National Cyber Security Centre
Ministry of Justice and Security

Managing edge devices

Five challenges and recommendations when using edge devices

10 June 2024, version 1.0

Introduction

In today's world, many organisations use edge devices. These systems operate at the edges of the network and involve security products such as firewalls, VPN servers, routers and SMTP servers. While these devices can make your organisation more productive, many users are not fully aware of the security aspects of the edge devices themselves. After all, edge devices often operate in areas that are not overseen by modern Endpoint Detection and Response (EDR) solutions. Recent incidents confirm the rising trend in attacks on edge devices by malicious actors. In this factsheet, the National Cyber Security Centre (NCSC) takes a closer look at some of the challenges and threats around edge devices and how organisations can manage the risks.

Introduction

Edge devices have become an integral part of our modern digital infrastructure. They are essential to secure hybrid working (VPN systems), monitoring network traffic (firewalls) and a host of other key operations. The nature of their features means that edge devices are usually publicly accessible over the internet, operating at the edges of the network. You may not be aware that internal systems that are publicly accessible over the internet can also be viewed as edge devices.

Despite the vital role these products play in network security, following recent incidents and vulnerabilities that have been identified in different edge devices, we can see that they often fail to satisfy modern security-by-design principles.

Edge devices have long been an attractive target for malicious actors. State actors invest capacity and resources in researching and identifying vulnerabilities. When attackers find and exploit vulnerabilities, they can gain access to an edge device – and thereby also to the wider network – without being discovered. According to the Dutch Military Intelligence and Security Service (MIVD) annual report, the COATHANGER spy software that was found on Fortigate devices at the Ministry of Defence[1] ultimately compromised more than 20,000 Fortigate devices around the world.[2]

Edge devices: a popular target

In our era, when practically every organisation uses at least one edge device, malicious actors can achieve big returns from conducting technical research and looking for vulnerabilities in these products. A wide range of vulnerabilities have been uncovered in edge devices.

In its 2023 annual report, the MIVD discusses how Chinese actors are increasingly targeting edge devices (specifically VPN systems),[3] but the threat is not limited to China: to give just two examples, Russian actors[4] and criminal ransomware groups such as Akira[5] have also achieved initial access by exploiting VPN systems.

Intended audience

This factsheet is intended for use by people working at a tactical level within the organisation that wants to understand the potential risks and threats of edge devices for their organisation.

[1] For more information on COATHANGER, please see: https://www.ncsc.nl/actueel/nieuws/2024/februari/6/ nieuwe-malware-benadrukt aanhoudende-interesse-in-edge-devices (in Dutch)

[2] MIVD annual report 2023: https://www.defensie.nl/downloads/jaarverslagen/2024/04/18/jaarverslag-mivd-2023 (in Dutch)

[3] MIVD annual report 2023: https://www.defensie.nl/downloads/jaarverslagen/2024/04/18/jaarverslag-mivd-2023

[4] APT44 blog by Google Mandiant: https://services.google.com/fh/files/misc/apt44-unearthing-sandworm.pdf

[5] The USA's Cybersecurity & Infrastructure Security Agency (CISA) and NCSC-NL advisory on Akira: https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-109a

Not the target – but still a victim

Several critical vulnerabilities in the Ivanti Connect Secure VPN solution were exposed in the first quarter of 2024. The security company Volexity found that malicious actors had exploited two unknown 'zero-day' vulnerabilities in certain systems.[6]

When alerts were published about these zero-day vulnerabilities, the actors switched to a new strategy, shifting from highly targeted zero-day exploits to a large-scale, global form of exploitation that was more opportunistic.
Every vulnerable device became a target. Sometime later, other malicious actors got the chance to exploit these vulnerabilities too when a public Proof of Concept (PoC) code was published.

This case shows how anyone can become a victim, even if your organisation is not the initial target. The simple availability of a vulnerable system or product that can be accessed over the internet opens up an organisation to the threat of opportunistic attacks.

Once an edge device has been successfully compromised

When malicious actors compromise an edge device, they have several options.

They may decide to install a backdoor on the device to achieve 'persistence' and continued access. These backdoors can persist, even after patches are installed and the device is rebooted. Attackers can also modify logging features, or even disable these features altogether, making it more difficult to detect suspicious behaviour.

Alternatively, an edge device can serve as an access point to the wider network. For example, a malicious actor can use 'lateral movement' to enter a database containing critical data or other crown jewels and then exfiltrate these data at a later time, whether through the compromised edge device or by other means.

Finally, edge devices often process sensitive data such as users' login details. Malicious actors can exfiltrate these data, then come back at any time and log into the system using an employee's credentials. This makes it difficult to distinguish between malicious and benign activity on the network, giving a malicious actor more freedom of movement.

Living-off-the-Land techniques hinder proper detection

Living off the Land (LOTL)[7] involves the use of legitimate tooling and applications in the victim's network to conduct malicious campaigns. For example, attackers can use PowerShell to execute malicious code or Active Directory tooling to obtain new or existing login details.
As these are legitimate applications, they are not usually automatically blocked, especially if the login credentials are valid. Edge devices frequently use these kinds of applications.

Malicious actors are making ever-greater use of LOTL techniques to effectively evade detection.
Many edge devices also contain multiple software libraries and applications that can be used for other purposes once the device has been compromised. If your organisation lacks the right tooling to detect anomalous behaviour in legitimate applications, you have a blind spot.

Monitoring and logging is complex and often poorly configured

If anything, the situation is even more urgent, as recent case histories show that edge devices operate beyond the scope of traditional EDR solutions. This means that organisations themselves have to configure the logging and monitoring processes for their own edge devices, as well as safeguarding the integrity of those logs.

Experience has shown that configuring logging and monitoring on edge devices is a complex operation, and in some cases, it requires extra customisation. To give one example, it can be challenging to collect and analyse all the logs. Not all organisations have the capacity or expertise to do this properly, and supplier instructions are not always comprehensive or applicable to each organisation's specific networks.

[6] Volexity blog post on Ivanti Connect Secure: https://www.volexity.com/blog/2024/01/10/ active-exploitation-of-two-zero-day-vulnerabilities-in-ivanti-connect-secure-vpn/

[7] More information on LOTL: https://www.cisa.gov/sites/default/files/2024-02/Joint-Guidance-Identifying-and-Mitigating-LOTL_V3508c.pdf

Edge devices: challenges

In view of the severity of current threats, the NCSC has distilled five challenges from recent incidents and vulnerabilities in edge devices. After examining these issues, we provide a framework for action to tighten up the management of edge devices.

Challenge 1: The organisation's attack surface is unknown

Today's systems are increasingly interconnected, and we have a growing need to work remotely and use mobile devices. Edge devices often act as a hub between internal and external networks. When the edges of the network are not clearly defined, and it is not sufficiently clear which edge devices are being used in this marginal area, there is a greater risk of exploitation.
For example, organisations can unknowingly leave systems running vulnerable software versions for longer than necessary.

It is also important to identify which edge devices – or which aspects of those devices – are publicly accessible online. The most commonly exploited vulnerabilities can often be accessed remotely, with no need for user interaction. This limited visibility makes it difficult to identify organisational risks and turns the edge of the network into a blind spot.

Challenge 2: Edge devices are 'black boxes'

After an ill-defined attack surface, a lack of understanding of an organisation's own edge devices presents a further risk. Edge device design has repeatedly been shown to fall short of modern security-by-design principles.
In many cases it is also complex – sometimes even impossible – to access the underlying source code. This hampers third parties (such as security experts and investigators) in their efforts to verify security levels or understand which software components are present.

When vulnerabilities in Ivanti Connect Secure were exploited in the first quarter of 2024, it became clear that the VPN devices were using software packages and libraries that were years out of date.[8]

Many of these packages contain multiple critical vulnerabilities that malicious actors can exploit once they gain access to a VPN system.

Some organisations must also depend on the supplier of the edge device to monitor and fix these issues. In the case of Ivanti Connect Secure, for example, system logs could not be read by users: the log files had to be sent to the vendor for decryption. Moreover, instead of being fully accessible, software updates are often only available to customers, which makes it harder to share information. Finally, patches are not always made available quickly enough, and suppliers may not be able to provide adequate support during severe incidents. These factors make it more difficult to produce an accurate risk assessment.

Challenge 3: Misconfiguration of edge devices can increase the risk of exploitation

Each generation of edge devices is more complex than its predecessors. Whereas in the past these types of devices were often relatively simple conduits for network traffic, today's edge devices are constantly benefiting from additional features. Although this can improve the user experience, it also involves new risks.

As some new features require organisations to open additional ports or make these functionalities accessible over the internet, it makes sense that this situation would also open up new opportunities for attackers. The use of these additional ports and features expands the potential attack surface, so organisations need to spend more time ensuring proper configuration and conducting the associated monitoring and logging.

Those additional features must then be correctly configured, an operation that requires organisations to have in-house capacity and expertise. Misconfiguration could cause features to be unnecessarily accessible over the internet or lead to data breaches. As in the previous challenges, adding more features can reduce the visibility of your edge devices, as these new functions also need to be configured and monitored. Malicious actors could exploit any misconfigurations to compromise your network.

[8] https://thehackernews.com/2024/02/ivanti-pulse-secure-found-using-11-year.html

Challenge 3 in real life: SSL/TLS offloading

The range of features for edge devices is constantly expanding. One example is SSL offloading, which allows users to decrypt traffic that was encrypted using the Secure Sockets Layer (SSL) or Transport Layer Security (TLS) protocol.[9] This feature makes edge devices more effective when it comes to inspecting traffic and detecting anomalies.

However, if malicious actors gain access to an edge device and SSL/TLS offloading is enabled, the attacker may be able to intercept all unencrypted traffic that passes through the device. This is a good example of LOTL exploitation. We recommend drafting a preliminary risk assessment that incorporates the activation and addition of new features such as SSL offloading.

Challenge 4: Patch management for edge devices is often inadequate

Patch management is an important part of any cyber security policy, and it is one of the basic measures recommended by the NCSC.[10]

Recent vulnerabilities in a range of different products have revealed that malicious actors need very little time to work out how to operationalise and widely exploit these weaknesses.[11, 12, 13]

This threat of exploitation only makes it more important to manage patches as early as possible.

Some factors that increase the risk of edge devices being compromised when organisations have inadequate patch management:

- Patching impacts organisations' continuity. Although downtime in an essential component such as a VPN system during business hours is undesirable – even unacceptable – delaying the installation of patches increases the risk that a system will be compromised, especially when there is a public Proof of Concept (PoC) that increases the chance of large-scale exploitation.
- Due to the urgency of the issue, it may be necessary to install security updates on edge devices in the evening or at weekends, but not all organisations have this capability.
- Service Level Agreements (SLAs) do not sufficiently specify the supplier's obligations in terms of security updates, communication about these updates and support to fix vulnerabilities. Sometimes there are no patches available, and organisations must rely on mitigating measures that may need to be tailored to their specific situation.

Challenge 5: Recovery after an edge device is compromised takes up a lot of capacity

If you are not sufficiently equipped to deal with these incidents, you run the risk that their impact can be much greater than it might initially appear. Exactly which data were accessed? What impact has the incident had on system availability? How should you communicate about this to your customers and partners? How will this affect your organisation's reputation?

When a device is compromised, organisations have to take measures such as bringing in external Incident Response (IR) specialists, often at great financial cost. Intensive work is also required to restore the confidentiality and integrity of edge devices and the wider network. Simply installing security updates or rebooting the device is often not enough to remove an attacker's access and eliminate them from the network. In extreme cases, the only way to restore system integrity is to replace the hardware itself. In addition to the financial impact, this also demands a lot of capacity from employees and has implications for the business continuity.

In the current threat landscape, any organisation could fall victim to an incident involving edge devices. This makes it essential for your organisation to be prepared for these scenarios, and to consider the impact of these challenges and risks on the continuity of your operations.

[9] More information on SSL/TLS offloading: https://cyberpedia.reasonlabs.com/EN/ssl%20offloading.html

[10] Introduce good patch management: https://www.ncsc.nl/wat-kun-je-zelf-doen/weerbaarheid/beschermen/basismaatregelen-cybersecurity/richt-patchmanagement-in (in Dutch)

[11] Security company Mandiant explains how multiple actors managed to exploit a vulnerability in Citrix following disclosure: https://www.mandiant.com/ resources/blog/session-hijacking-citrix-cve-2023-4966

[12] Security company Volexity explains how malicious actors were able to exploit vulnerabilities in Ivanti Connect Secure on a global scale within 3 days after publication: https://www.volexity.com/blog/2024/01/15/ivanti-connect-secure-vpn-exploitation-goes-global/

[13] Blog on exploitation that targeted Palo Alto firewalls: https://www.volexity.com/blog/2024/05/15/ detecting-compromise-of-cve-2024-3400-on-palo-alto-networks-globalprotect-devices/

Framework for action

What can you do to manage these challenges? This chapter offers some advice on how to identify and manage the risk that an edge device will be exploited, and to mitigate the impact of an attack.

Recommendation 1: Make a list of the edge devices within your organisation and determine how accessible they are over the internet

The ability to manage organisational risks arising from the exploitation of edge devices starts with a clear overview of all the edge devices your organisation is using. This overview will help you understand not only *which* edge devices are in use, but also *where* they are located. Armed with this information, you can determine whether and to what extent these edge devices can be accessed over the internet.

Some ways to do this:

- Identify the attack surface by scanning for internet-based vulnerabilities. This is also known as Attack Surface Management (ASM).
- Consult different departments within your organisation, from SOC staff and technical specialists to the procurement department. Ask questions such as:
  - What edge devices does the organisation have or plan to acquire?
  - Is there an existing list of edge devices?
  - Is it kept actively up to date? And who is responsible for this?
  - Where are edge devices located in the network? To what extent are they accessible over the internet?
  - Who configures and manages the edge devices?
  - How are the edge devices managed?

Wherever possible, ensure that your overview aligns with your organisation's existing asset management processes. This makes it more likely that your organisation will continue to update the edge device overview in the future.

Recommendation 2: Understand the edge devices in your network

Once you have identified all your edge devices and the corresponding attack surface, the next step is to understand the edge device itself. This will help you decide on potential extra configuration and hardening measures, as well as showing up the areas where you are most dependent on your supplier. You can work on this together with the technical specialists within your organisation and with your suppliers.

Consider questions such as:

- Which version is currently running?
- Which features are enabled by default? Are any features not being used?
  Are these features accessible over the internet?
- Can I find out how the edge device is configured? Can high-risk features be disabled?
- Can the supplier demonstrate that the edge device was developed according to modern security-by-design principles?
- Can the supplier demonstrate that the device has been subjected to an independent test (such as a penetration test or security assessment)?
- Can the supplier guarantee that the relevant software and libraries are up to date? Can the supplier provide a software bill of materials (SBOM)?
- Am I dependent on the supplier for support during incidents, or can I find the necessary information and take action myself? How does the supplier communicate about vulnerabilities and incidents? Do I know their out-of-hours contact details?
- Where can the organisation find instructions for how to configure or download security updates? Add these instructions to a manual, and store that manual offline.
- How dependent am I on the supplier to install patches quickly?

Recommendation 3: Monitor edge devices

Edge devices are an attractive target for malicious actors. Active monitoring and detection are important security measures that will help you identify when edge devices have been exploited. As mentioned above, logging and monitoring can be complex, and you may well need to tailor these processes to your organisational context. If your organisation does not have the necessary in-house expertise, the NCSC recommends exploring the possibility of hiring external experts.

Things to consider when introducing a monitoring procedure for edge devices:

- Are periodic checks carried out on the integrity of configurations (are all the settings still appropriate?) and on new security updates?
- Are the logs collected by edge devices sent to separate, segmented storage to ensure their integrity?
- Is there active monitoring in place for suspicious behaviour on the edge device, or on the associated endpoints if it is not possible to monitor the edge device itself? Before you can identify anomalous behaviour, you first have to determine what constitutes normal behaviour. You can develop a baseline to help you do this.

Recommendation 4: Ensure you have specific patch management for edge devices

Your organisation may well already have a patch management policy in place. In view of the challenges set out above, the NCSC recommends focusing specifically on patch management for edge devices. If your organisation does not yet have a patch management policy in place, the NCSC recommends that you draft one. Patch management is an important basic measure. Check to see if your current patch management for edge devices is fit for purpose.

For example, you can find out:

- whether your organisation has considered what 'timely' patching would look like in your specific context. In practice, this is usually linked to the impact on the continuity of the organisation: for example, has the organisation decided whether this should be done within 24, 48 or 72 hours? And is there a distinction between critical vulnerabilities and regular updates?
- how the organisation receives information about new vulnerabilities. Does this information come from the supplier or through other channels?
- whether mandates have been allocated, and whether it is clear who holds which mandate. Who is authorised to decide that security updates should be implemented as soon as possible, even if this impacts the organisation's operations?
- whether there is sufficient capacity to both prepare and implement security updates.
- whether your organisation has considered the scenario: 'What if no security update is available, but a critical vulnerability exists?' What is the recommended approach in such a situation, and what impact will this have on the organisation?
- whether you have an SLA that sets out clear agreements with the edge device supplier about patching. Are there clauses the supplier must comply with, based on your organisation's patching policy? Has the supplier formally agreed to this, and are they actually capable of complying with these clauses and providing the agreed support?

Recommendation 5: Limit the impact of edge device exploitation on your organisation

There is a risk that your organisation could fall victim to a compromised edge device. If this happens, you will need to be able to act quickly to limit the damage to your organisation. Developing a victimisation or 'assume breach' scenario will help your organisation prepare for an incident.

How you can do this:

- Assume that a breach will occur. It is increasingly necessary to have a 'defence-in-depth' security strategy in place. The 'assume breach' scenario works on the assumption that the organisation will face a successful attack and serves as a starting point to test a defence-in-depth security strategy. Consider questions such as:
  - Does the organisation currently rely on a security measure that would prove to be a 'single point of failure' (SPOF) if an edge device were to be compromised?
- Are the surrounding network infrastructure and network components also set up along defence-in-depth principles? For example, what is the level of trust and rights between these components and edge devices? How are the components of the network segmented?
  - Does your organisation use its own hardening for edge devices, or do you need to consult with the supplier? Suppliers often provide instructions for these processes.
- Include the specific scenario of a successful exploitation of edge devices in your Business Continuity Management (BCM). Draft an Incident Response Plan (IRP) describing the first and most important actions to be taken during an incident.
  Here are some examples:
  - Has the organisation considered which critical processes and/or information to secure first and prioritise? Has this been agreed with the executive board?
  - Which mandates, roles, tasks and responsibilities have been assigned during an incident, and to whom?
  - Have you determined the maximum acceptable duration of an interruption to organisational continuity?
  - Is there sufficient capacity within your own organisation to handle an incident, or do you need to hire an incident response partner? What agreements can you make with this partner in terms of their availability when quick action is needed? Is your organisation capable of providing this partner with the necessary logging, or would you need to depend on the supplier to do this? Set out these agreements with a suitable partner in advance to avoid having to wait to complete a tendering process when an incident is already in progress.
- Carry out tabletop exercises and red teaming on an 'edge devices compromised' scenario to simulate, test and improve your incident response plans. This will help your organisation assess whether potential incident response plans include appropriate triggers, and whether these plans will deliver the results you want.

Published by:
National Cyber Security Centre (NCSC)
PO Box 117, 2501 CC The Hague
Turfmarkt 147, 2511 DP The Hague
070 751 5555

More information
www.ncsc.nl
info@ncsc.nl
@ncsc_en

June 2024