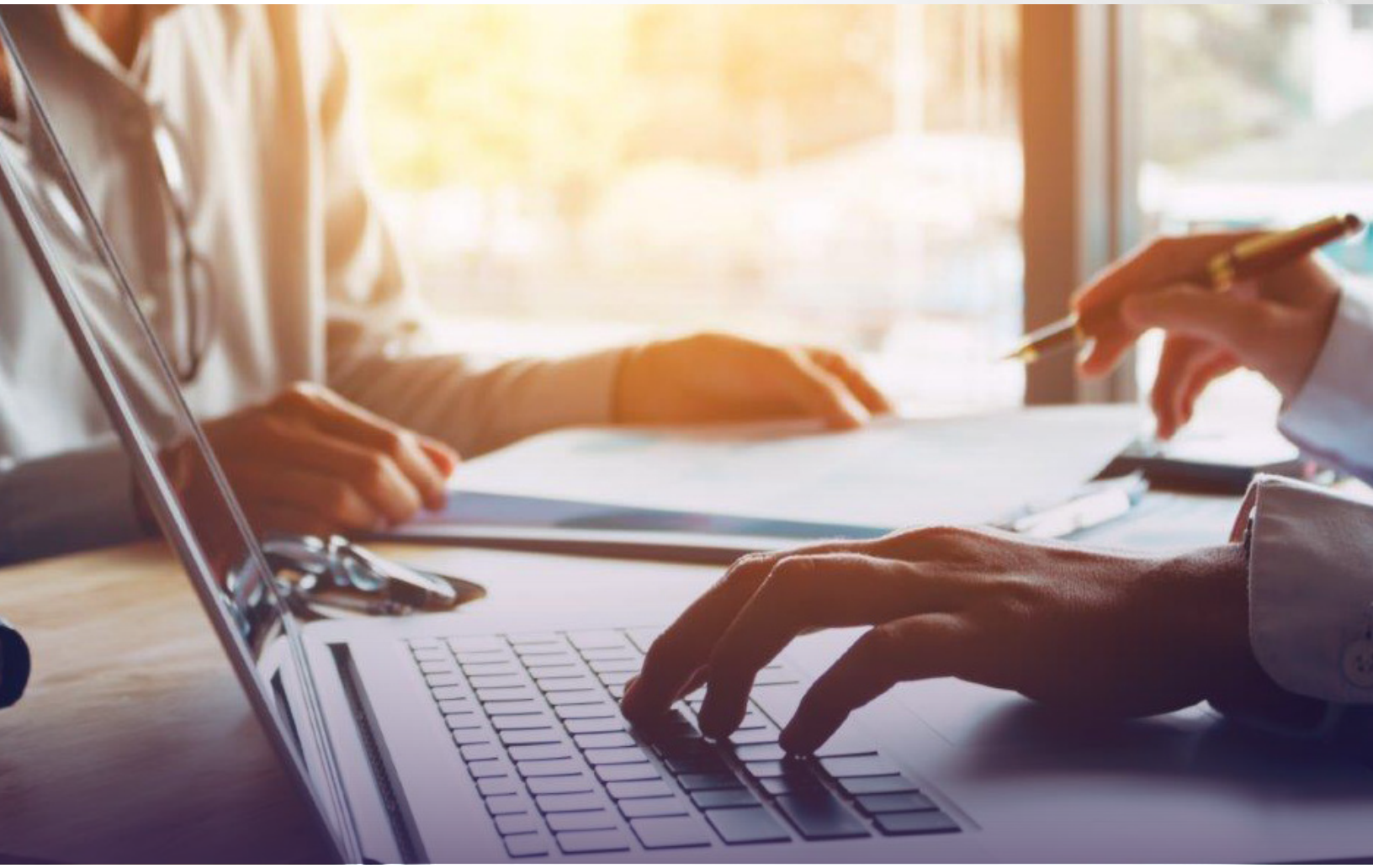
WHITE PAPER



Offline Certification Authority Best Practices

Candidate architectural design solutions and best practices for deploying offline certification authorities secured with Entrust nShield hardware security modules

ENTRUST

SECURING A WORLD IN MOTION

Contents

Introduction

Anchoring Trust With Root Certification Authorities

Root certification authorities (CAs) are deployed as trust anchors in public key infrastructure (PKI) solutions used to issue digital certificates in an organization. Root CAs in a simple PKI hierarchy certify sub-ordinate CAs (often referred to as Issuing CAs), which then issue digital certificates to "end entities" such as web servers, workstations, or individuals. These entities use digital certificates, and the public/private keys associated with them, to identify themselves on the network or perform other functions such as encrypt and/or sign information or messages.

The physical separation of trust anchor (root CA) and issuing CAs in the PKI hierarchy is designed to minimize the risk of compromise of the root CA and therefore of the entire PKI. As part of the risk mitigation of the root CA it is protected using both physical and logical controls.

The root CA is unlike any other information technology asset deployed within an enterprise; implementing a root CA demands critical security considerations first and foremost. To ensure suitable physical isolation, a root CA is ordinarily deployed offline (not connected to any networks). The private key material[1] that constitutes the primary security enforcing component of a root CA (it is used to sign certificates issued by the CA) must be protected within a tamper-proof environment for cryptographic isolation – generally established by the use of a hardware security module (HSM). Furthermore, the HSM should provide suitable controls to ensure correct authorization of access to the protected key material.

HSMs are cryptographic devices that are connected to a system to provide extremely strong protection of cryptographic key material and to overcome the inherent weaknesses of performing cryptographic operations in software. HSMs provide a physically tamper-proof security envelope within which key material can be stored and used securely subject to multiuser operational controls that enforce separation of duties and authorization policies. Despite a CA being deployed offline, it is still essential to employ HSMs to protect signing keys.

1. Compromise of a CA's private signing key would make it relatively easy for a hostile party to impersonate the CA and issue bogus certificates to untrusted entities.

It is essential to always be mindful that the PKI is often used to issue certificates where ultimate reliance is placed in the trustworthiness of the credential – whether that is to assert identity to high-value systems (authentication), encrypt security sensitive data (confidentiality), or impart confidence that security of sensitive data has not been modified after a signing operation has been performed on it (integrity). Loss of confidence in the integrity of a root CA could force some or, quite often, all issued credentials to be revoked and reissued – which implies a high level of business disruption for potentially a prolonged period of time, incurring significant direct and indirect costs.

It should be stated that nowhere is it written in PKI lore that root CAs must be offline[2] – it's a design approach influenced by the assurance required of the trust anchor, largely derived from the potential value of the operations outlined previously. Whether a root CA is implemented online or offline in no way structurally affects the logical PKI design – such as the chain of trust from a leaf certificate to a root CA. Storage of root CA keys in an appropriately rated (FIPS 140-3 Level 3) HSM adds a further level of physical protection to the logical protection of the root CA concept. Combined, these two tactics yield what has come to be regarded as an industry standard for commercial strength root CA deployment.

Being deployed "offline" ensures there's absolutely no opportunity for network-based attacks directly on the root CA. It's worth noting however that while root CAs are deployed offline, they periodically publish a CA certificate and certificate revocation list (CRL), which must be distributed to online repositories[4] and retrievable by relying parties.[5] Unavailability of this material can be a potential denial of service threat against parties relying upon certificates issued by the PKI.

This document focuses on describing the architectural choices and considerations in deploying a CA hierarchy rather than pure PKI security elements such as key lengths or CRL distribution points – the exception being the short discussion regarding the applicability of a "three-tier PKI" in the next section. The purpose of this document is to articulate trade-offs in cost, utility, and security involved in offline CA architectural design.

2. This paper only considers the offline CA scenario – with the CA's private key protected by an Entrust nShield HSM.

4. Such as HTTP servers or LDAP directories.

5. A relying party is an entity that must make a decision upon whether to trust a certificate they are presented.

The Case for Three-Tier PKI

Before getting into the nuts and bolts of architecting offline CAs, it's worthwhile considering whether there is justification for a two- or three-tier solution. As a basic premise, a PKI design should start with a two-tier solution – consisting of a single root CA and however many issuing CAs are deemed necessary.[6] Ability to cross-certify with other organizations and ability to enforce strict certificate policy constraints either intra-organization or via the cross-certification can all be achieved using a two-tier approach as illustrated in Figure 1.

There are situations where an intermediate tier of policy CAs is justifiable; for example you may have an extensive infrastructure of high-assurance CAs that which needs to be trusted via a cross-certificate[7] as shown in Figure 2; in this circumstance it may be sensible to cross-certify at a policy CA rather than have multiple cross-certification "instances."

Many people consider it a best practice to implement a three-tier PKI based and a cursory read of older Microsoft PKI documentation on the topic would lead one toward three-tier architecture. However, an excerpt on the following page from the latest Microsoft implementation planning and design guide for Active Directory Certificate Services[8], clearly illustrates that there are trade-offs and the designer should consider both approaches.

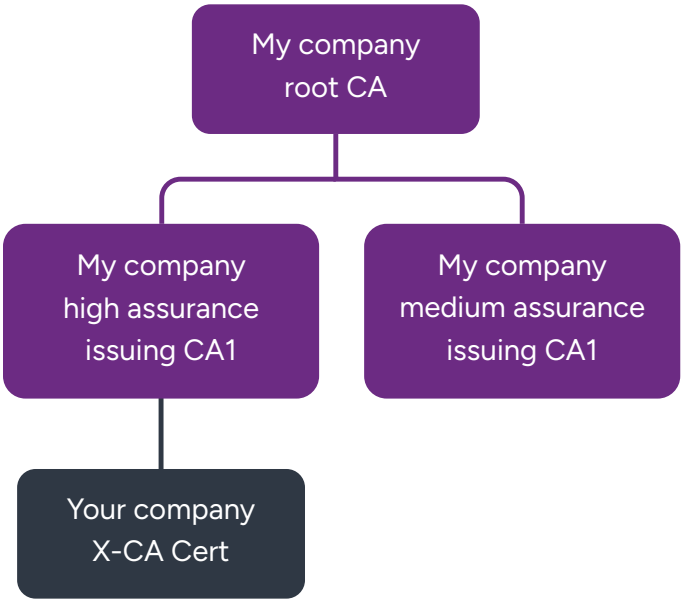


Figure 1: two-tier PKI

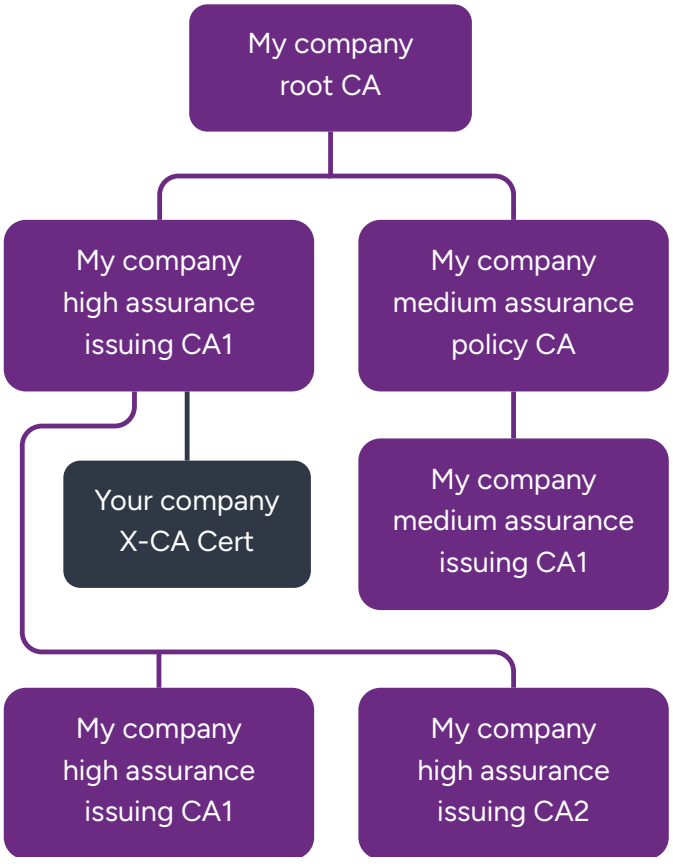


Figure 2: three-tier PKI

6. The number of issuing CAs is driven by location, volume, response time, and other considerations too numerous to detail in this document.

7. Another good use of cross-certification is to migrate from legacy PKI to a new PKI instantiation.

8. See https://techcommunity.microsoft.com/t5/ask-thedirectory- services-team/designing-and-implementing-apki- part-i-design-and-planning/ba-p/396953.

"Designing a three-tier hierarchy with intermediate CAs increases the complexity of the environment. Requirements to implement different policies can be implemented in a two-tier hierarchy with additional issuing CAs. The Windows Server product group states that there are no scale limitations that require a middle tier, so avoid using intermediate CAs unless there is a compelling business reason for doing so."

One final thought on the subject of tiers… if you are simply implementing a PKI for a short-term tactical objective or for a low-assurance situation such as a dedicated CA for issuing "health certificates" to support network access protection (NAP), etc. it may not even be necessary to go beyond a single tier. There's nothing mandating that you must have root CAs purely as trust anchors; in tactical situations it may be advantageous to issue end entity certificates from a "root issuing CA" – not a scenario which this paper considers.

Offline CA Solution Design Approaches

When considering the best approach for designing your offline CA solution, a number of basic elements should be considered:

- **Server hardware or portable hardware**
  There's certainly two schools of thought here… one aligned with deploying offline CA services on enterprise class server hardware and the other minded to deploy offline CA services on a portable, small form-factor platform such as a laptop.
- **Virtualization or "physical server" platforms**
  While virtualized server platforms (such as Microsoft Hyper-V or VMware solutions) have obvious benefits in the IT estate, we look at how applicable this approach is in the use case of offline CAs.
- **HSM form-factor**
  The Entrust nShield® HSM product line consists of three different form-factors[9]: USB attached (nShield Edge), PCI Express (PCIe) card (nShield 5s), and network-attached appliance (nShield 5c). Note that Entrust has also allowed deploying the nShield 5s, PCIe card in third-party PCIe extension cases, which conform to Thunderbolt and our hardware environmental requirements (e.g., cooling), supported.
- **Planned obsolescence**
  It should be borne in mind that during a root CA's lifetime (typically 20 years) the server hardware platform, HSM hardware module, operating system, and CA application software will all have become obsolete several times over.

9. The various features and specifications that differentiate members of the Entrust nShield family of HSMs are outlined in the appendix of this document.

Finally, and by no means least, consideration should be made to ensure the commissioning processes and activities of deploying the PKI in what is often referred to as a key signing ceremony[10] are fully documented. Indeed, there is almost a mythology around deploying offline CAs with the rigorous processes and strict accountability – unnecessarily complex solutions can have serious ramifications for ensuring the integrity of the initial instantiation as well as on-going maintenance and potential recovery scenarios. CA implementations that are rushed and / or poorly documented, no matter how much expense is lavished on technology, are likely to be brittle in times of emergency (e.g. system restore under pressure) or to demonstrate integrity to an auditor.

From a practical point of view the fact that offline CAs have no requirement to integrate with network services, in some ways, reduces the complexity of the system. It can be argued however, that the balance of complexity essentially transfers to suitable physical access controls to provide the requisite assurance.

material (such as with a CA) whereby extremely detailed installation documentation is required and every action isprecisely executed and observed So, although it can be recognized that offline CA design is perhaps fundamentally less technically complex than that of network connected CAs – it does not mean deploying them is necessarily straightforward. The bulk of this paper is intended to inform the reader of the type of decisions needing to be made when considering physical deployment of offline CAs and give insight into potential best practice approaches.

The remainder of this document considers the following three offline CA deployment options:

- Server with Entrust nShield 5s HSM
- or an nShield 5s deployed in a 3rd party Thunderbolt PCIe extension case.
- Virtualized server with Entrust nShield 5c HSM

10. A key signing ceremony is an approach used when commissioning systems utilizing sensitive cryptographic keys.

Server-Based Offline CAs

Overview

Traditionally, the platform of choice for deployment of an offline CA has been a dedicated server; with resilient hardware, enhanced supportability, and high-performance processing. While there has clearly been a significant shift in the market toward consolidating servers using virtualization technology, there are still many situations (and an offline CA is one) where dedicated "tin" is still a very legitimate approach to provide the requisite security, longevity, and manageability platform for an offline CA.

Architecture

The Entrust nShield 5s is an HSM with the form-factor of a PCIe[11] card and is generally the preferred member of the nShield family when an enterprise server chassis approach is taken to deploying offline CA capability. The use of a server chassis with easily pluggable disk drives makes viable the deployment of each offline CA (assuming there are policy CAs as well as a root CA) on independent disks, which would be bootable in their own right and never present in the server chassis simultaneously.

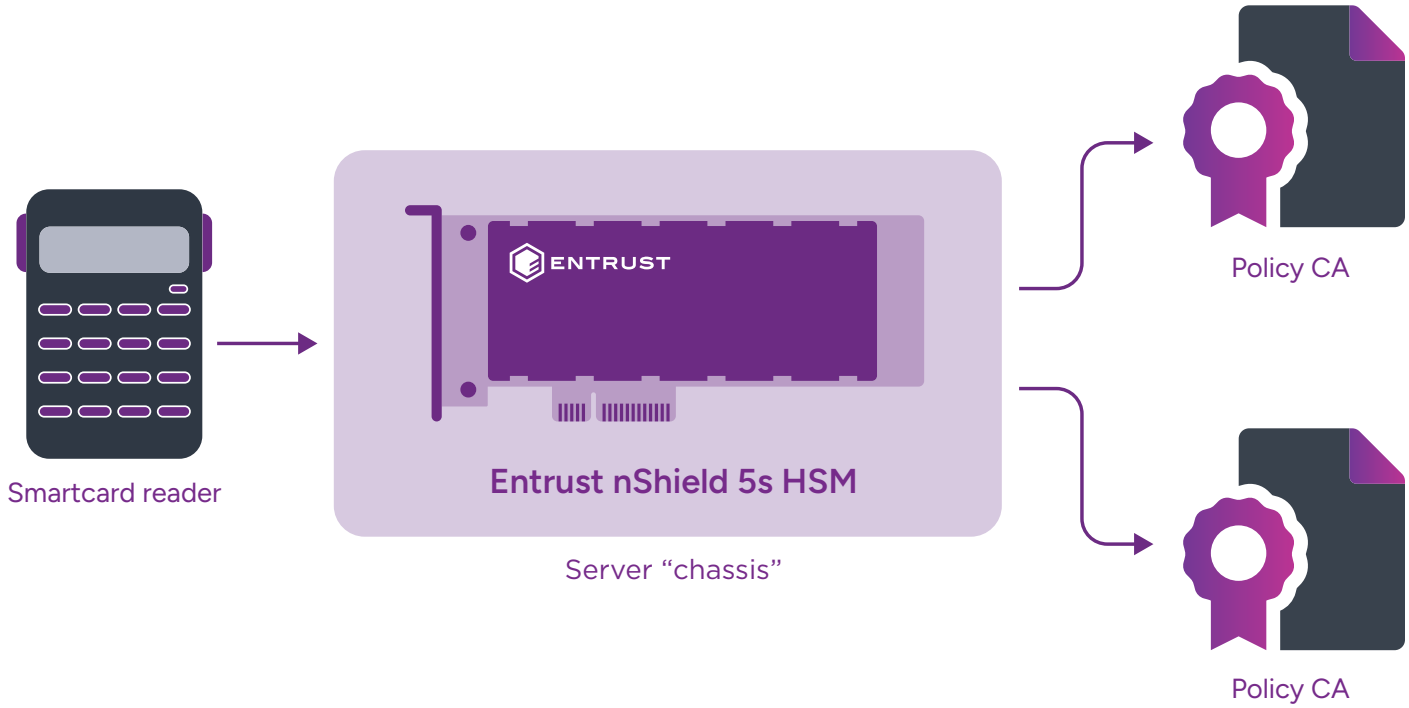


Figure 3: server chassis hosted offline CAs

11. The peripheral component interconnect express (PCIe) standard is a high-speed serial computer expansion connection used in motherboard interfaces.

Deploying offline CAs on a server chassis means that there may be physical access challenges to be overcome, which may necessitate establishing dedicated server racks, etc., with suitable controls to ensure that the high-assurance benefits of implementing the CAs offline are not negated.

The server chassis approach for implementing offline CAs is illustrated in figure 3; although a rack form-factor server chassis is illustrated here, there's no reason why a tower server could not be used.

Discussion

As far as simplicity of server build/demarcation between offline CAs are concerned, the server chassis approach has great merit and is entirely valid. For many years the server-centric approach combined with the Entrust nShield 5s HSM cards has been the mainstay of offline CA deployments leveraging Entrust nShield HSMs.

A further consideration when evaluating how to design longevity into your offline CA solution is the backup/recovery methodology employed. One positive school of thought is to make it relatively simple to recover/move/migrate your offline CA from one platform to another – termed a component-based recovery, it is described in the following section.

Component-Based Offline CA Recovery [12]

Introduction

It's likely that during its 20-year lifetime (typically), the root CA will go through the following events:

- Three or more hardware refreshes (especially so for laptops)
- A hardware failure
- Two or more operating system platform (e.g. Windows Server version) cycles[13]

For these reasons, it's imperative that a solution is in place for a component-by-component recovery/migration installation of a CA to complement any "full system" type restore methods employed.

While this paper is CA vendor agnostic, it is still instructive to illustrate what is meant here by a component-by-component-based recovery by using the example of an Active Directory Certificate Services (AD CS) implementation on Windows Server 2025 or other supported version. Other Certification Authority platforms would require a different backup regime than that described.

Material to back up

In the case of AD CS, the solution should be designed with a component-based recovery approach in mind and this is made easier by separating the components to be recovered onto a second logical volume, e.g. "D". This second drive doesn't need to be on a separate physical disk and the principal reason for its use is solely to make the backup/recovery process clearer. You'd store the following material on "D":

When the aforementioned components are established, backup essentially consists of copying a few megabytes of data in folders on your "D" drive to a CD/DVD.

Recovery is straightforward; the process outlined below can recover/redeploy/migrate an offline CA in a matter of minutes once the Windows server platform is re-established.

- Install Windows and configure the platform as before (hostname, logical drives, etc.)
- Copy the files from the component backup (CD/DVD) to "D"
- Install Entrust middleware, then connect to the HSM
- Re-associate the CA certificate with its HSM protected private key (CSP/KSP repair)
- Install AD CS, specifying the existing certificate and private key
- Replay the AD CS registry configuration,[15] then restore the CA database

While the sequence may seem more complex than simply "insert a DVD and press the restore button," there's total confidence in the outcome as the anxiety of waiting for a restored system to come back up without errors is avoided.

12. The recovery process described here is applicable to all offline CA deployment strategies described in this paper.

13. While it's true that upgrades could be done cumulatively (i.e. in place upgrades) – there would almost certainly come a time when it's desirable to have a clean build.

15. Best practice is to configure CA registry settings programmatically (e.g. with batch files) rather than manually editing the registry or using the certification authority management snap-in.

Laptop-Based Offline CAs

Overview

The offline root CA scenario has strong requirements for physical isolation of its computer hardware. Rather, discuss the ability to store a latop in a safe for additional protection of the root CA. Given the cost and logistical advantages of this relatively new model it is an approach worthy of serious consideration.

Architecture

While the network-attached, appliance-based Entrust nShield 5c could legitimately be used in combination with a laptop, the clearest and most practicable HSM form factor is either the USB-attached nShield Edge or the nShield 5s, deployed in a suitable Thunderbolt case. The nShield Edge provides all the capability required to support an offline root CA and shares the same level of FIPS certification and software platform as other members of the nShield HSM family. The primary concessions of the nShield Edge over its most immediate sibling, the nShield 5s, are its signing performance and lack of support for post-quantum algorithms. If either of these are a concern, the 5s should be the HSM form factor of choice. It is also worth noting that the nShield Edge ended its sales life in 2025 and can no longer be purchased. So, for a new root CA, the nShield 5s installed directly in a server, PC, or Thunderbolt case may be the only option for new deployments until the Edge replacement is available in 2027. Which in the context of a root CA is a negligible factor. The architectural simplicity of connecting a laptop to an nShield Edge or nShield 5s in a Thunderbolt case is illustrated in Figure 4. The next consideration is whether to implement the server operating system hosting the root CA natively on the laptop, e.g. install Windows Server on the laptop or to choose a virtualization path. Attention must be paid to whether the Windows Server OS is supported on the laptop hardware of choice. There are multiple laptop offerings from the major equipment vendors that are certified interoperable with MS Windows Server, any of which would be a suitable platform for an offline CA.

Another variation of the laptop architecture is for situations where you have multiple offline CAs, as in the circumstance of implementing policy CAs as well as a root CA. In this situation, you might choose to virtualize your CAs within the same physical system and deploy the nShield 5s in its “multi-CA” version. This version allows multiple, completely isolated CA tenants to be configured within the nShield 5s and “rotated” by a system administrator (who has no access to the CA keys) for convenient roots of trust for up to 5 CAs.

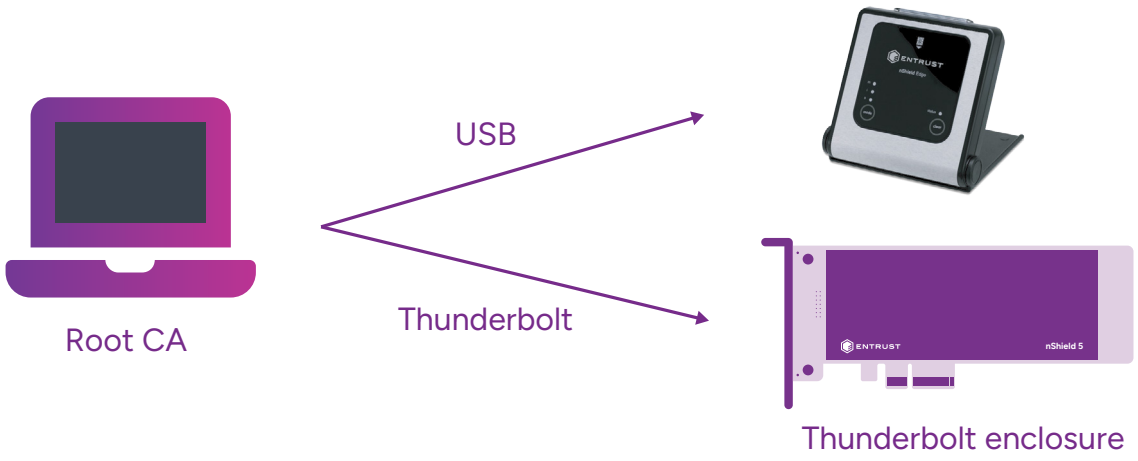


Figure 4: Laptop native root CA using either nShield Edge or nShield 5s HSM in Thunderbolt enclosure

Another approach while still considering the laptop architecture is to install Windows 11 Pro/Enterprise on the laptop, then create a Hyper-V role with a Windows Server VM running CA services.

Discussion

Implementing the component-based recovery process described earlier ensures that it's relatively straightforward to move the offline CA from one platform to another – whether the circumstance is demanded due to a hardware failure, hardware refresh, operating system upgrade, etc.

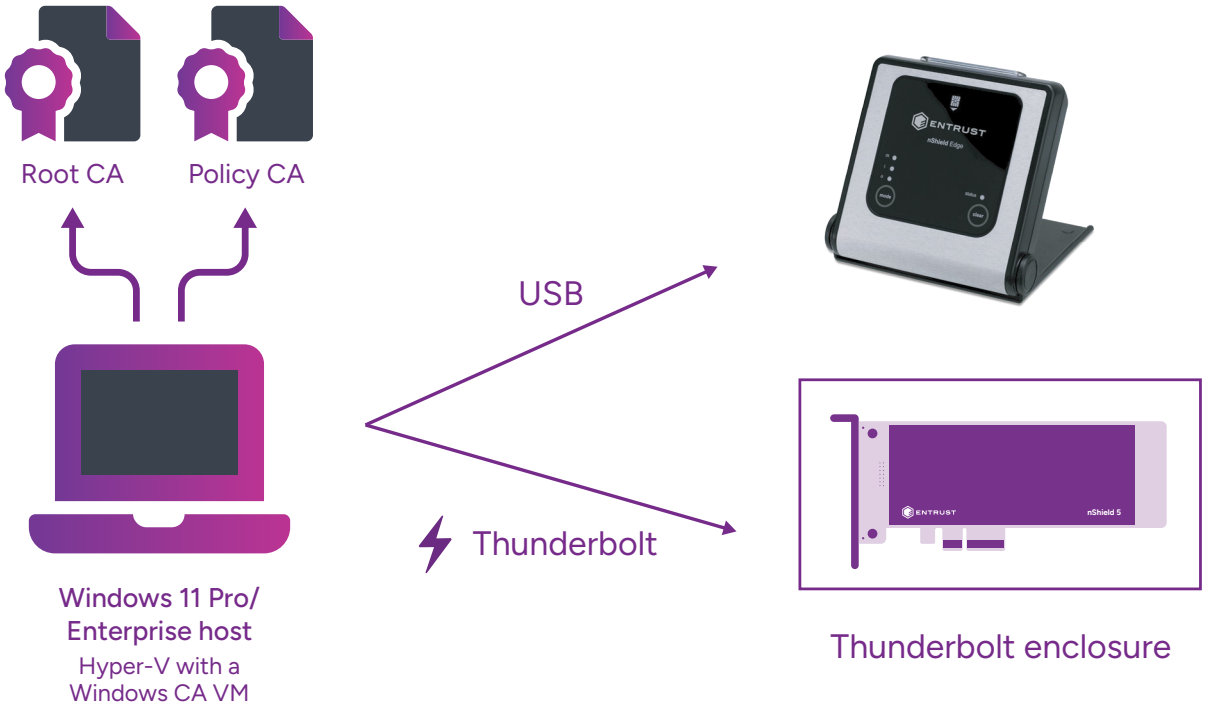


Figure 5: Virtualized root CA using either nShield Edge or nShield 5s HSM in Thunderbolt enclosure

Virtualization of Offline CAs

Overview

Virtualization has obviously made massive inroads into IT infrastructure and it's not the intent here to repeat all of the potential benefits that can be realized in a general context; however, there's purpose in examining the value proposition of virtualization for "server-based" offline CAs.

Architecture

Firstly, it's worth noting which HSM form factors are available in the context of server virtualization. For example, Thunderbolt drivers are not available directly in Windows Server operating environments.

So, the most appropriate HSM form factor when choosing to go down the virtualization path is the network-attached nShield HSM. Taking the basis of a single offline CA (root CA), Figure 6 illustrates the principal components required to support deploying the aforementioned CA in a VMware ESX solution. One thing is – this approach has a relatively high "platform overhead" in the context that a single root CA depends upon:

- A VMware ESX host
- A (Windows) management console for the ESX host
- A remote file system (RFS) guest operating system (OS) on the ESX host – this is a requirement when deploying an nShield Connect HSM
- The root CA guest OS itself

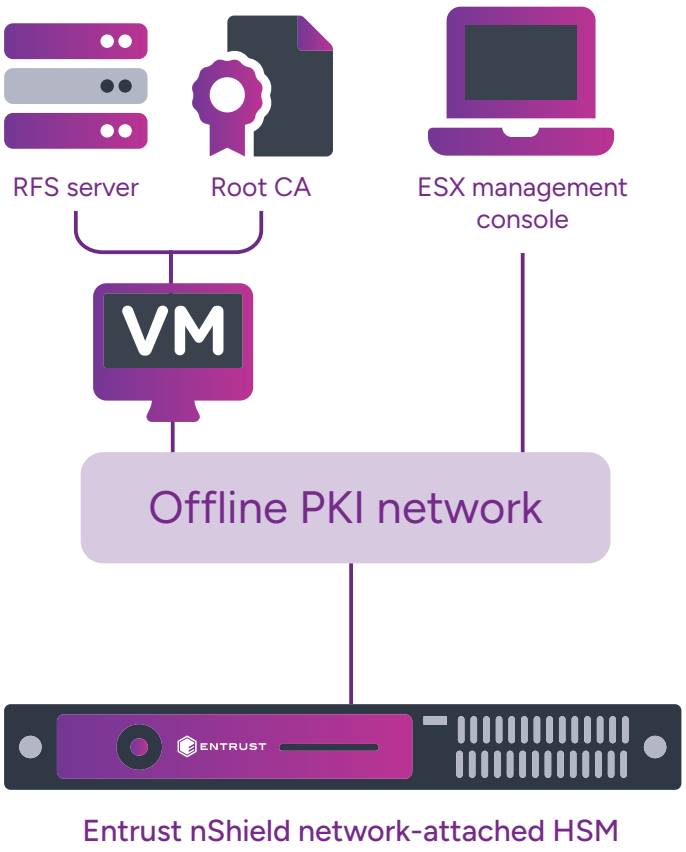


Figure 6: nShield network-attached HSM for offline CA virtualization

There may also be a network switch required to plumb the components together; generally speaking the cost and complexity of this type of solution to satisfy instantiation of a single root CA might be deemed unjustifiable.

To realize greater value from the network-attached nShield HSM deployment, customers often state a desire to leverage the network-attached nShield HSM by utilizing the same unit for their online CA infrastructure as well, an example of which is illustrated in Figure 7.

There are a number of ways of implementing the leveraged network-attached HSM; it is possible to take advantage of the fact that it has a second (nonroutable) network interface that can be connected to the offline network? while the first interface is connected to the online network (as in Figure 7). Another option would be to deploy the network-attached nShield HSM in a DMZ off a firewall that straddles the two "networks" and configure appropriate rules to isolate the offline network from the online network.

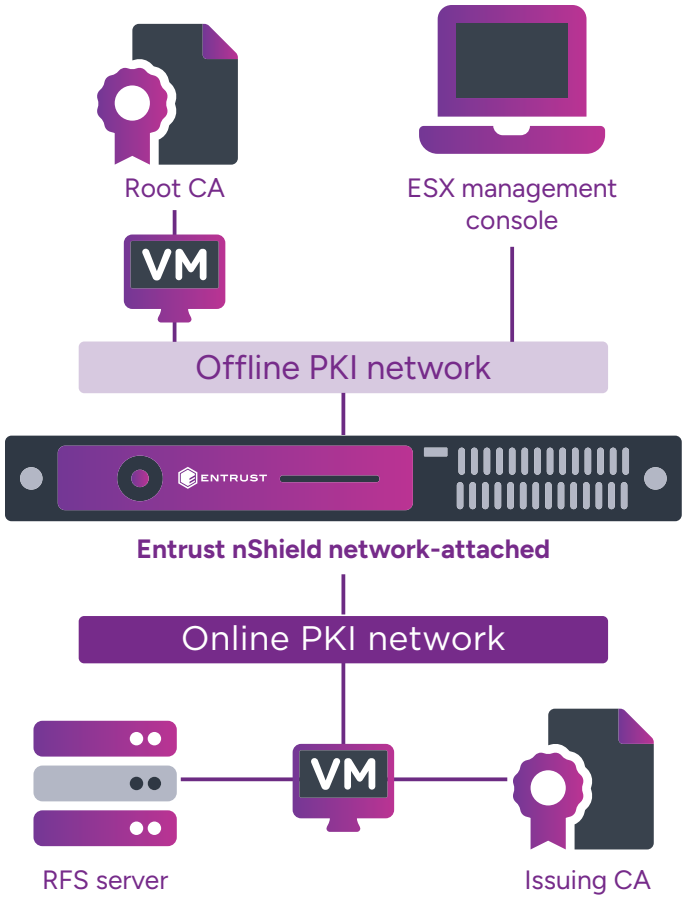


Figure 7: leveraging network-attached nShield HSM for offline and online CA virtualization

So, could deploying a suitably rated and configured firewall between the Root CA and extended network suffice? Typically no – PKI is governed by certificate policy (CP), which for anything higher than a low-assurance scenario will ordinarily demand that proper separation (airgapping) is rigidly enforced. There are no hard and fast rules around this and it may be that for a PKI being deployed solely for "internal use" that a firewall break is sufficient; however, if future requirements demand the PKI extends via cross-certification to a wider base, partners may not be satisfied with the firewall approach and their certificate policy may disallow the aforementioned cross-certification.

Discussion

It's important to ensure evaluation of the relative merits of architectural solutions are not simply validated by their technical viability, but more so by their applicability in the context of an extremely demanding operational and management PKI regime. The solutions described in this section are often passionately defended by "techies" and they certainly have much merit when building demonstrators or development rigs to flesh out the bones of a solution. However, put yourself in the position of developing a key signing ceremony around the aforementioned approaches and it puts a whole different slant on proceedings.

Let's look at the primary perceived advantages that are mooted for the virtualization approach for server-based offline CAs; these are more conservative than those benefits that would be identified for online CAs where virtualization provides huge opportunities for extracting value.

- **Leveraging network-attached nShield HSM**
  Can the root CA truly be described as offline if there is clear evidence of connectivity to wider networks as shown in Figure 7? Furthermore, it is harder to apply additional levels of physical security in this scenario, leading to an increased exposure to tamper attempts. To mitigate this, the virtualization platform hosting the root CA could legitimately be powered off and removed to a secure location and only re-introduced at such time that it needs to be operated.
- **Hardware abstraction**
  Having the root CA virtualized makes moving it from one server to another a relatively straightforward task. But, it's also likely that during its typically 20-year lifetime, the root CA will go through two or three operating system platform cycles and also the virtualization platform will go through a similar sequence. A component-based recovery solution (described earlier) offers greater flexibility to mitigate the potential for hardware/software failures/planned upgrades at the root CA.

Conclusions

Best practices

Hopefully this paper has given food for thought regarding different options available for deploying offline CAs in combination with Entrust nShield HSMs and perhaps dispelled a few myths. Clearly, this paper can't address every possible approach that can be taken, but it does provide best practice guidance and candidate solutions that will help you deploy offline CAs with Entrust nShield HSMs for you and your customers.

All three presented options are legitimate approaches and given suitable parameters can be justified; some of the influencing parameters that may drive you to take a certain approach are included here:

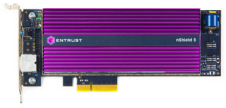
- **Small form-factor host:** If the over-riding requirement is to use a laptop for hosting the offline CA, then the nShield Edge HSM is generally the best choice.
- **Virtualization technology:** In this situation you would ordinarily use an nShield Edge HSM if you have a virtualization platform that supports attaching USB devices to guests (typically the laptop virtualization scenario); otherwise a network-attached nShield HSM is the realistic choice.
- **Performance:** Ordinarily performance comes far down the list of requirements for an offline CA; however, if this was ever the case then the nShield PCIe HSM or network-attached nShield HSM provide a high cryptographic performance platform.
- **Isolation:** In situations where "embedding" of the HSM into the server unit is required, the nShield PCIe form factor HSM provides unambiguous coupling between the HSM and host.
- **Leverage network-attached nShield Unit(s):** Customers are generally keen to get maximum utilization out of their HSM investment; in this circumstance it may be advantageous to connect an offline CA to the non-routable network interface of the nShield HSM.

A couple of further important points to bear in mind when making your design choices:

- Consider the applicability of your solution in the context of a key signing ceremony/ongoing rigorous management regime. Server virtualization may appear to be a technically adept solution, but the additional complexity introduced may negate this benefit – this is always difficult to portray to anyone who has never been responsible for authoring/directing a key ceremony or managing a tightly controlled and operated PKI estate.
- Use a neutral backup format where possible, such as DVD (which must be stored securely); in combination with a component-based recovery process this affords significant flexibility for maintaining the offline CA over the course of its lifetime.

Regardless of the offline CA platform architecture – it's clear that the Entrust nShield family of HSMs provides an extensive portfolio that can fit your solution.

Appendix – Entrust nShield HSM product matrix



Feature	nShield Edge	nShield 5s	nShield 5c
Interface	USB	PCIe	Network-attached form factor
Shareable	No	No	Yes
Power Supplies	N/a	N/a	2
Speed variants	10	nShield 5s Base, Mid, High	nShield 5s Base, Mid, High
Certifications	FIPS 140-2 L2 &L3	FIPS 140-3 Level 3 CC EAL4+	FIPS 140-3 Level 3 CC EAL4+

ABOUT ENTRUST

Entrust fights fraud and cyber threats with identity-centric security that protects people, devices, and data. Our comprehensive solutions help organizations secure every step of the identity lifecycle, from verifying identity at onboarding to securing connections and fighting fraud in everyday transactions. Ongoing monitoring supports compliance and safeguards keys, secrets, and certificates. With a foundation of identity-centric security, our customers can transact and grow with confidence. Entrust has a global partner network and supports customers in over 150 countries.

For more information, visit entrust.com.

entrust.com Toll-Free: 888.690.2424 | International: +1.952.933.1223 | sales@entrust.com

ENTRUST
SECURING A WORLD IN MOTION