere on the distributed network with crippling effect on high-availability, mission critical applications.

A notable form of a network borne DoS is the Distributed DoS (DDoS) attack, where attack traffic may be generated simultaneously from multiple points on the network from machines that have been "hijacked" or "captured" by the attacker. This attack traffic, when directed at a victim host or network, can inundate the victim and deplete its computational and communication resources and preclude it from providing services to legitimate users. It is imperative that mission-critical applications operating over the GIG be protected against such DDoS attacks.

The goal of the ongoing research reported here is to develop organic survivability techniques that can be used to build highly available, real-time, mission critical GIG applications that are resistant to DDoS attacks. We define organic survivability techniques as those mechanisms that can be implemented using only the equipment (i.e., computing and communications resources) owned and operated by the customers of the GIG network service. Consider, for instance, the scenario where a networked application supported by a GIG customer relies on the services of a network service provider on the GIG for wide area information transport services necessary to interconnect the geographically separated edge networks owned by the customer. In this case, although

the customer owns and controls the different edge networks, the customer has no direct control over the computing and communications resources of the network service provider upon which it relies. In this scenario, the organic survivability techniques developed by us would enable the customer to build DoS-resistant survivable GIG applications using mechanisms implemented within the edge networks owned by it without requiring any additional mechanisms or changes within the infrastructure owned by the network service provider.

1.1 Target Network Environment

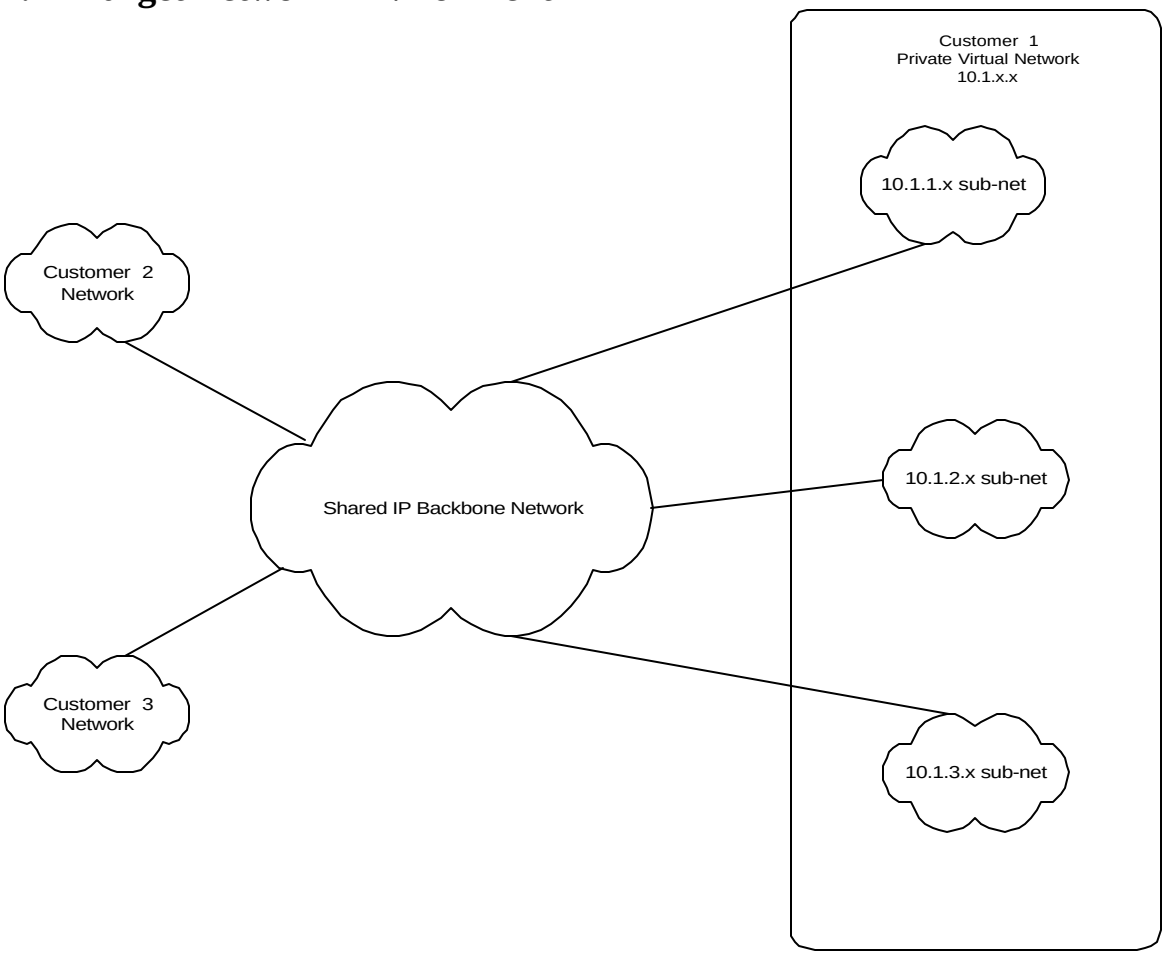


Figure 1: Example of a Notional Architecture of Target Network Environment

Figure 1 depicts a target network environment representative of that within the GIG that will be used to illustrate the types of DoS attacks that our survivability techniques are designed to handle. Shown in the figure is a shared, global IP based infrastructure owned and operated by one or more network service providers (NSPs). These NSPs may either be DoD organizations or commercial enterprises providing network access services purchased by the military. Such use of commercially available services is expected within the DoD's GIG.

As shown in the figure, the shared global IP infrastructure provides the information transport services (where the transported information includes voice, video, and data)

needed to interconnect customer-owned edge networks together. These edge networks may be owned by different customers, each providing application services to its subscribers on the GIG. Furthermore, a single customer's network may consist of multiple distributed elements or subnets that may be interconnected by VPNs implemented over the shared IP infrastructure into a single virtual customer network.

Referring to Figure 1, customer 1 owns and operates a virtual private IP network (10.1.*x*.*x*) that is composed of three geographically distributed sub-nets (10.1.1.*x*, 10.1.2.*x*, and 10.1.3.*x*) tied together using VPNs implemented over the shared global IP network. The distributed application implemented over this virtual private IP network is made accessible to the other subscribers on the GIG (i.e., customers 2 and 3) through a globally reachable IP address (say 128.42.72.3). For example, the 10.1.2.*x* and 10.1.3.*x* subnets of the virtual IP network of customer 1 might represent forward deployed networks that collect and pre-process data from field sensors and forward this information stream to the 10.1.1.*x* sub-net over secured "virtual leased lines" implemented by VPNs. In the 10.1.1.*x* network, the sensor streams are fused to develop a common tactical picture of the battlefield. The common tactical picture developed by this application is then made available to other subscribers of this information through a globally reachable IP address (*e.g.* using a publish-subscribe mechanism). In the figure, customers 2 and 3 are the subscribers of the application service provided by customer 1.

1.2 DoS Attacks in Target Networks

The survivability techniques being developed by the RFITS effort are primarily targeted towards thwarting two kinds of network borne DDoS attacks in the environment described above: 1) flooding attacks; and 2) host take-down attacks. Either of these attacks may be initiated from one or more points of attachment on the shared IP infrastructure of the GIG and may be directed at one or more victim GIG customers.

A flooding attack may take one of two forms: *access link flooding* and *service request flooding*. In the former case, the attack traffic consists of spurious IP packets directed at one or more customer networks, potentially with spoofed IP addresses, originating from one or more points in the network. This bogus traffic may inundate the access link connecting the customer network to the shared backbone network and thereby usurp valuable bandwidth from legitimate traffic using that link. The overloading of this network access link by the attack traffic may result in partial or total denial of service to subscribers of the customer network. Further compounding the situation, spoofed attack traffic may "seep" through perimeter defense mechanisms, such as firewalls, installed within the customer network to reach hosts (servers) implementing the application service. Substantial processing resources at these hosts may be consumed in handling this flood of spurious packets thereby diverting these processing resources from legitimate subscribers of the application service.

In service request flooding, the attacker sends spurious application level service requests to the customer network thereby overloading the processing capacity of the hosts implementing the service and causing an unacceptable degradation in the level of service provided to legitimate users. Unlike packet flooding, service request flooding may not necessarily result in the overloading of the link connecting the customer network to the backbone IP network.

Host take-down attacks exploit known vulnerabilities or bugs in the implementation of the host software to direct messages (or packets) to victims that cause the host or application program to crash. An example of a host take-down attack that plagued the Internet not long ago is the "ping of death" attack which exploited deficiencies in certain operating system implementations by sending them malformed packets that they could not handle and that would cause them to crash.

1.3 The RFITS Approach

Our approach for protecting GIG applications against network borne DDoS attacks is based on the concept of *randomized failover*. Survivable systems built using this approach are termed Randomized Failover Intrusion Tolerant Systems (RFITS). The RFITS approach for survivability is predicated upon making the failover mechanisms that are invoked by a fault-tolerant system appear to be an unpredicatable, random process from the perspective of the attacker. In order to launch a successful DoS attack on GIG customer in the network environment described above, the attacker needs to know two items of information about the system: 1) known vulnerabilities or limitations of the system that can be exploited by the attack; and 2) the current configuration of the victim, *i.e.*, the customer. In the GIG network environment, one of the known vulnerabilities that can be exploited to launch DoS attacks is that spoofed traffic flows can easily be transported over the shared IP network to the victim. This is because the shared network has no built in mechanisms to check the authenticity of packets flowing through it and to filter out spurious traffic flows. In addition to this knowledge, the attacker of a customer network needs such system configuration information as the globally reachable IP address of the customer network to which spoofed packets can be sent. In the RFITS approach, the failover mechanism that is invoked by the victim system upon detection of a DoS attack reconfigures the system. The intent is to render the attack that was directed upon the old configuration of the system ineffective through this reconfiguration.

The RFITS-based survivability techniques being developed by this effort are primarily targeted towards building two kinds of "middleware" services for survivable GIG applications:

1. Survivable Information Transport Services
2. Survivable Server Group

Survivable Information Transport Services (SITS) are implemented over IP and support a highly available, DoS-resistant one-to-one, one-to-many, many-to-one, and many-to-many packet delivery service for GIG applications implemented over the shared IP infrastructure. Survivable Server Group (SSG), as the name implies, is a middleware service implemented using RFITS-based survivability techniques that protects highly-available replicated server groups against DoS attacks. Several alternative techniques for implementing SITS and SSG have been defined by the RFITS effort thus far. These are described in the RFITS Application Handbook [1] along with guidance on the design tradeoffs that must be made in selecting the one that meets the requirements of a given survivable GIG application. To date, the Architecture Technology Corporation (ATC) team has successfully built and demonstrated a prototype of an operational survivable

virtual private network (VPN) service using RFITS techniques described in the handbook.

2. Assumptions

In order for DoS attacks to be successful in a dynamically reconfigurable RFITS-based system, the attacker needs to be able to accurately track system configurations as they change and adapt the attack accordingly. The RFITS approach uses the computing analog of defensive battlefield tactics such as "camouflage", "concealment", and "deception" to implement a failover process or system reconfiguration process that makes it difficult for an external attacker to determine the new configuration of the system. Of course, given sufficient time an attacker might potentially be able to deduce this information. However, the goal of our RFITS based organic survivability mechanisms is to deter the attacker long enough to allow other complementary attack neutralization techniques to pin-point the source of the attack so that actions can be taken to neutralize the attacker before further damage is inflicted on the system.

Neutralization of an attack may take several forms. These include (1) eliminating vulnerabilities or design flaws in the system or system components that was exploited by the attacker(s) to launch the attack on the victim(s); and (2) use of offensive techniques to disable or destroy the source of the attack.

RFITS survivability techniques are targeted towards countering attacks launched from edge networks attached to the GIG core. Insider attacks from within the GIG backbone infrastructure are not addressed.

2.1 Residual Risks, Limitations, and Caveats

RFITS survivability techniques for ensuring system availability in the face of DoS attacks rely on the existence of complementary techniques for identifying the source of the attack and the neutralization of the attack source. These survivability mechanisms are intended to sustain continued operation of the system in the face of DoS attacks for a period of time that is long enough for complementary attack neutralization techniques to identify and remove the attack source (or the system vulnerability exploited by the attack). The effectiveness of the survivability techniques is therefore dependent on the ability of the attack neutralization techniques to take effect within the window of opportunity afforded by the RFITS techniques.

3. Vulnerabilities and Attacks

The survivability techniques being developed by the RFITS effort are primarily targeted towards thwarting two kinds of network borne DoS attacks, i.e., 1) flooding attacks; and 2) host take-down attacks. Such attacks may be facilitated by inherent vulnerabilities within the system. We are not concerned about how these attacks were originated. The focus of these survivability techniques is to enable the system to cope with the attacks when they occur and provide continued service in spite of the attacks.

Attacks addressed by RFITS survivability techniques can be further categorized into four types:

TAV-1	Access link flooding
TAV-2	Dial port flooding
TAV-3	Service request flooding
TAV-4	Host takedown

4. Information Assurance and Survivability Attributes

The goal of RFITS survivability techniques is to ensure high system availability in spite of network borne denial of service attacks launched on the system (i.e., the consumer network) by attackers residing on the edges of the GIG.

5. Comparison with Other Systems

Existing work on approaches for dealing with DoS flooding attacks primarily focus on one of the following two areas: *fail-soft* mechanisms and *traceback* mechanisms. Fail-soft mechanisms are aimed at mitigating the deleterious effects of flooding attacks on the victim hosts. They are designed to enable a victim host to provide at least a minimal level of acceptable service in the face of flooding attacks targeted at it. As noted by Savage *et al* [2], this approach serves at best as a stopgap measure. It does not eliminate the problem to allow the victim host to resume providing full service nor does it serve as a deterrent to attackers. More importantly, these fail-soft mechanisms focus on the end hosts only and do not address the problem of attacks that are designed to disrupt the operation of entire edge networks by inundating the link connecting the edge network to the shared IP backbone infrastructure.

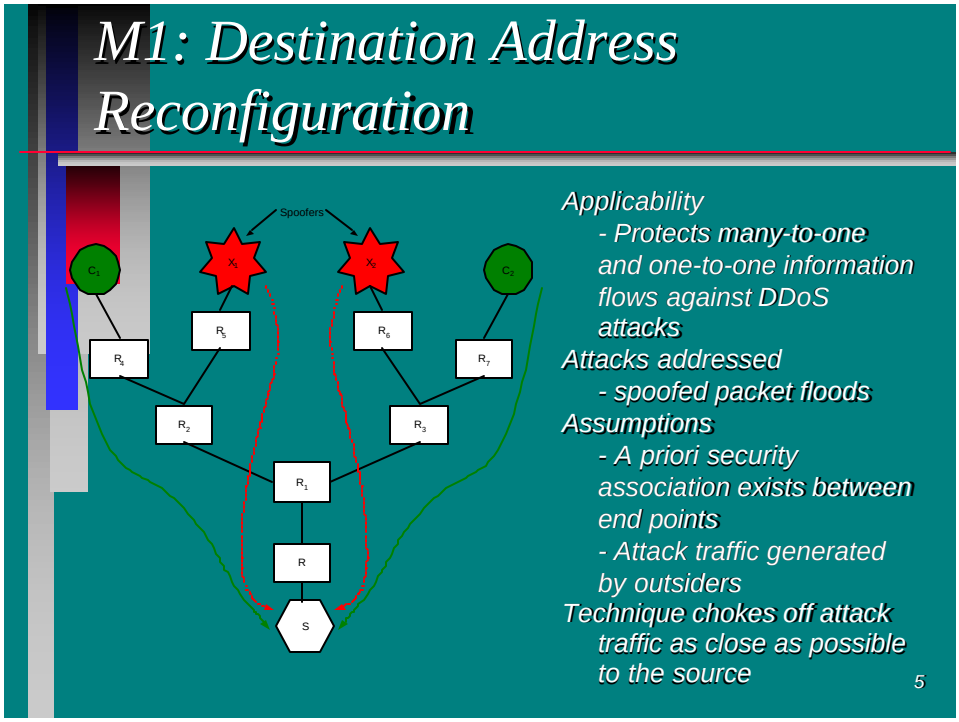
Traceback mechanisms [2] attempt to trace the attack traffic towards their origin. This is motivated by the fact that the effectiveness of techniques such as packet filtering would be greatly enhanced by applying them as close to the generation points of the traffic flood. A limitation that is common to all existing traceback mechanisms is that they provide no means for a network to easily differentiate spurious packets from legitimate packets originated by a spoofed source so that selective filtering of traffic can be performed. Consequently, all packets from a source IP address destined for the victim are filtered out, including legitimate traffic. Furthermore, for high-availability mission-critical network applications that cannot tolerate disruptions of network service beyond a few seconds, the latency associated with network recovery using traceback mechanisms may be unacceptable. For large-scale distributed DoS flooding attacks, the amount of time taken by traceback mechanisms to trace all the distributed attack sources and squelch the spurious traffic may be in the order of minutes to hours, even assuming that the entire traceback and network response process is automated. Thus, traceback mechanisms by themselves are insufficient for meeting the needs of such network applications and must be augmented by other mechanisms that provide immediate restoration of network services upon detection of a DDoS attack.

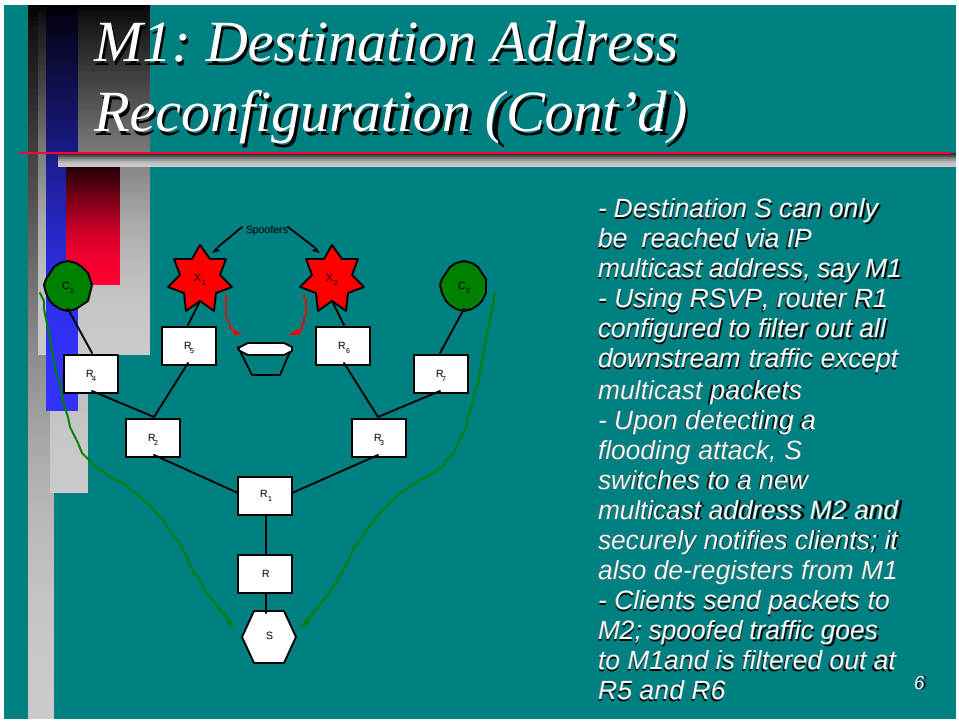
Existing approaches for dealing with host take-down DoS attacks are primarily based on perimeter defense mechanisms (*e.g.* firewalls) that examine the packet stream passing through them to cull out or block "suspect" packets. Packets may be marked as "suspects" if their structure is similar those packets that have been known to crash end hosts that process them. However, this approach is only effective dealing with known vulnerabilities of a host system and all the known ways in which they could be exploited by an attacker to crash the host.

6. Survivability Mechanisms

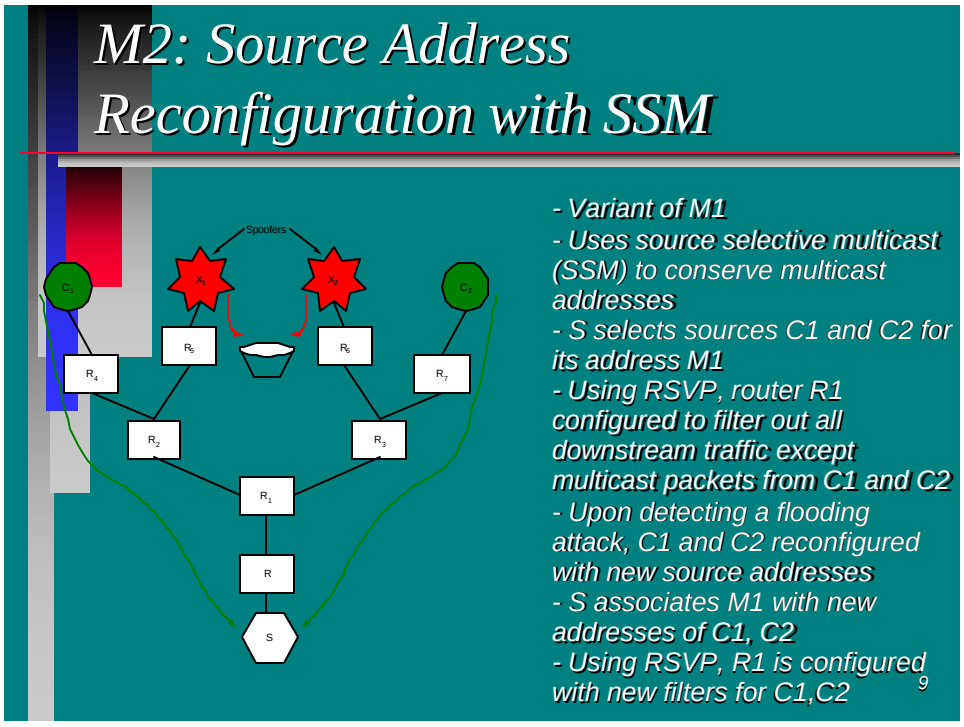
The RFITS Application Handbook [1] documents over 20 different survivability design patterns that may be employed, either individually or in combination, to build survivable information systems that are resistant to network-borne DoS attacks. Below, we summarize six of those techniques. The first five techniques presented below (i.e., M1-M5) address DoS flooding attacks while the sixth technique, M6, addresses host takedown attacks.

M1: Destination Address Reconfiguration

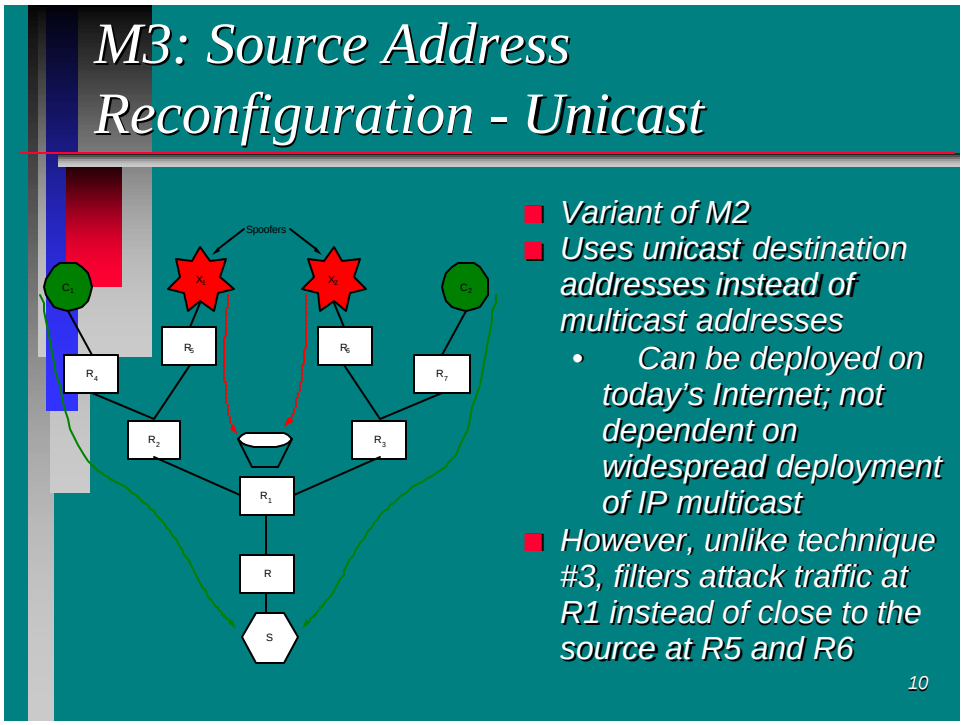




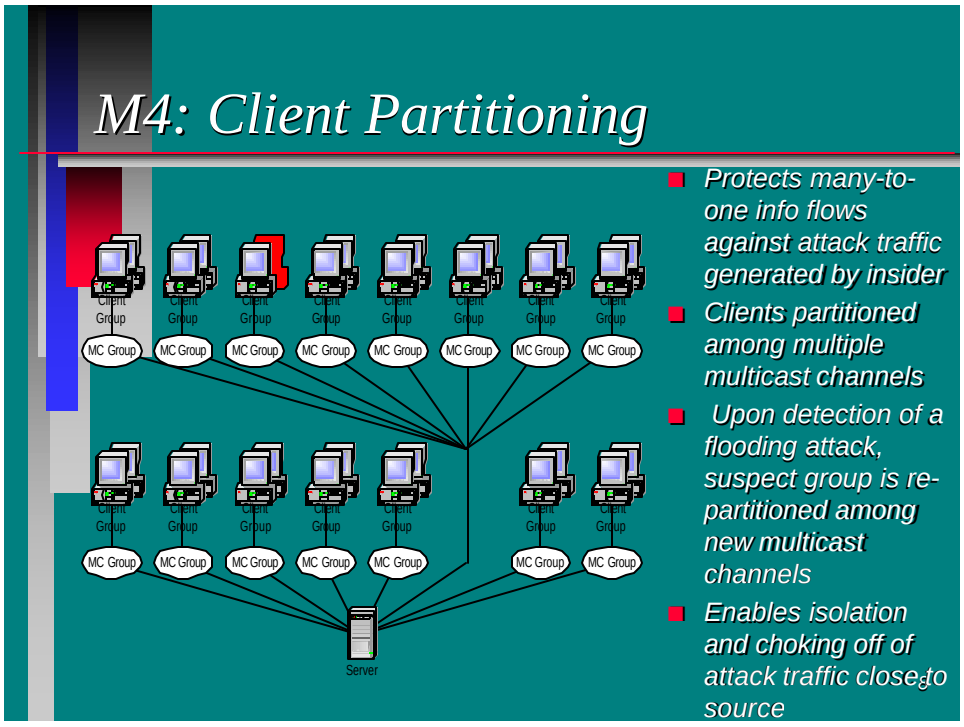
M2: Source Address Reconfiguration with Source-Selective Multicast (SSM)



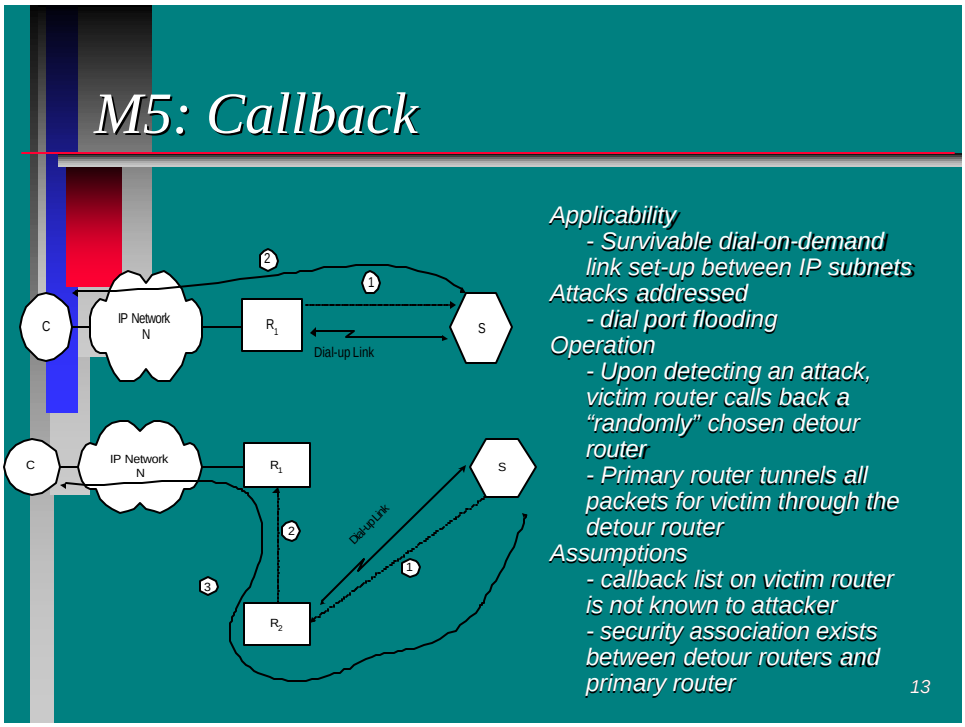
M3: Source Address Reconfiguration - Unicast



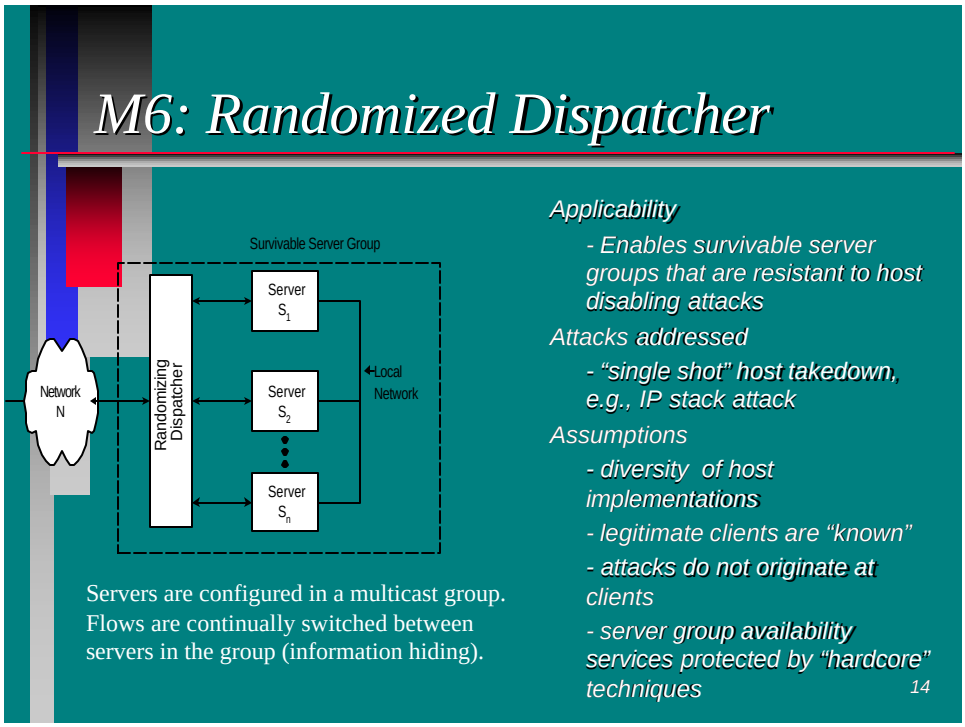
M4: Client Partitioning



M5: Callback



M6: Randomized Dispatcher



7. Rationale

Stage	TAV	AV (Availability)
Operation	TAV-1	M1 or M2 or M3 or M4
Operation	TAV-2	M5
Operation	TAV-3	(M1 or M2 or M3, or M4) and (M6)
Operation	TAV-4	M6

For verifying and validating the effectiveness of RFITS survivability techniques, we plan to use a combination of red team testing and analysis. Some of the techniques described in the RFITS Application Handbook [1] are being implemented as software prototypes. These will be amenable to red team testing. The remaining techniques in the handbook will be subjected to qualitative analysis.

8. Cost and Benefits

DDoS attacks represent a serious threat to enterprises operating over the Internet. A recent study by researchers at the University of California–San Diego found that nearly 4,000 DoS attacks are launched each week with 20 to 40 attacks at any instant in time. Some systems that were studied were attacked as often as once per minute with attacks ranging from 1000 packets per second to 600,000 packets per second. The barrier for entry for launching DDoS attacks on the Internet is relatively low given the ease with which well known attack tools can be obtained and operated [2]. For the GIG, these barriers could be raised higher by restricting access to the backbone network to trusted edge networks and by requiring the DoD owners of the edge networks to implement security measures designed to make host capture or hijacking very difficult. Despite these measures, GIG applications would still remain vulnerable to network borne DDoS attacks. This is because, some of the edge networks on the GIG may be forward deployed, deeply networked system (*e.g.* unmanned sensor networks). The physical capture of such nodes by an adversary cannot be ruled out. Such captured elements of the GIG could be used by a sophisticated adversary to launch debilitating DoS attacks on mission-critical GIG applications. RFITS survivability mechanisms enable edge networks on the GIG to provide continued services in spite of DoS attacks targeted at them. These survivability mechanisms thus enable the implementation and deployment of high-availability, mission critical GIG applications.

The costs associated with RFITS varies with the individual mechanisms that are employed. A common component of the cost across all these mechanisms is the use of some form hardware and software redundancy for implementing them. Some mechanisms, such as that for survivable information transfer services (SITS), also employ logical redundancy in the form of a pool of IP addresses maintained by the system to support dynamic address reconfiguration. The RFITS Application Handbook [1] describes the implementation and performance overhead costs associated with each of the survivability techniques in more detail.

References

1. “Architecture and Application Handbook for Building Survivable (Intrusion-Tolerant) Systems”, Interim Project Report submitted to DARPA by ATC, June 2001.
2. D. Moore, G. Voelker, and S. Savage, “Inferring Internet Denial-of-Service Activity,” Proceedings of the 2001 Usenix Security Symposium, Washington, D.C., Aug. 2001.