SYSTEM SECURITY ASSESSMENT USING A CYBER RANGE

H. Winter

Northrop Grumman Information Systems Europe, United Kingdom, Leander House, 4600 Parkway, Solent Business Park Fareham, PO15 7AZ (hwinter@ngms.eu.com)

Keywords: Cyber-Range, System Security, Virtualisation, Simulation, Models.

Abstract

The paper explains the concept of a cyber range and its use for performing system security assessments. It shows the advantage of evaluating security from a whole-system perspective rather than individual components and undertaking this with no risks of contamination, damage or degradation of the actual system. Moving the architecture of a real system into a cyber range in a meaningful and cost-effective way is the key challenge for performing security assessments. A solution is to use representative models and virtualisation however the paper explains that it is necessary to be clear what side effects this might have.

1 Introduction

The world's infrastructure increasingly depends on networked computer systems. Networks connect mobile and home users, companies, industry, organisations, and governments and their departments. Critical infrastructure and utilities such as the health service, police, civil defence, control and distribution of gas, water, electricity, phone, radio and TV are all connected.

A lone attacker using an inexpensive PC and free tools can cause extensive damage to vulnerable systems in any part of the world. Attacked systems may become unstable or even unavailable but consequences may also include physical damage, financial loss, pollution of the environment (e.g. through spills and leaks of harmful substances), and collateral damage to dependent systems (e.g. wide spread loss of electrical power causing loss of telecommunication systems and financial services).

1.1 Making systems safer from attack

We can make systems safer from attack from the internet by not connecting them in the first place, surrounding them with perimeter defences, or fixing the vulnerabilities. We also need to consider attacks from the inside, for example by disgruntled users.

Safety critical systems were traditionally not connected to the internet but this is changing. Internet connectivity offers significant benefits such as using the internet as a convenient and inexpensive carrier to connect geographically dispersed installations, allowing remote control and maintenance without travel costs, or providing management with instant feedback on the system output for planning and marketing. Even systems that still rely on dedicated networks for connectivity are in danger because frequently somewhere in the network a machine exists that, usually inadvertently and unplanned, bridges the public internet with the private network. As in the case of Stuxnet the fatal "connection" may just be the shared use of USB sticks between two computers on different networks.

Because the simple "don't connect" strategy does not work in practice, security improvements use the other two measures instead. We can protect networks by adding network defence devices such as firewalls, anti-virus and intrusion prevention systems and we can try to fix vulnerabilities. Both measures need to be applied with care.

Adding too many network defence devices is not only expensive and carries the risk of making the system unusable for its intended purpose, it also adds complexity, support costs and opens the door to configuration errors with the very real possibility of leaving the system less secure than before. Fixing vulnerabilities on the other hand is a never ending cycle of downloading, evaluating and applying updates and hoping that they do not break essential system functionality and contain less exploitable bugs than before.

1.2 When is a system secure enough?

A key question for the owner is when the system is secure enough. Conventionally, this point is reached when the costs of mitigating against the remaining risk exposure is higher than the potential loss caused should the events happen. This assessment requires that one knows which risks have been (successfully) mitigated, which are still remaining, and the consequences of them happening.

This paper shows that a cyber range has many advantages over traditional methods when performing such a comprehensive assessment of a system. Moreover, it enables testing of different mitigation strategies without risk to the actual system to find an optimized balance between costs to secure a system and remaining risks.

2 System security assessments

For systems of even just moderate complexity performing a comprehensive security assessment is surprisingly hard. The reasons for this difficulty are briefly:

1. Emergent behaviour of complex systems is caused by dynamic interaction of system components with each other and their environment. This frequently creates unexpected side effects, subtle dependencies and couplings which modify the overall behaviour in ways that are impossible to predict by static analysis or testing individual components in isolation.
2. System behaviour depends on policies and processes, which should be properly planned, documented and implemented in terms of dependencies and consequences, but frequently are not. Moreover policies and processes change even more often than system hardware and configurations do and the act of changing a policy or process in a running system is a frequent source of hard to predict problems.
3. Systems interact with other external systems and load and timing issues of these interactions may change the way the system behaves and expose complex dependencies and cascading failures.
4. By accident or malicious action, system users, administrators and maintainers may act in ways that are completely unexpected and thereby inadvertently breach some of the inherent assumptions of the system design.

Traditionally, security assessment on a system-wide scale uses audits and penetration testing. Both methods have shortcomings that led to the concept of the cyber range.

2.1 The cyber range concept

A cyber range is a facility allowing a model of an IT system to run in a simulated environment to perform tests and measurements that are applicable to the real world.

Cyber ranges come in various forms. The distinguishing factors are the size and complexity of the supported model, the level of detail and realism of the simulated environment and the scope of possible tests and measurements. It is important to note that the usefulness and suitability of a cyber range for testing security of a particular system is not just a function of the range hardware; a key factor for success is the skill and experience of the range engineers in creating the model, the environment, and the instrumentation.

Many cyber ranges are purpose-built for a very narrow band of tests. For example anti-virus companies have labs that allow them to investigate the behaviour of new malware in typical scenarios such as home computers or offices. While these labs can be called cyber ranges, they are too small and specialized to be able to test security of a wide range of systems and infrastructure.

At the core of a general-purpose commercial cyber range such as Northrop Grumman's Federated Cyber Range (FCR) in Fareham, UK, is a computational grid formed by powerful general-purpose servers interconnected through high-speed networks. Cyber range engineers transform this grid into a representative model of the system to be tested and a simulation of the environment required by the system's external interfaces.

Specialist tools to record, analyse, and replay network traffic as well as high-volume network traffic generators help with the simulation of the environment. A key capability is to effectively wipe the computational grid clean and quickly redeploy the model and the environment simulation. This "reset" capability is an important feature of cyber ranges because it allows tests to be unconstrained by possible malware infections or system crashes compromising data integrity.

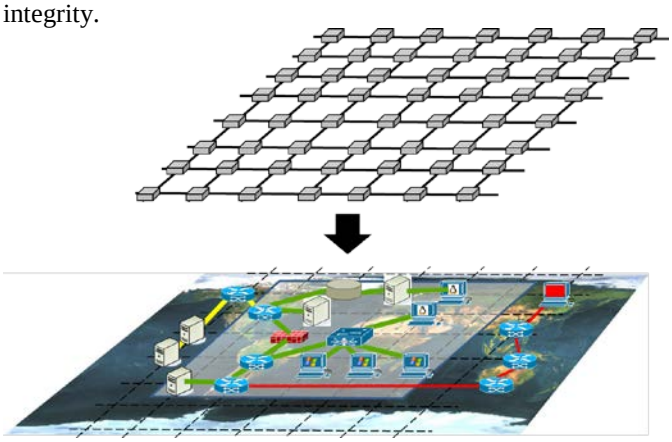

Figure 1: Transformation of a computational grid (top) into a representative model of a system and its environment (bottom). The boundary of system model is represented by the shaded area while the simulated environment is represented by the other devices.

2.2 Security assessment using cyber ranges

If implemented correctly, cyber ranges such as the FCR can be very useful in assessing the security of a system, or its resilience against attacks, overloads, or cascading failures, because cyber range testing addresses the previously cited reasons that make traditional system testing hard.

2.2.1 Testing effects of emergent behaviour

An accurate system model in the cyber range can exhibit similar emergent behaviour as the real system. This enables testing of what-if scenarios, for example by deliberately introducing component failures or disconnections, or by adding new components that were planned for future expansion.

A good example of such tests is investigating shifts in traffic workload patterns. Adding new components such as workstations in an office or new sensors or reporting thresholds in an industrial control system can cause

unexpected critical traffic flow changes in parts of the system quite remote from the location where the new components were added. This in turn can make a single router in yet another part of the system a critical component. If it fails or is shut down for maintenance it triggers a cascading chain of routers trying to redirect traffic around the problem causing other routers to fail in overload and potentially a collapse of most of the system's network in seconds.

Cyber ranges are ideal for such studies of traffic bottlenecks, hot spots, failovers and what-if scenarios because they have the required tools and instrumentation and crashing the system model does not harm the actual system. It is straight forward to simulate adding new components to a system and measure the resulting shifts in traffic flow across the system. Without any costs to the real system, it is then possible to study the effects of, for example adding a backup path for the traffic.

2.2.2 Testing effects of policies and procedures

Testing in the cyber range can include examining system policies on two levels:

Firstly, the range allows testing the effects of policy settings on system behaviour; for example, if access rights to a shared file system are tightened, it may have the desired effect of preventing access by rogue employees or infected workstations. However, it may also stop legitimate users or legacy applications from working. In the cyber range such policy trials can be performed safely without causing a storm of upset users or customers.

The second level of testing verifies whether policy settings are actually reaching all places in the system. In larger systems the policies are centrally managed and rolled out; however, it is often not clear or verified whether they reach all remote corners of the system. There are a surprising number of failure points, ranging from misconfigurations of filters and time inconsistencies, to making policy definitions accidentally inaccessible to some parts of the system. The problem is compounded if systems use a diverse set of hardware and operating systems, such as network equipment from different manufactures and mixtures of Windows, UNIX and other operating system flavours. This happens when two completely different system infrastructures are combined into a larger system; for example, after a company merger or acquisition. The compactness of the cyber range and the high level of observation it allows, enables these problems to be investigated and resolved much more easily than it would be in a geographically distributed real system.

2.2.3 Testing overloads and malware effects on interfaces

Another area where cyber range testing is ideal is evaluating the system's response to unplanned activities coming from its external and internal interfaces. Traditional penetration testing of the real system would run the significant risk of damaging the operational system particularly if they tried interface attacks in the same unconstrained way a real attacker would. In real life, external systems hardly oblige to test requests, especially for those requiring overloading or malicious traffic. A cyber range can easily test internal and external system interfaces in overload and with potentially damaging levels of malware such as packets that may cause crashes in poorly implemented network interfaces, packets that contain viruses, Trojans and other exploits of system vulnerabilities.

A recent example of such testing in the FCR found an intrusion prevention device deployed in a system model that could be made to fail open when subjected to the right kind of overloading. It would simply give up and pass all traffic through, good or bad. This is not something you would want to find out in a real system under attack.

2.2.4 Testing effects of user interactions

One of the big advantages of system testing in a cyber range is that it allows users to interact with the system just as if they would be in the real world. For example, the FCR contains dedicated offices with workstations that can be made part of the model. This allows real users and administrators to take part in tests.

There are two common scenarios that are frequently used. One is to have administrators defend the model of their own system using their own tools and processes against attacks staged by simulated hackers. This improves their abilities and training and puts response plans and disaster recovery procedures to the test.

The other case involves testing the system behaviour, incident handling and recovery procedures against insider actions by malicious or careless system users. Examples of such tests that have been performed in the FCR are sending deliberately infected emails that can pass the system's anti-virus defences and then have users open them as they might in the real world. Other examples include crafting infected web servers in the simulated external world and studying the effects of users being lured to these pages, or simply giving users an infected USB stick "found in the parking lot", and examine the effects of it being plugged in.

2.3 The challenges to cyber range system testing

Testing in the cyber range uses a model of the system: Therefore it is essential to assure this model is accurate or the results of the test will not be applicable to the real system.

The first challenge is to find out enough detail about the real system's composition. In the FCR we call this the discovery phase. We have found that this is surprisingly hard; typically because systems have often grown over long periods and are poorly documented, IT departments and their institutional knowledge have shrunk or outsourced, and the architecture may have gone through several acquisition and merger phases.

As a result, there is often a significant effort required during discovery to re-create or validate a reasonably accurate description of the architecture, software versions and patch levels. In some cases, this involves actual probing of the real system with network mapping tools.

The second challenge is deciding on how much virtualisation will be used. There are great advantages in using virtual machines (VM). In many cases, the real system will use VMs already and using copies of the original VMs for the model in the cyber range ensures that all versions, patches, and configuration settings are identical to the real system.

There are, however two issues that must be checked before deciding on using copies of original VMs. Firstly, the VM may contain confidential or personal data that must be removed and replaced with test information, and secondly, it is important to check if temporarily using a copy of the VM for testing is permitted under the licensing agreements of the operating system and applications.

Unfortunately, it is sometimes easier to just build a new VM from scratch in the cyber range and apply the same patches and configurations than trying to solve these issues and using a copy.

Virtualisation is generally used in cyber ranges to create the simulated environment around the system model or to create portions of the model for which the level of detail is not very important. For example, in the FCR we often create whole office departments as collections of hundreds of individual workstation VMs running on a few physical servers from the computational grid.

When deciding on the level of virtualisation to be used in a model, it is important to consider its side effects: Running multiple VMs on the same physical machines or using multiple virtual LANs on the physical LAN inevitably causes some undesired coupling. If the CPU load of a VM goes up, there might be a slowdown in the processing of the other machines on the same hardware. Similarly, sending large volumes of traffic over a virtual LAN reduces the bandwidth available for all other virtual LANs on the same cable.

For items that cannot or should not be virtualised, it might be necessary to use the actual or identical hardware in the FCR. One case when this is generally necessary is to measure physical performance parameters such as CPU load or packet throughput. For this reason, a cyber range usually has provisions to temporarily house additional hardware and the necessary connectivity to include it into the model.

A consequence of the need to include some real hardware into a model is that physical performance tests of very large systems may have to be done in parts because not all additional hardware can be mounted in the range. In such cases, a number of models are created for the system. Each of these models specialises in testing a particular aspect of the system for which the model is designed to be highly accurate, including additional hardware as needed. The system parts outside the test focus are then represented in the model in a less detailed, virtual manner.

Discovering the system architecture and creating representative system models clearly requires some effort. The economics of cyber range testing become much better if the models are reused multiple times. For this reason, a cyber range usually has the capability to archive models offline and restore them when needed. A year later, when a system reapplies for a repeat security assessment, it is much faster to update an already existing model with the latest changes than to recreate a model from scratch.

The final challenge to commercially available cyber ranges is to ensure security and privacy of the current user of the range. Not only does this mean physical security and strict access control, it also requires thorough procedures to clean the range between users. Depending on the security level of the test, these measures start with secure erasing of disks between activities to using separate ranges for higher classification levels.

2.4 Beyond cyber security

Cyber ranges have a significant use in research. Since its opening in 2010, the FCR has been used for a number of research studies and cyber experiments. Examples of unclassified studies using the FCR include research in building emergency management [1] and evacuation strategies [2,3]. Another study concerned providing reliable communications in emergencies over largely unknown networks [4,5,6,7].

In these studies, the computing grid of the FCR was used to run a model simulating buildings, their transport and communication infrastructure and the actions of simulated people during emergencies.

4 Conclusion

It is clear that system security assessments in a cyber range are superior to the traditional methods if a sufficiently accurate model can be created.

However, the challenges to the model designers are significant. There are many system architectures ranging from legacy mainframe systems to cloud computing. Similarly, there are many flavours of security assessments from resilience against external or internal attacks, verifying compliance with regulations, to testing effective incident response and recovery by system administrators.

One of the key challenges in our experience is the lack of sufficiently detailed knowledge by the system owners about their own system. We need this information to create the model and therefore added the discovery phase into our process for cyber range engagements. Somewhat

unexpectedly, the first benefit of doing a security assessment in the cyber range comes even before the model is built, just from our discovery work investigating what the real system actually looks like. Once the model is created, the cyber range can deliver a much more comprehensive assessment than would be safely possible by traditional system testing.

Acknowledgements

This paper is based on the experience of building and running Northrop Grumman's Federated Cyber Range (FCR) in Fareham, UK. The FCR is based on Northrop Grumman's USA-based cyber ranges which have been operational since 1998. The FCR was partly funded by the UK Technology Strategy Board with sponsorship from both the Centre for the Protection of the National Infrastructure (CPNI) and the Engineering and Physical Sciences Research Council (EPSRC) as part of the SATURN (Self-organising Adaptive Technology Underlying Resilient Networks) programme. The SATURN programme is led by BT and includes Imperial College London and the University of Oxford as academic partners.

References

[1] N. Dimakis, A. Filippoupolitis, and E. Gelenbe. Distributed building evacuation simulator for smart emergency management. The Computer Journal, 53(9): 1384–1400, 2010.

[2] A. Filippoupolitis, G. Gorbil, and E. Gelenbe. Autonomous navigation systems for emergency management in buildings. Proceedings of the 54th Annual IEEE Global Telecomunications Conference (GLOBECOM '11), pp. 1–6, Houston, Texas, USA, December 2011.

[3] A. Filippopoulitis, G. Gorbil and E. Gelenbe. Pervasive emergency support systems for building evacuation. Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on, pp. 525-527, 19-23 March 2012.

[4] G. Gorbil and E. Gelenbe. Resilient emergency evacuation using opportunistic communications. Proceedings of the 27th International Symposium on Computer and Information Systems ISCIS 2012, October 3-4, 2012, Springer Lecture Notes in Electrical Engineering, to appear.

[5] E. Gelenbe. Networks in Emergency Cyber-Physical-Human Systems. Opening Keynote Paper, ICCCN 2012, July 30-Aug. 2, 2012, Munich, Germany.

[6] E. Gelenbe. Steps toward self-aware networks. CACM, 52 (7), pp. 66-75, July 2009, DOI="http://doi.acm.org/10.1145/1538788.1538809"

[7] R. Lent and E. Gelenbe. Cognitive packets in large virtual networks. Proceedings of the 27th International Symposium on Computer and Information Systems ISCIS 2012, October 3-4, 2012, Springer Lecture Notes in Electrical Engineering, to appear.