Trust No One? A Framework for Assisting Healthcare Organisations in Transitioning to a Zero-Trust Network Architecture

Dan Tyler and Thiago Viana

Abstract: Traditional networks are designed to be hard on the outside and soft on the inside. It is this soft inside which has made the traditional perimeter model laughable to attackers, who can easily breach a network and run away with the data without even having to deal with the hardened perimeter. The zero-trust security model, created by John Kindervag in 2010, addresses the security flaws of the traditional perimeter model and asserts that all network traffic on the inside should not be trusted by default. Other core principles of zero trust include verification and continuous monitoring of all communication, as well as encryption of all data in transit and data at rest, since the goal of zero trust is to focus on protecting data. Although the zero- trust model was created in 2010, with some of the associated security practices existing even before that, many healthcare organisations are still choosing to focus primarily on securing the perimeter instead of focusing on the vulnerabilities within them. The current COVID-19 pandemic which healthcare providers are struggling with further highlights the need for improvements to security within the network perimeter, as many healthcare providers and vaccine developers are still using vulnerable, outdated legacy systems which could become compromised and indirectly have a detrimental effect on patient care. Legacy systems which are technologically limited, as well as medical devices which cannot be controlled or managed by network administrators, create boundaries to transitioning to a zero-trust architecture. It is challenges like this that have been explored during the research phase of this project in order to gain a better understanding of how a health organisation can adopt zero-trust practices despite the limitations of their current architecture. From the information gathered during this research, a framework was developed to allow a health organisation to transition to a more secure architecture based on the concept of zero-trust. Aspects of the proposed framework were tested in Cisco Modelling Labs (CML), and the results were evaluated to ensure the validity of some of the recommendations laid out in the framework. The main objective of this research was to prove that if a host within the local area network (LAN) were to be compromised, the damage would be limited to that host and would not spread throughout the rest of the network. This was successful after the qualitative research performed in CML. One of the other takeaways from testing the framework in CML was that medical devices could be secured by placing firewalls directly in front of them. This placement of firewalls may seem like an unorthodox approach and was shown to increase latency, but the blocking of all unnecessary traffic on the rest of the network will result in a performance boost and should balance it out in a real-world application.

Keywords: zero-trust networks; healthcare; legacy systems

1. Introduction

The purpose of this study was to identify the greatest problems faced by healthcare organisations when attempting to adopt a zero-trust network architecture so that a frame-work could be developed to present practical solutions to these problems and aid the healthcare industry in securing their networks using zero-trust practices. Research has shown that although organisations would benefit from the increased security provided by adopting a zero-trust architecture, many remain hesitant to make the move. For many organisations the hesitance is due to the financial implications associated with making such a huge change to their network architecture, but for healthcare organisations, it is less of a hesitance and more of an inability to change due to the limitations of legacy systems and the medical devices which reside on a healthcare provider's network. Through research and review of the literature by various authors, it is evident that there is a research gap concerning reliable solutions to the challenges faced by healthcare organisations who wish to transition to zero-trust network architectures. The zero-trust concept encompasses not only the trust relationships between devices but also humans and the decisions they make about trust. However, this study was only focused on delving into the challenges associated with the technological implementation of zero trust in an attempt to develop solutions to these challenges. While there is an abundance of network simulation software available to be used to test zero-trust designs, these applications are limited in the fact that they are only simulations. Due to these limitations, care was taken during the analysis of the results collected from these simulations. The specifications of the Amazon Web Services (AWS) instance were also taken into account. The proposed solution is intended to be appropriate for a small- to medium-sized health organisation that wishes to implement a zero-trust architecture.

2. Literature Review

This section contains a literature review of the most relevant and important papers related to the research problem.

2.1. *The Pitfalls of the Perimeter Model and the Benefit Zero Trust Offers*

Many healthcare organisations focus on security at the perimeter and leave themselves vulnerable to attacks from the inside. In Protenus' 2020 report, it was revealed that the number of breached patient records due to attacks that originated from the inside of the net-work in 2019 was over 3.8 million, up 26% from 2018 [1]. An infamous real-world example which can be used to argue for prioritising security within the perimeter is the WannaCry ransomware attack, which had a devastating impact on the National Health Service (NHS). The WannaCry worm was able to spread across NHS devices due to the outdated Windows operating systems the NHS were using. Keeping these outdated and vulnerable systems separated from the rest of the network through microsegmentation, a zero-trust practice, would have prevented the WannaCry worm from infecting other systems [2]. While it is impossible to prevent every threat, examples like the WannaCry attack show that adopting zero-trust practices can deny attackers of a dangerous attack vector.

2.2. *Scepticism of Zero Trust within the Healthcare Industry*

Executives often perceive tight security controls such as microsegmentation to be a hindrance rather than a benefit, and they are reluctant to approve changes to infrastructure unless there is a clear advantage to doing so [3]. While zero-trust practices can feel like unnecessary obstacles initially, the benefit of the additional security against inside attacks far outweighs the cost of the practices feeling like inconveniences. However, there is no one-size- fits-all solution to isolating legacy systems on the network, and healthcare organisations with many legacy systems are naturally sceptical about adopting the zero-trust model. To alleviate scepticism over zero trust, healthcare organisations would benefit from a framework which provides tried and tested solutions to securing their legacy devices using zero-trust practices. A zero-trust framework for a healthcare organisation must be designed with additional care and attention due to the diversity of medical devices which exist within a healthcare building. These devices can include magnetic resonance imaging (MRI) scanners which, due to their age, must be operated via outdated and vulnerable operating

systems such as Windows 7 or Windows XP. Other devices such as receptionists' workstations may be more up to date, but the point is that zero-trust concepts would have to be applied on a device-by-device basis due to some systems being more limited than others.

2.3.Microsegmentation and Legacy Systems within healthcare Organisations

A full zero-trust implementation involves microsegmentation of the network, often down to the individual hosts, in order to create firewall-esque boundaries in appropriate locations. Health organisations wishing to transition to a zero-trust architecture may struggle to microsegment legacy systems which reside on their existing network. TO support these legacy systems, Uttecht [4] proposed that where integrating policy enforcement within a legacy application itself is impossible, policy enforcement should instead take place at the next best location. This could include taking a more centralised approach by traffic [4]. Haber [5] argues that while there are workarounds to including legacy systems in the zero-trust architecture, these workarounds only restrict the external access to the resources and cannot interact with the systems themselves, thus defeating the premise of the zero -trust model and being left with more resources to manage.

2.3.1 Microsegmentation of Legacy Systems Using Proxy Servers

To circumvent the management issues created by adding more hardware into the network in order to enable legacy systems, Gilman and Barth [6] proposed that proxy servers could be deployed at the application level in order to bridge the gap between legacy systems and the rest of the zero-trust network. This solution, visualised in Figure 1, involves using a forward proxy to allow legacy systems to communicate with a zero-trust system and a reverse proxy to allow the zero-trust system to communicate back to the legacy system.

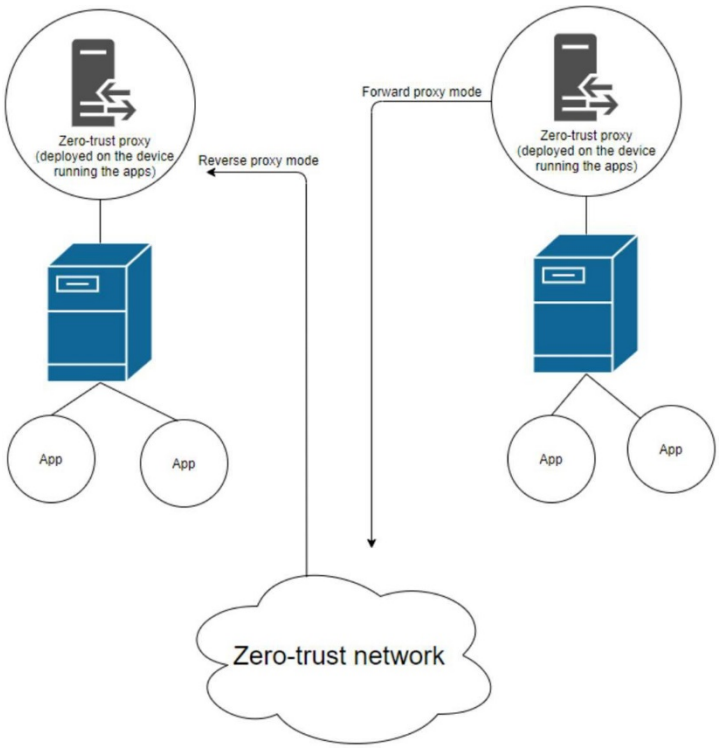

Figure 1. Proxied solution to including legacy devices in the rest of the zero-trust network [6].

2.3.2 Benefits and Drawbacks of the Proxied Solution

Using proxies to enable an unmodifiable legacy system to participate in the rest of the zero-trust network would meet the "encrypt all communication" requirement of the zero-trust model, albeit with less guarantee of its security compared with systems fully compatible with zero-trust [7]. While the proxied solution would not provide the advantage of stateful inspection like a firewall would, the performance load would be significantly reduced, which would benefit a healthcare organisation given the importance of their work. Security is unhelpful if it comes at the cost of the usability of the network, which is the main priority for healthcare organisations given the nature of the service they provide. The proposed framework will address this by putting forward a hybrid design which uses the most effective zero-trust practices in order to increase security, but it still retains some traditional network design practices in order to create a balance between usability and security.

2.3.3 Microsegmentation of Legacy Devices Using Firewalls

It is imperative that policy enforcement is applied correctly, as accidentally blocking communication from medical devices could result in the interruption of patient care [8]. As network administrators are unable to manage of control medical devices themselves, the only option is to apply policy enforcement on an external device, such as a firewall [9]. However, one firewall placed at the edge of a medical device creates a point of failure, which would have devastating consequences on patient care if the firewall were to fail. By placing a cluster of firewalls at the edge of the most important medical devices, network redundancy is ensured, and these devices are able to comply with the zero-trust model [10]. Figure 2 shows a visualisation of this proposed solution, which utilises multiple Juniper SRX340 firewalls to create a cluster of firewalls which logically act as one device.

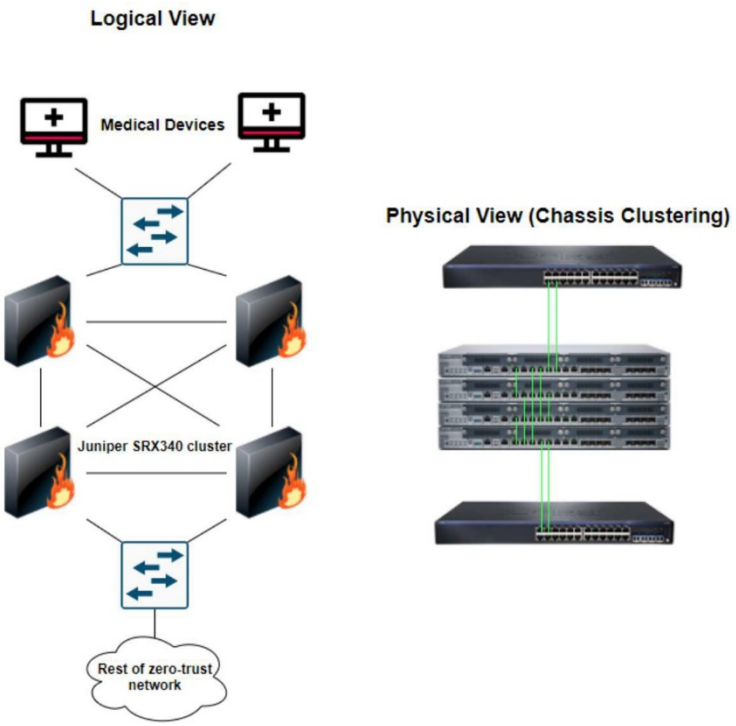

Figure 2. Fault-tolerant firewall cluster to secure medical devices [11].

2.3.4 Availability Issues Associated with Firewall Clustering

A pitfall of firewall clustering was identified by Hamilton et al. [12] when they discovered through testing that a cluster of Juniper SRX340s with deep packet inspection (DPI) enabled increased latency and reduced the speed of the network by more than 10%. It was also found that using a cluster of firewalls without DPI enabled increased latency, but considerably less so than with DPI enabled [12]. Having high latency within a network is unhelpful to a healthcare organisation, as it indirectly affects the speed at which clinicians can treat patients. Due to the redundancy issues with only having one firewall and the latency issued caused by firewall clustering, a compromise between the two will be found which allows for redundancy and a minimum round-trip time for packets sent to and from medical devices. Once an optimum solution is identified, it will be presented in the proposed framework along with the results and analysis of the testing.

2.3.5. Microsegmentation and Protection of Virtual Machines

Microsegmenting virtual servers on a physical server is possible via the use of tools such as VMWare NSX, where virtual servers can be microsegmented at the virtual network interface card (vNIC) level with firewall rules. However, as Kwon et al. [13] explained, traditional security measures such as firewalls are still vulnerable to zero-day exploits. Intrusion tolerant systems (ITSs) are able to solve this problem by maintaining the quality of service even while the systems are under attack. Virtual machine (VM)-based ITSs, such as the self-cleansing intrusion tolerance (SCIT) system, utilise a central controller and multiple hosts loaded with duplicated VMs in order to provide service during the time that the system is under attack. Kwon et al. [13] proposed an optimal cluster expansion scheme for ITSs, by implementing a robust cluster controller (RCC) with multiple hosts. The RCC is able to determine cluster sizes and the redistribution of packets by monitoring the duplicated VMs housed inside each host. This proposed expansion scheme has been tested to ensure that it can ensure the uptime of systems during attacks, such as a giant DoS attack [13]. This solution could be incorporated into a zero-trust-inspired framework for healthcare organisations, particularly as a defence-in-depth strategy.

2.4. Authentication Challenges Associated with BYOD Policies in Zero-Trust Architectures

While incorporating legacy systems into a zero-trust architecture is not straightfor-ward, devices that leave the building present different challenges compared with static devices. The prevalence of Bring-Your-Own-Device (BYOD) systems varies in the health-care industry, but organisations that do utilise BYOD may struggle to enable BYOD devices to participate in a zero-trust network because the zero-trust model is heavily reliant on user and device authentication [14]. The NHS' official BYOD policy states that BYOD is optional, but any personal devices are limited to basic systems such as email, calendar, rotas and scheduling, and web browsing. The policy also states that any personal devices are not permitted to connect to the corporate network and instead must use the guest WiFi [15]. On the opposite end of the spectrum, The Ottowa Hospital (TOH) in Canada allows clinicians to access the main repositories for patient history and diagnostic images via mobile devices. Mobile devices are assigned to employees so that they can be wiped in the case of loss or theft. However, in some cases, senior staff who wished to use their own personal devices permitted TOH to wipe their devices in the case of loss or theft [16]. Marshall [17] lauded TOH's decision to assign employees corporate devices but criticised the lack of enhanced authentication methods for the few staff who were allowed to use personal devices [17].

2.5. Issues Surrounding Current Authentication Methods for BYOD Devices within Zero-Trust Architectures

One method of enhanced authentication is through behavioural analytics, which allows continuous authentication through methods such as recognising a user's typing pattern and their usual network usage,

face recognition, and potentially voice recognition [18]. While behavioural analytics provide a second barrier to pass through after login credentials, the data would take months or potentially years to build up in order to be used reliably. Another issue with continuously monitoring and analysing BYOD devices is the issue of user privacy, as users may not feel comfortable having their personal devices constantly monitored. In institutions such as the NHS where BYOD is a low priority, a zero-trust solution could be to not allow personal devices onto the network at all. For small- to medium-sized private healthcare organisations who are reliant on BYOD, a definitive solution to including BYOD devices in a zero-trust network remains to be seen. By experimenting with different security controls in a simulated environment, a method of securely including personal devices into a zero-trust architecture will be devised and presented in the proposed framework. These security tools would include firewalls placed and configured in a certain manner and proxy servers, as previously discussed. Other traditional security controls such as access control lists (ACLs) and VLANs will still play a part.

2.6.Research Gap

Zero trust is a relatively new concept which has been hailed as the model of the future for most organisations, but as Uttecht [4] stressed, there has been a lack of inde-pendent testing for zero-trust networking practices. More specifically, there is a research gap concerning the implementation of zero-trust networking practices in healthcare or-ganisations [4]. Healthcare organisations are naturally the most vulnerable to cyberattacks due to the inherent weaknesses in their networks, which house many legacy systems and unmanageable medical devices [19]. The research conducted during this study shows that there is an abundance of general guidance for adopting a zero-trust network architecture but an absence of specific guidance tailored for healthcare organisations. As mentioned earlier, there is also a lack of independent testing for zero-trust network architectures, let alone network architectures which are suited to healthcare organisations.

The existing approaches to allowing legacy devices to participate in a zero-trust network have the advantage of securing outdated devices which are unable to be replaced, which is common in a healthcare setting. However, a disadvantage of using the firewall and router approach is that more devices have to be added into the network, which results in more devices to manage. The proxied solution addresses this issue, and this means that more devices do not have to be added to the network, but there is a lesser guarantee of security compared with the firewall and router approach. The current research suggests that these methods have not been thoroughly tested and that there may be room for a different approach.

Advocates for zero-trust networks insist that legacy devices can be included in a zero-trust network through methods such as those previously mentioned, but the current research is lacking in methods of securing medical devices through zero-trust concepts [20]. As medical devices are essential tools which must suffer as little downtime as possible, a solution to include them in a zero-trust network must be fault tolerant and reliable enough to be used in a real-world implementation.

Research has shown that BYOD policies differ depending on the organisation. For organisations that allow BYOD, there needs to be a one-size-fits-all solution for safely including personal devices in the rest of the zero-trust network. The proposed framework will provide guidance on the most effective way of including personal devices into a zero-trust network without using techniques such as behavioural analytics which invade the privacy of employees.

3 Materials and Methods

In order to create the framework, different security protocols, microsegmentation methods, access control lists, and security topologies were tested in Cisco Modelling Labs (CML) to identify zero-trust implementation strategies which would be suitable for the healthcare industry. Packets were sent between devices in order to test the reachability of each device to ensure that every device was only allowed to reach the devices they were permitted to. This was to ensure that the identified methods were compliant with the

zero-trust model. Each valid method was compiled into a framework with tailored guidance on implementing a zero-trust architecture within healthcare organisations. Guidance on how to implement each method within a healthcare setting was written, along with explanations of the workings behind each method.

To test the framework, a "current" network architecture for a hypothetical healthcare organisation was built in CML. This acted as an example of a traditional network architecture which the hypothetical healthcare organisation was transitioning from. The previously designed framework was used to implement zero-trust concepts into the "current" network architecture of the hypothetical healthcare organisation in order to create a new network architecture which met the criteria of the zero-trust model. Packet flows between devices were analysed using Wireshark to determine whether the network was behaving in accordance with the zero-trust model. Packets sent from medical devices to the rest of the zero-trust network were randomly sampled before and after the implementation of the framework. These samples were compared in an independent sample *t*-test in order to determine whether there was a significant increase or decrease in latency after the guidance from the framework was implemented into the network. Table 1 details the specifications of the AWS EC2 Bare Metal instance used to run the CML simulations.

Table 1. AWS EC2 Bare Metal instance specifications for CML experiments.

Component	Specification
Server	AWS EC2 Bare Metal instance m5. metal Processor Intel Xeon Platinum 8000 Series
RAM	384 GB
Cores	96 Logical Cores
Secondary Storage	AWS S3 Bucket 1TB

4. Results

This section discusses how the hypothetical network was built in CML, along with explanations of the zero-trust concepts that were implemented in order to create a hybrid model composed of the perimeter model and the zero-trust model.

4.1 Designing a Hypothetical Network for a Healthcare Organisation

To design the framework and to be able to build a virtual network in network simulation software, a hypothetical healthcare network had to be designed first. Due to security reasons, it was unlikely that any healthcare organisations would make their network topology publicly available, so the hypothetical network was based on an already existing framework for a hospital designed by Cisco [21]. This framework was not a zero-trust framework, and it was only used as a starting point for the topology. After viewing this design for a healthcare network and combining this with the knowledge obtained during the initial research stage, it was determined which types of devices would be included in the network topology. Figure 3 shows a brainstorm graph detailing the physical rooms to be included in this hypothetical healthcare organisation, along with potential devices that could be situated in these rooms. Figure 4 shows the base network topology using the traditional perimeter model, while Figure 5 shows the new network topology created using the framework. Table 2 gives a quick summary of each stage of the framework.

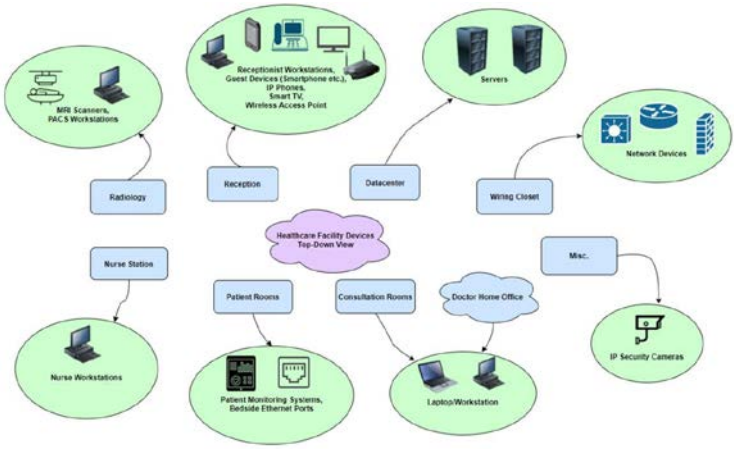

Figure 3. Physical rooms and devices contained in these rooms within the hypothetical healthcare organisation

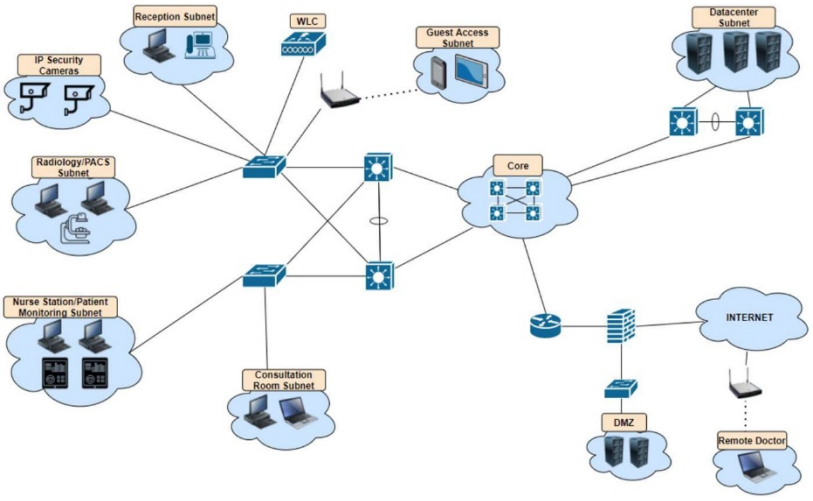

Figure 4. Base network topology for a hypothetical healthcare organisation.

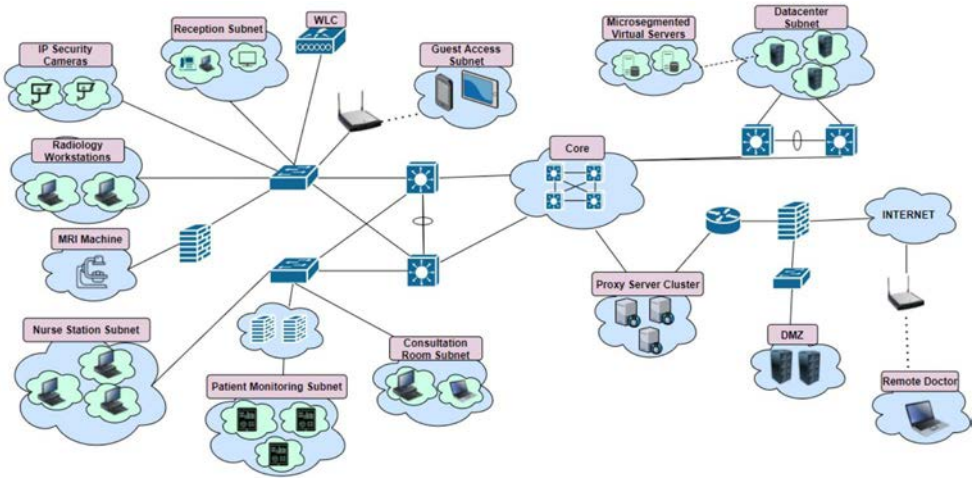

Figure 5. Summary of the framework to support healthcare organisations in transitioning to a zero-t trust-inspired network architecture with an example network topology.

Table 2. Summary of the stages of the framework from basic security to defence in depth.

Stage of Framework	Summary
Change all default credentials MFA for everyone including remote workers	
Stage 1: Basic Security	Keep operating systems and applications up to date Run anti-virus scans Change default lockout time for accounts Encrypted data at rest on the servers
Stage 2: Logging and Monitoring	Prior to applying framework, use NetFlow collector to build a map of the network traffic Identify which traffic is absolutely necessary and which traffic is not needed Continue to log all network traffic after framework is applied using NetFlow and IDS
Stage 3: Microsegmentation	Use physical firewalls to segment old medical devices (e.g., MRI, scanners, patient monitoring systems, infusion pumps, and anything which is outdated and vulnerable) Use ACLs to restrict workstations from talking to each other directly Use ACLs to ensure hosts only have access to required resources (Principle of least privilege) If virtual servers are used, microsegment them at the vNIC level using a product such as VMware NSX distributed firewall

Table 2. *Cont.*

Stage of Framework	Summary
No internet access for anyone unless it is required (e.g., radiology department does not need internet access)	
Stage 4: Further Access Control	All internet-bound traffic to go through proxy server cluster first Users on the guest WiFi network can access the internet but nothing else BYOD to be restricted to company-owned devices only, as the risk is too high for casual BYOD
Stage 5: Defence in Depth	DNS sinkholing if a host attempts to connect to a malicious domain Behavioural analytics on DMZ servers via an application such as VMware AppDefense to prevent attacks on 80 and 443 Behavioural analytics on company-owned devices used by employees from home

4.2.Stage 1 of the Framework: Basic Security

Although the first stage in this framework consists of only the most basic security controls, it is worth briefly mentioning because zero trust as a concept focuses partly on cyber human factors. While the framework focuses mostly on security at the network layer, zero trust as a concept relies heavily on multi-factor authentication (MFA) in order to be able to succeed. The network design and configuration in the framework only does half of the job; the rest lies within the hands of the humans who have access to the systems within the organisation [22]. Due to the limitations of Cisco Modelling Labs (CML), MFA was unable to be implemented into the final network design, with the focus instead primarily being on the network layer. It is still worth mentioning that basic security practices such as implementing MFA, changing default credentials, and changing lockout times for devices play a significant part in a zero-trust architecture.

4.3.Stage 2 of the Framework: Logging and Monitoring

This stage is related to the need for monitoring the network traffic before the implementation of the framework, as well as the importance of continuous monitoring of traffic after the framework has been applied.

4.3.1. Building a Map of Ordinary Network Traffic

Logging and monitoring of all network traffic is an essential part of a zero-trust architecture both before and after implementing zero-trust concepts. Before a healthcare organisation implements the proposed framework, they should log all network traffic for a period of time in order to determine the normal flow of packets between devices. While the traffic logging takes place, the network administrators should work with the rest of the staff in the healthcare organisation to create a list of what systems need to be able to communicate with each other. Once these two steps are completed, the network administrators will be able to examine the recorded traffic to discover which devices are unnecessarily communicating with each other. The goal of this is to identify the absolute minimum resources staff require to do their job in order to enforce the principle of least privilege (PoLP) when implementation time arrives. The second goal is to identify traffic which makes no sense and is better off being blocked to avoid security risks, as well as adding the benefit of increased network performance.

4.3.2. Running Mock Malware Tests

Alongside logging the network traffic and creating a map of packet flows, a test should be run involving mock malware to identify how far real malware could spread throughout the network. Many healthcare organisations still use outdated and vulnerable Windows operating systems to run legacy applications or control medical devices which are not supported in more current operating systems. Knowing this fact, it is recommended that the network administrators run a mock malware incident by introducing the EICAR test file onto one of the Windows workstations to see how far it spreads throughout the network with the current architecture and controls in place. The European Institute for Computer Antivirus Research (EICAR) test file is a DOS program containing a 68-byte ASCII string which, when executed, will print “EICAR-STANDARD-ANTIVIRUS-TEST-FILE” on the screen. The file can be downloaded to any workstation, but a script has to be written in order to allow the EICAR test file to spread between hosts [23,24].

Multiple tests should take place where the EICAR test file is introduced to one host in a different subnet each time in order to identify which sections of the network are most prone to causing malware to spread. In the case of the network simulations on CML, the EICAR test file was unable to be introduced to the hosts due to the limitation of only having access to Linux machines. The EICAR test file would not execute on Linux machines as it is a DOS program, but manual tests such as port scans were still carried out to showcase what information an attacker on a compromised machine would have access to. As the base network was built using traditional networking practices with the bulk of the security being at the perimeter firewall, an attacker on a compromised machine within the LAN could port scan subnets to discover the hosts, open ports, and operating systems being used. Because it was difficult to build a network architecture that would match perfectly with

that of a real healthcare organisation, the security controls in the base network design were based on assumptions. It is recommended that healthcare organisations run their own mock malware incidents to identify how much damage would be caused to their network, as it will differ from organisation to organisation depending on their network device configurations.

4.3.3. Continuous Network Monitoring

While using this framework will significantly limit the likelihood of a host becoming compromised and limit the damage potential for the rest of the network, it is still essential to monitor all network traffic in case of an attack. The network monitoring should not just take place at the perimeter or in certain sections of the network; it should take place everywhere, and all packet flows should be sent to a NetFlow collector to be analysed by network administrators. Intrusion detection systems (IDSs) are a common feature of traditional network designs, and this should also be a feature in a zero-trust design. IDSs would usually be deployed close to the perimeter. This should continue but with extra deployments around the network to cover intra-host traffic. In the case of medical devices shielded by a firewall, the firewall should be configured to actively perform threat detection and shun hosts that display malicious activity.

4.4. Stage 3 of the Framework: Microsegmentation

This stage is related to how healthcare organisations can microsegment their networks and overcome the challenges
presented by outdated operating systems and medical devices.

4.4.1. Implementing the Principle of Least Privilege (PoLP)

Stage 2 of the framework makes clear the importance of monitoring the network and communicating with non- technical employees in order to build a map of which traffic should be allowed. Because this framework is designed for the healthcare industry, the importance of building this map cannot be stressed enough, as lives can and do depend on the network administrators ensuring that essential traffic is not blocked. Once it has been identified which systems the healthcare professionals need access to in order to carry out their job, network access control can be implemented using access control lists (ACLs) and port-based access control. Care should be taken to ensure that network traffic from the nurse stations and doctor consultation rooms can only pass through to the servers on the appropriate ports. Apart from the data sitting on the servers in the data centre, no workstation in the building should be allowed to directly communicate with another workstation. This is to prevent malware spreading in the event of an inside attack, and these rules can also be configured via ACLs. In Figure 5, it can be noted that in some sections of the network, each workstation has a bubble around it. This is to represent that while the workstations may be in the same room, they cannot communicate directly with each other.

4.4.2. Physical Firewalls to Microsegment Medical Devices

One of the challenges the healthcare industry faces in regard to zero trust is how to secure medical devices which are usually outdated and vulnerable. It was discussed earlier how there were possible solutions to this problem, with one involving proxy servers and the other involving physical firewalls. The benefit of the proxied solution was to decrease the performance load, which would result from using firewalls, but the downside was that firewalls are superior in security and can perform stateful packet inspection.

Both of these solutions were tested in CML to determine the security they offered, as well as any increase in latency. After analysing the results, the firewall solution came out on top for security and latency. Proxy servers still found a place in the framework by being used to provide secure internet access. The problem with using proxies to secure medical devices was the high latency, which is not ideal given the nature of the industry this framework applies to. The latency displayed in CML was significant enough to rule the proxied solution out due to concern about how high the latency could be in a real-world implementation of this framework. Latency increased when using the firewall solution, but this was nowhere near the level of latency displayed

by the proxied solution. Because devices such as patient monitoring systems and infusion pumps are life-saving and any interruption could be detrimental to patient care, a firewall cluster using two firewalls was tested. There was no significant increase in latency when both of the firewalls were operating in load-balancing mode. Having said this, there was a significant increase in packet loss when one of the firewalls was shut down and all of the traffic was directed through the failover firewall. After testing this scenario many times, it was unclear why this happened. Having said this, a firewall cluster causing some packet loss is better than one firewall suffering problems and causing complete packet loss. In any case, a healthcare organisation wishing to implement this framework should test both of these methods in their own networks before committing to a choice.

While the workstations within the nurse station are blocked from communicating directly with each other, the firewall should be configured to permit traffic from all of the nurse workstations on the appropriate ports and to deny all other traffic. Despite the name "zero trust", this configuration is technically placing a lot of trust in the nurse workstations. To defend against one of these hosts becoming compromised, the firewall should be configured to only allow access from users which are in the Active Directory. Because the nurses will have used MFA to log in already, the risk of a malicious user operating one of the machines is lower. Nonetheless, the firewall should still monitor the outside interface for malicious activity such as removing access for hosts with inconsistent MAC addresses.

4.4.3. Microsegmenting Virtual Servers

The server subnet should be microsegmented in the same way as the workstations. For example, the picture archiving and communication server (PACS) should not be able to communicate with anyone but the workstations in the radiology department. There is no need for the PACS server to be able to communicate with a web server, an email server, and so on. To make the segmentation of the network as granular as possible, healthcare organisations hosting virtual servers on their physical servers should use a product such as VMware NSX's distributed firewall to control access at the vNIC level. The NSX-distributed firewall can also authenticate users from the Active Directory, meaning that no physical firewall is required for this section of the network and resources can be saved [25]. VMware is a paid product which was unable to be used during the CML simulations, but the product is well documented by VMware and is easy to implement.

4.5.Stage 4 of the Framework: Further Access Control

This stage is related to secure internet access, what devices should have access to the internet, and solutions for healthcare organisations who use BYOD.

4.5.1. What Devices Should Have Access to the Internet?

In 2019, it was reported that more than 36,000 medical devices were easily discoverable on Shodan, a search engine which can be used to find specific types of devices which are connected to the internet [26]. If medical devices are not protected properly, Shodan will list all the information attackers will need to be able to break into the network. In this framework, the focus is more on whether certain devices within a healthcare building actually need to connect to the internet at all. In early 2021, attackers were able to gain access to 150,000 IP security cameras, some of which belonged to hospitals and health clinics. The question is this: why do IP security cameras need to have internet access? In this scenario, this framework would recommend that IP security cameras should send the 27 videos they record to an internal server which is only accessible by the appropriate staff members, such as security or management. This same thinking should be applied to the rest of the devices in the building. The radiology department does not need internet access, the nurse station workstations may need internet access depending on the requirements of the organisation, and the receptionist and doctor workstations will likely need internet access. When it has been determined which devices need internet access, the next step is making sure they do not bring anything unwanted back onto the network.

4.5.2. Secure Internet Access

Hosts which require internet access, such as some of the examples listed above, should have their internet-bound traffic passed through a proxy cluster before leaving the perimeter of the network. The reason for a cluster of proxies is for load balancing, as using proxies can slow down internet browsing. However, the proxy servers will cache frequently visited websites to increase browsing speed for future visits. Even with caching, the latency will still be higher than if no proxy server is used, but the benefit of improved privacy and security is worth this small cost. The life-saving devices and important data can be accessed quickly over the high-speed zero-trust internal network, but any data accessed over the internet can be considered less important and therefore not require a high-speed connection.

4.5.3. BYOD Solution

To address the issue of BYOD policies within healthcare organisations using a zero-trust architecture, it was decided that the devices employees could bring to and from work should only be company-issued devices. The first purpose this serves is being able to configure the devices beforehand to ensure only certain applications and websites can be used or accessed, as well as being able to remotely wipe the device in the event of it becoming lost or stolen. The second purpose this serves is that the devices can be analysed ethically using behavioural analytics to ensure the person using it is an employee of the healthcare organisation. Monitoring the behaviour of an employee's personal device would be unethical, as it would infringe on their privacy.

Advocates for zero trust often suggest that VPNs have no place in a zero-trust architecture due to the risk associated with a remote device having access to the internal network, as it is difficult to authenticate the employee using it. This framework challenges this and encourages the continuation of remote VPN access, because other aspects of the framework such as access control and behavioural analytics lower the risk of an attack and the damage an attack would cause. Considering that BYOD policies in healthcare organisations vary so much, it is up to each organisation to consider whether this BYOD solution would be useful for them or whether it can be ignored.

4.6.Stage 5 of the Framework: Defence in Depth

This stage is related to implementing additional security measures to create defence in depth for the network. The measures listed here can be considered the last line of defence if all else fails.

4.6.1. DNS Sinkholing

Using object management on the firewalls, a DNS sinkhole should be configured and applied to the DNS policy for internal and remote users. If all other security mechanisms fail and an attacker manages to gain access to one of the hosts on the network, the DNS sinkhole will intercept DNS requests sent to malicious domains and return a false result. This DNS sinkhole also serves as a security mechanism for the staff in the event that an employee accidentally visits a malicious domain.

4.6.2. Secure Internet Access for BYOD Devices and Remote Users

Users who take company-owned devices home for work should be able to access the internet as well, but this should not be implemented via a split VPN tunnel. Instead, the remote user will be assigned an IP address from the internal network, and they should have their internet-bound traffic directed through the proxy server cluster and out of the perimeter firewall in the same way that all internal users do. By using this security mechanism alongside behavioural analytics of company-owned devices, the devices taken home by employees will behave as similarly as possible to devices hosted inside the internal network.

4.6.3. Securing DMZ Servers

A DMZ is a feature of most traditional networks, and this feature is carried over to this zero-trust

framework. The only problem associated with servers on the DMZ subnet is that while they will have almost all of their ports closed, ports 80 and 443 will still be open, and attacks can potentially come through via these ports. Although this is unlikely due to the rest of the security controls in place and any damage would only be limited to the server itself, defence in depth can still be applied by using behavioural analytics on the web servers. This involves using a product such as VMware's AppDefense, which will monitor the processes running on the server for a few weeks and then restrict the device if any unfamiliar processes pop up.

4.7. T-Test Results

The quantitative research carried out during the creation of the framework was focused on measuring the latency medical devices would have when different microsegmentation methods were used. For this, ping tests were used to measure the roundtrip time (RTT) of packets before and after the addition of a fully configured firewall or proxy server. The RTT was measured in milliseconds, and the results from this analysis should not be presumed to be an exact representation of the latency in a real-world implementation of the framework, as a real network would have a constant flow of traffic. The results from these tests will instead give healthcare organisations a better idea of what percentage the latency can be expected to increase by.

The objective of testing the latency was to find out which was the most appropriate method of microsegmenting medical devices and vulnerable workstations. In this scenario, the tests were pinging the PACS workstation from the device representing the MRI scanner and pinging one of the nurse workstations from the device representing one of the patient monitoring systems.

The tests to measure latency involved sending 60 packets from one of the devices representing the medical device to a workstation representing either the PACS workstation or a nurse workstation. To make it clearer where the devices are in this topology, Figure 6 shows the flow of packets in these tests. Only the patient monitoring device was used in the tests, but since the topology was identical to that of the radiology department, there was no need to use two different devices.

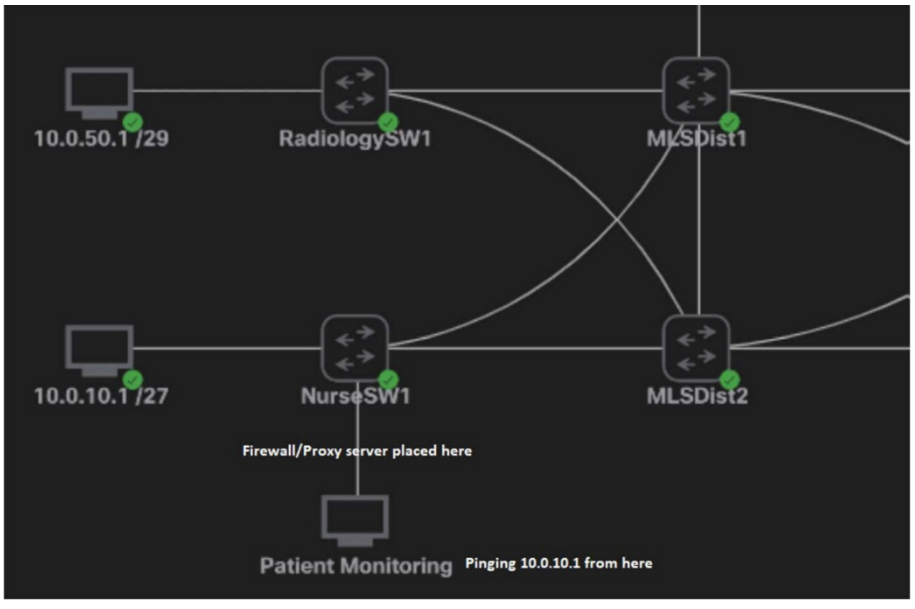

Figure 6. Packet flow and topology for the tests.

A two-tailed *t*-test was used to compare the RTT data from packets sent with and without a proxy server in place. After this, a t-test was used to compare the RTT data from packets sent with and without a firewall and with and without a firewall cluster. Table 3 provides a list of the null and alternative hypotheses formed before the *t*-test, while Table 4 shows the results of the *t*-tests.

5. Discussion

The research used a hybrid approach, combining both qualitative and quantitative research. The majority of the research conducted was qualitative, consisting of experimentation in CML to find ways of implementing zero-trust concepts using standard network devices such as routers, switches, and firewalls. The fact that the hardware and software in CML was limited was actually a benefit to the research, considering most healthcare organisations will be using older equipment, and a lot of zero-trust specific software will not be compatible with their systems running outdated operating systems. The aim of this qualitative research was to show that a traditional network architecture with an average security posture could be transformed into a hybrid architecture comprised of zero-trust concepts and traditional networking practices. The results were successful, and the proof of concept was exhibited during the simulations. If the framework is able to work using basic devices in simulations, there is a very high probability that this framework could transfer over to the real world and be used in healthcare organisations. The quantitative research was used to evaluate certain aspects of the solutions provided in the framework. It focused on measuring the difference in latency when different microsegmentation methods were used. This was important to consider because the latency for medical devices should be low enough to not impact patient care.

Table 3. Null and alternative hypotheses for each *t*-test.

Test Description	Null Hypothesis	Alternative Hypothesis
Test 1: Comparing the latency without a proxy server vs. with a proxy server	There is a significant increase in latency when using a proxy server.	There is not a significant increase when using a proxy server.
Test 2: Comparing the latency without a firewall vs. with a firewall	There is a significant increase in latency when using a firewall.	There is not a significant increase in latency when using a firewall.
Test 3: Comparing the latency with one firewall vs. with a firewall cluster	There is a significant increase in latency when using a firewall cluster.	There is not a significant increase in latency when using a firewall cluster.
Test 4: Comparing the latency with a firewall cluster vs. with a firewall cluster with one node recently shut down	There is not a significant increase in latency when a firewall cluster node converges.	There is a significant increase in latency when a firewall cluster node converges.

Table 4. *T*-test results.

	No FW or Proxy vs. Proxy	No FW or Proxy vs. FW	FW vs. FW Cluster	FW Cluster vs. FW Cluster Converged
T-Value	−36.08188	−16.21913	1.3295	−0.92
p-Value	<0.00001	<0.00001	0.186768	0.361821
Significant?	Yes	Yes	No	No

5.1. CML Experimentation Results Discussion

The initial base model for an example of a healthcare network was designed using a traditional topology and traditional security controls. Controls such as ACLs, demilitarised zones (DMZs), and VLANs were put in place, as these are all features of traditional network designs in healthcare organisations. Although the EICAR test file could not be used during the initial monitoring of the base network, scenarios were created where each host was accessed with the mind of an attacker who had compromised hosts from each department. Network scans and ping tests were carried out on these devices in order to determine how far the attacker's reach was within the network. After identifying this, the objective was to implement zero-trust concepts as effectively as possible by segmenting the network at a very granular level.

While the only available host operating system in CML was Alpine Linux, this was not a problem, as the Linux hosts were only acting as placeholders for real medical devices and outdated operating systems on a healthcare network. Having access to virtual Windows XP machines would have been ideal, but due to licensing

reasons, this was not possible. Because the devices were microsegmented using firewalls, it no longer mattered what operating system was running on the virtual hosts, as the firewall protecting the devices effectively became part of the hosts at this level of microsegmentation. To make this clearer, when a device attempts to access a host behind a firewall, they are not accessing the Linux machine, the outdated Windows XP machine, or the 20-year-old MRI scanner. Their first point of contact is the firewall, and they are effectively communicating with the firewall instead of communicating with the host machine.

All of the recommendations laid out in the framework which were able to be implemented in CML were implemented successfully. Although the recommendations such as using VMware products could not be implemented, it is safe to say that if the other recommendations were successfully implemented, then there is no reason why the rest could not be implemented in a real-world implementation. The main objective was to microsegment the network, and this was completed using firewalls which were configured using Cisco's Adaptive Security Device Manager (ASDM), proxy servers which were con-figured on Linux machines using the Squid proxy, and other traditional security controls such as ACLs which were configured on switches and routers. To test the BYOD solution, a VPN was configured on the perimeter firewall to make sure that secure internet access for the remote user worked. This solution was tested using a real internet connection by connecting the CML server to the internet via the perimeter firewall and connecting to the VPN using a Windows 10 PC. The behavioural analytics of the remote devices as well as the DMZ servers could not be tested due to costs and the fact that the network was not actually real and did not have any genuine network activity taking place every day. There are many well-respected behavioural analytics products available, and implementing these on top of the rest of the framework would not be an issue in a real-world scenario.

5.2.*T-Test Results Discussion*

In the first *t*-test, the alternative hypothesis was rejected because there was a significant increase in latency when using a proxy server to microsegment medical devices. For this reason, the proxied solution was ruled out and instead found another role to play in the framework for providing secure internet access. In the second *t*-test, the alternative hypothesis was also rejected. Having said this, upon manually reviewing the packet RTT data, it was found that the increase was not nearly as great as when using a proxy server. It was for this reason, along with the added security controls the firewall provides such as stateful inspection, that the firewall solution was used in the framework of protected medical devices and hosts running outdated operating systems. Because patient monitoring requires low latency to keep up to date with the health of patients, it would be too risky to use proxy servers which generated high latency. The latency was at an acceptable level when using a firewall, and this is the ideal solution.

Medical devices such as MRI scanners which are not actively monitoring a sick patient's status do not require heavy network redundancy. For this reason, it is safe to only use one firewall to segment MRI scanners from the PACS workstations. In the case of patient monitoring machines and infusion pumps, redundancy is critical because lives are depending on the medical machines. Firewall clustering was implemented in CML, and the RTT was measured, comparing the latency with one firewall to the latency with a firewall cluster consisting of two firewalls. In this third *t*-test, the null hypothesis was rejected, as there was not a significant increase in latency. This was interesting to note because, as was mentioned during the literature review, it had been reported that firewall clustering did cause a significant increase in latency.

In the fourth *t*-test, the null hypothesis was also rejected, as there was not a significant difference in latency when one of the firewalls was shut down and the secondary node converged to act as the primary firewall. However, these results have to be classed as null and void because there was 55% packet loss after the firewall had converged. This is not ideal for patient monitoring systems, but testing this in CML showed that this latency was only temporary and packet loss returned to a normal rate after a period of time. As mentioned during the framework discussion, healthcare organisations should test this on their own networks to identify packet loss when one firewall is shut down. Even though the packet loss was high when one firewall was shut down, this is still better than only using one firewall which could experience problems, resulting in all packets being dropped.

6. Conclusions

This framework proposes a hybrid network architecture model using mostly zero-trust concepts while allowing healthcare organisations to continue using traditional implementations such as VPNs. This factor was important during the design of the framework, as security has to stop somewhere or else a point will be reached where security is so tight that the staff are unable to carry out their work. In terms of a zero-trust future for the healthcare industry, it is likely that medical devices being built today with high security standards will still be used when they are 10, 15, or 20 years out of date and new vulnerabilities have been discovered. It is for this reason that the framework should stay relevant to the healthcare industry for a long time. In 2016, the FDA issued a document laying out security recommendations for future medical devices. If medical device manufacturers listen to this guidance, the healthcare industry may end up being much safer in the future compared with present times [27]. Unfortunately, the FDA guidance does not focus enough on se-curing medical devices which are currently being used in healthcare organisations. The framework was designed to fill this gap in the guidance not just within the FDA document, but within the cyber security industry as a whole. The quantitative research results will also provide peace of mind for healthcare professionals, as the framework was designed to be secure but not so restrictive that patient care is interrupted. While the framework was tested as thoroughly as it could be in the simulation software, it is only a proof of concept, and results may vary in real-world implementations. For example, healthcare organisations may struggle to finance so many firewalls at once, and the implementation process may take a long time because they cannot risk interrupting patient care for too long. Having said this, it is better to make a start toward a zero-trust architecture than to not move forward at all.

References

1. Protenus. *Breach Barometer 2020*; Protenus: Baltimore, MD, USA, 2020.
2. Walker-Roberts, S.; Hammoudeh, M.; Dehghantanha, A. A Systematic Review of the Availability and Efficacy of Counter-measures to Internal Threats in Healthcare Critical Infrastructure. *IEEE Access* 2018, *6*, 25167–25177. [CrossRef]
3. Mansfield-Devine, S. Identity crisis: The disconnect between business and IT executives. *Comput. Fraud. Secur.* 2018, *2018*, 16–20. [CrossRef]
4. Uttecht, K.D. Zero Trust (ZT) Concepts for Federal Government Agencies. Lexington, Massachusetts. 2020. Available online: https://apps.dtic.mil/sti/pdfs/AD1106904.pdf (accessed on 12 December 2020).
5. Haber, M.J. *Privileged Attack Vectors*; Apress: Berkeley, CA, USA, 2020; pp. 295–304. [CrossRef]
6. Gilman, E.; Barth, D. *Zero Trust Networks*, 1st ed.; O'Reilly: Sebastopol, CA, USA, 2017.
7. Sommerlad, P. Reverse Proxy Patterns. In Proceedings of the EuroPLoP, Irsee, Germany, 25–29 June 2003; pp. 431–458.
8. Langer, J. What Zero Trust Should Look Like in Healthcare, Medigate. 2020. Available online: https://www.medigate.io/what- zero-trust-should-look-like-in-healthcare/ (accessed on 18 January 2021).
9. Cunningham, C.; Holmes, D.; Pollard, J. The Eight Business and Security Benefits of Zero Trust. 2016. Available online: https://www.akamai.com/us/en/multimedia/documents/white-paper/the-6-business-and-security-benefits-of-zero-trust- white-paper.pdf (accessed on 24 January 2021).
10. Ayuso, P.N.; Gasca, R.M.; Lefèvre, L. Demystifying Cluster-Based Fault-Tolerant Firewalls. *IEEE Internet Comput.* 2009, *13*, 31–38. [CrossRef]
11. Juniper. Configuring Chassis Clustering on an SRX Series Devices. 2020. Available online: https://www.juniper.net/ documentation/en_US/junos/topics/topic-map/security-chassis-cluster-verification.html (accessed on 18 January 2021).
12. Hamilton, R.; Gray, W.; Sibanda, C.; Kandasamy, S.; Kirner, R.; Tsokanos, A. Deep Packet Inspection in Firewall Clusters. In Proceedings of the 2020 28th Telecommunications Forum (TELFOR), Belgrade, Serbia, 24–25 November 2020; pp. 1–4.
13. Kwon, H.; Kim, Y.; Yoon, H.; Choi, D. Optimal cluster expansion-based intrusion tolerant system to prevent denial of service attacks. *Appl. Sci.* 2017, *7*, 1186. [CrossRef]
14. Flanigan, J. *Zero Trust Network Model*; Tufts University: Medford, MA, USA, 2018.
15. NHS. Bring Your Own Device (BYOD) Policy. 2020. Available online: https://www.nhsx.nhs.uk/key-tools-and-info/ procurement-frameworks/clinical-communications-procurement-framework/bring-your-own-device-policy/ (accessed on 18 January 2021).
16. Sobers, A. BYOD and the Mobile Enterprise—Organisational Challenges and Solutions to Adopt BYOD. *arXiv* 2015, arXiv:1512.03911. Available online: http://arxiv.org/abs/1512.03911 (accessed on 22 January 2021).
17. Marshall, S. IT Consumerization: A Case Study of BYOD in a Healthcare Setting. Technol. *Innov. Manag. Rev.* 2014, *4*, 14–18. [CrossRef]
18. Embrey, B. The top three factors driving zero trust adoption. *Comput. Fraud. Secur.* 2020, *2020*, 13–15. [CrossRef]
19. Martin, G.; Martin, P.; Hankin, C.; Darzi, A.; Kinross, J. Cybersecurity and Healthcare: How Safe are We? *BMJ* 2017, *358*, j3179. [CrossRef] [PubMed]

20. Li, S. Editorial: Zero Trust based Internet of Things. EAI Endorsed Trans. *Internet Things* 2020, *5*, 165168. [CrossRef]
21. Cisco. Cisco Medical Grade Network, Cisco. 2012. Available online: https://www.cisco.com/c/dam/global/en_ca/solutions/ strategy/healthcare/assets/docs/09CS2124-MGN.pdf (accessed on 24 February 2021).
22. Mehraj, S.; Banday, M.T. Establishing a zero trust strategy in cloud computing environment. In Proceedings of the 2020 In- ternational Conference on Computer Communication and Informatics, ICCCI 2020, Coimbatore, India, 22–24 January 2020; pp. 20–25. [CrossRef]
23. EICAR. Anti Malware Testfile. 2006. Available online: https://www.eicar.org/?page_id=3950 (accessed on 7 April 2021).
24. Frenz, C. Mock Malware Outbreak. 2020. Available online: https://dhinsights.org/wp-content/uploads/2020/07/Mock_ Malware1.1.pdf (accessed on 4 May 2021).
25. VMWare. Network & Micro-Segmentation Solutions. 2021. Available online: https://www.vmware.com/uk/solutions/micro- segmentation.html (accessed on 13 August 2021).
26. Gralla, P. Medical IoT Devices: The Security Nightmare That Keeps CIOs up Late at Night, Hewlett Packard Enterprise. 2017. Available online: https://www.hpe.com/us/en/insights/articles/medical-iot-devices-the-security-nightmare-that-keeps-cios- up-late-at-night-1709.html (accessed on 13 August 2021).
27. FDA. Postmarket Management of Cybersecurity in Medical Devices. Guidance for Industry and Food and Drug Administration Staff. Food and Drug Administration. 2016; pp. 1–30. Available online: https://www.fda.gov/downloads/medicaldevices/ deviceregulationandguidance/guidancedocuments/ucm482022.pdf (accessed on 16 May 2021).