

The Crux of NT Security - Phase 4

Aaron Sullivan 2000-12-18

The Crux of NT Security Phase Four: Network D - High Availability, High Speed, High Security, (High Cost)

by Aaron Sullivan

[SBC DataComm Security](#)

last updated Dec. 18, 2000

This is the fourth in a series on NT security. In the last [article](#) we discussed three common network designs referred to as Networks A, B and C. Hence, this article will discuss a last design, Network D, for those with more performance and security demands, as well as a high availability feature, and the additional budget required to implement it. Network D has all the bells and whistles, costs and complexities that the other networks were lacking for a truly high-security, high-availability network. To those working in organizations with very large budgets, staff, and a business residing nearly 100% on the Internet - this one's for you.

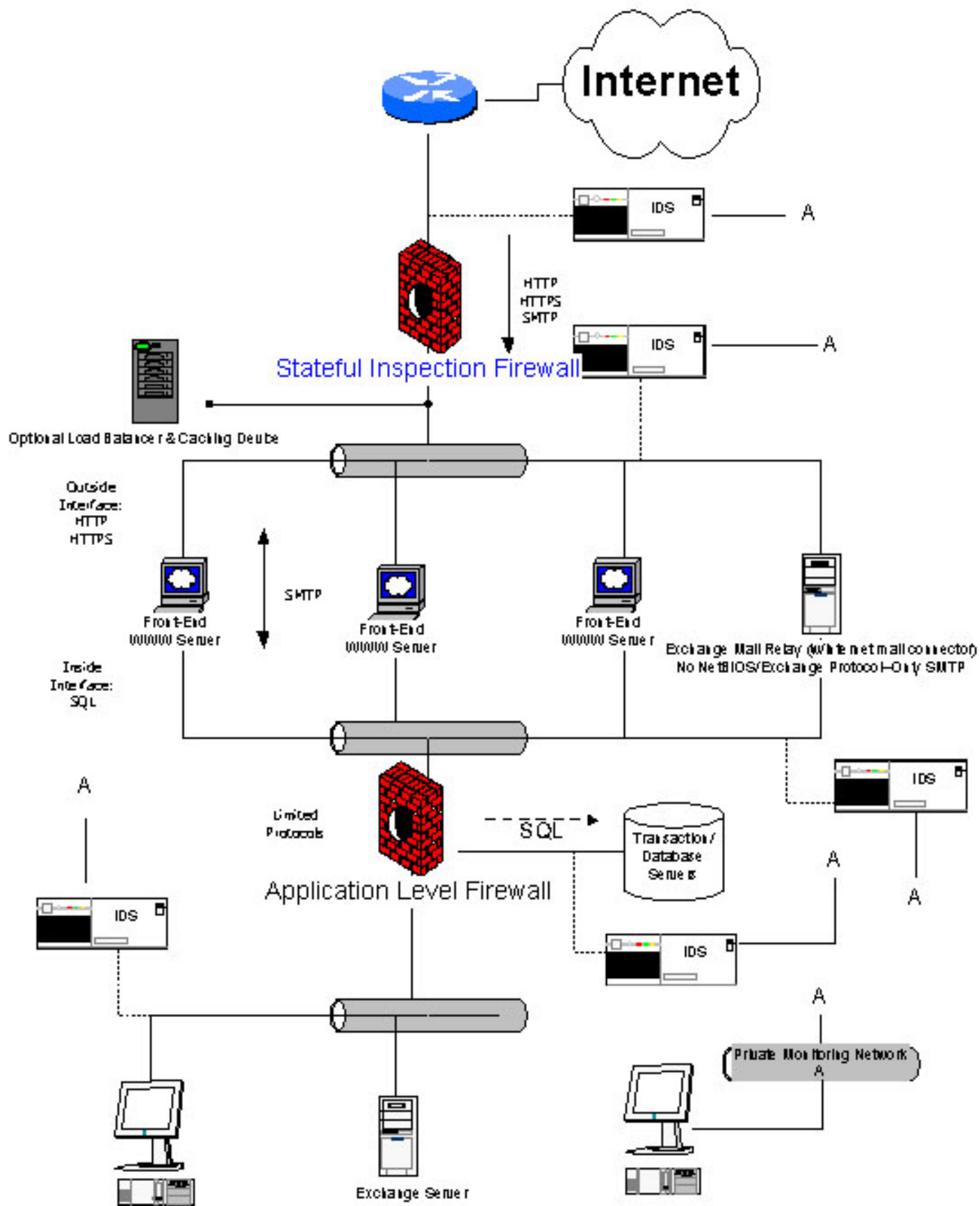
As with many potentially hazardous items available to the public at large, this design comes with a warning:

Implementation of Network D by those unprepared to deal with its complexities in terms of expertise requirements, staffing requirements, general maintenance, and time and monetary investment will likely result in the creation of a large, expensive bombshell that will likewise result in harsh reprimand and/or termination of employment.

Implementing this network to specifications has costs in every way mentioned above. Those costs are likely to be higher than even the most generous initial projections that one might make after pricing out the hardware involved. To the sys-admins out there who have just received a sizeable budget for implementing security at a small-to-midsize firm: I realize that it looks very cool, but think carefully before climbing aboard Network D's ship. As a general security implementation guideline, if the risk involved outweighs the cost, begin to consider implementation. And remember, implementing high availability in the firewalls will result in twice the number of devices and network lines, thus costing twice as much.

Below is an illustration of Network D. Before we begin the discussion, I'd like to note that those unfamiliar with secure network design principles would do well to go back and read the first

three articles before reading this one. Certain things about this design are based on and should be understood by implied logic gained from the previous articles in the series.





Let's talk about what's different in Network D and why. Probably the most distinct feature is the presence of two firewalls instead of one. The first firewall is labeled with a different color than the second. This is meant to imply that they are separate types of firewall. There are a few reasons for this. First, having two separate types of firewalls lowers the risk of an attacker getting into your internal network or your transaction servers. Why? The number of vulnerabilities found in firewalls tends to be lower than most other networked objects. The likelihood of an attacker bypassing both firewalls if they are of differing make is much lower than if they are of the same.

Another second reason for implementing two different types of firewalls is performance. It is likely that the majority of traffic passed into the DMZ between the two firewalls is for static content or that content that does not need to be dynamically generated by databases queried on a transaction server. Because stateful-inspection firewalls tend to be much higher-performance than application-level firewalls, performance for static content traffic will be much higher, thus lightening total performance restraint. Once again, rules at the firewalls should permit only the required protocols to be allowed both inbound and outbound permission. If there seem to be any red flags at this point, don't worry; we'll cover them later in the critique of this network.

Another notable characteristic of Network D is that all servers in the DMZ are multi-homed. Sys-admins who are implementing the network should make sure that routing/forwarding is disabled on them and that they are only allowed to pass the specified types of traffic per interface in the diagram. The types of traffic that they should be allowed to pass should be enforced both at the host level and at the firewalls (i.e. create specific rules for these hosts on the firewall and specify explicitly the traffic they are allowed to pass). Hosts on the internal network may be allowed to pass more types of protocols out of the network, but the types they are allowed to pass should also be explicitly defined per their NATed IP addresses. You may also note that the transaction server is now layered behind two firewalls, but has the same rules for passing query traffic as in previous designs.

Intrusion Detection Systems should be implemented to monitor traffic according to what's important on their segments. The first IDS en route into the network should be able to set off an alarm indicating a precursor to an attack. All IDS's should be designed with at least two

interfaces: one for monitoring attacks and the second for the management link. On the monitoring link, ideally there would be specially constructed cable with the transmit pair cut. The monitoring link would not have an IP address and with no transmit link, it would be impossible to remotely break into the IDS (although it is still possible to disable the IDS). Also, the management link should be to an entirely different network, disconnected from the Internet. Such a design makes the IDS's incredibly difficult to defeat. There's also an option to install HIDS on each of the front-end web servers and the mail relay.

One question that some might be asking, "what about a VPN?" It seems that every vendor has its own endorsed design for a VPN relative to the firewall(s). Some vendors say that a VPN device should be implemented at the same vertical level on a network as the firewall. In other words, the VPN device should provide a second point of entry so that one can come into the network either via the firewall, or the VPN. Both as a matter of personal opinion and from experience in lab work, I believe that this is a very bad idea for two reasons. The first reason is that VPN equipment technology has not been around long enough to be tested and proven to act as a fully protective gateway to a network. The second reason is that one of our labs has discovered a major vulnerability in all models of one of the largest VPN vendors on the market! While we can't disclose the vulnerability (because there is no fix for it at the time this article was written), I can suggest that one solution to the problem is to put the VPN device on a DMZ of a firewall, thus using the more proven, tested and true security of that firewall to protect it more fully. This also follows the layering ideology that has been discussed throughout this series. To sum up, put the VPN device on a zone of the firewall separate unto itself. In this design, we recommend putting it on an extra zone of the application level firewall. As an implementation tip, don't run any address translation for that zone and, of course, limit the inbound protocols to only those necessary to negotiate and maintain a VPN session.

Now, let's talk about the problems with this network. There are two major issues that I see for this network. The first issue is that the outside firewall does not filter up to the application level. Application firewalls are a double-edged sword: they provide maximum security at minimum speed. For the networks that we discussed in the third article of this series, traffic load is likely not at a point where the slower performance of an application firewall is made apparent. However, in Network D, the design is based on an assumption that traffic load will be both constant and intense. One way of getting around having an application-based firewall at the outside might be to install a caching server to take most of the requests for static content from the servers and centralize them to a cache. Since the caching device will serve up most of the requests in the end, the conversations will generally be reduced to one device talking to the

firewall. This is done with the intent of minimizing the amount of information that the firewall has to keep in its state tables about concurrent connections. However, in some cases and with some kinds of caching devices, there will be no performance gain. The choice between security and performance is yours.

Another major issue, as stated in the warning at the beginning of this document, is the incredible level of complexity brought on by trying to make all of these networks part of one, single corporate network. Consider that each security function of this network will likely require at least one very qualified person to administer-especially the IDSs and web servers. Hiring someone who is not fully qualified to administer any security component of this network will likely result in overall lower security and performance than by running a more simple design with less hardware. Such complexity also requires fabulous policy and management staffing to administer properly. In essence, running a network like Network D properly requires people for whom the management of each of these technologies is a simple task. This network may be a dream for designers and management wishing to show-off the company's high technology; but without the proper resources, it can be a nightmare for administrators. It might be wiser to build two separate networks with one for public use and one for corporate use. In this way, the complexity is divided into two, (probably) easier-to-swallow chunks. However, if you link them together through any means, you might as well go with this design.

As a final note, some may wonder why one of the subtitles of this document mentions high-availability when there doesn't appear to be any redundant hardware in the chart. This design is built to be easily upgradeable to high-availability features. The redundant devices are not displayed in the chart because doing so would have made it too cluttered. Remember that if you are going to be using switches as part of the high availability system, port "mirroring" or "SPANing" needs to be configured for every port that the IDS's are plugged into.

It's been a pleasure doing these articles. Comments, questions, alternative information, or flames are welcome. If you are thoroughly impressed or you think I'm completely off my rocker, email me (aaron.r.sullivan@sbc.com) and I will respond in as timely a manner as possible.

Aaron Sullivan is an engineer (specializing in building, securing, and penetrating the Microsoft NOS environment) for the fine, fine SBC Datacomm Security Team. SBC DataComm offers managed firewall services, IDS monitoring services, vulnerability assessment, disaster recovery, network design services, and penetration testing. If you would like to request the services of SBC's security team, you may visit the web-site (<http://www.sbcdata.com/>) or call (800) 433 1006 and leave the request there.

The Crux of NT Security - Phase One: The Approach

Aaron Sullivan

The Crux of NT Security - Phase Two: Securing The Hosts

Aaron Sullivan

The Crux of NT Security - Phase Three: Controlling and Monitoring Communications

Aaron Sullivan

[Privacy Statement](#)

Copyright 2006, SecurityFocus