IT and DevOps

- Using Smart Cards for Remote Git Instances
- OpenZFS Issue #15526 Patch and Mitigation for Older Versions
- Docker and Docker Compose v2 in Fedora CoreOS
- Importing VMs from TrueNAS Core (Bhyve) to Proxmox
- Image Credit - Book Cover Art

Using Smart Cards for Remote Git Instances

Summary

This is a coauthored article by Ryukodev and myself.

While this feature is undocumented, it is possible to use a smart card to authenticate to a Git repository behind Mutual Transport Layer Security (mTLS). This product describes how to do this given a Gitea instance, a Yubikey PIV smart card and a GNU+Linux information system. The solution is to simply add the following to a user's `~/.gitconfig` file:

```
[http "https://git.example.com"]
	sslCert =
"pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someUserna
me"
	sslKey =
"pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someUserna
me"
	sslbackend = openssl
	sslkeytype = ENG
	sslcerttype = ENG
	sslCertPasswordProtected = true
```

Additionally, an authentication token must be generated for the user within Gitea (or equivalent). This is how a git pull looks like with this configuration:

```
$ git pull origin master
Password for
'cert:///pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someU
sername':
Username for 'https://git.example.com': someUsername
Password for 'https://someUsername@git.example.com':
From https://git.example.com/someUsername/someRepository
```

```
 * branch            master     -> FETCH_HEAD
Already up to date.
```

Details

Background

Mutual TLS

In TLS, the client authenticates the server by verifying that the certificate has been cryptographically signed by a trusted certificate authority. Most commonly, this is seen when a user browses to a website. A lock icon appears next to their address bar if the server holds a proper certificate and associated private key; otherwise, an error occurs, and with the advent of HSTS, a user may be unable to browse to the website at all.

In mTLS, the server also requests the client to provide a client certificate, proving that it is signed by a certificate authority the server trusts. This is done at the protocol level, disavowing unauthorized users to even see the web app. It bears repeating: even if the server is running vCenter Server 7.0, vulnerable to CVE-2021-21985, an unauthenticated remote code execution (RCE) exploit, the attacker cannot exploit the server unless they have a private key and underlying certificate authorizing them to access vCenter.

Smart Cards

It is possible to simply hold a private key and certificate in an operating system's certificate store, allowing a signed-in user to use the certificate via a simple prompt. It's also possible to require the user to enter the encryption passphrase for the key, though in most environments, the user has the ability to turn this requirement off. Overall, one major issue remains: a password requirement for a private key used for mTLS is a form of single-factor authentication.

Enter NIST's FIPS 201-3, *Personal Identify Verification (PIV) of Federal Employees and Contractors*. Initially published in 2005, the standard defines a two-factor smart card, allowing an information system to request the user to enter an (up to) eight digit PIN before allowing the use of the private key and certificate. This effective makes PIV devices two-factor authentication security tokens. Yubikeys officially support PIV, and this will be the primary focus of this product.

System Configuration

Prerequisites

Ensure that the `gnutls` and `openssl-pkcs11` packages are installed on your system. The former is used for `p11tool`, and the latter is used for `cURL` (the HTTP backend for `git`). For RHEL-based distros, use the following command to install:

```
sudo dnf install gnutls openssl-pkcs11 -y
```

Creating .gitconfig

The `.gitconfig` requires the PKCS11 URI of the certificate and key. To get the URI, use `p11tool`, like so:

```
$ p11tool --list-all-certs
warning: no token URL was provided for this operation; the available tokens are:

pkcs11:model=p11-kit-trust;manufacturer=PKCS%2311%20Kit;serial=1;token=System%20Trust
pkcs11:model=p11-kit-trust;manufacturer=PKCS%2311%20Kit;serial=1;token=Default%20Trust
pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someUsernam
e
```

In this case, the bottom-most line is our PIV device. This can be easily identified by the token value, which for Yubikeys is the subject name.

Create a new file, `.gitconfig`, in your home directory, containing the following data:

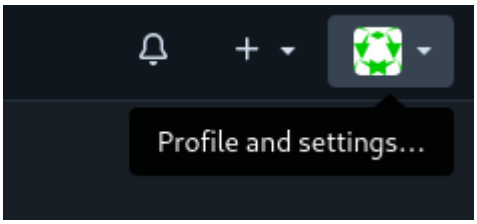
```
[http "https://git.example.com"]
	sslCert =
"pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someUserna
me"
	sslKey =
"pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someUserna
me"
	sslbackend = openssl
	sslkeytype = ENG
	sslcerttype = ENG
	sslCertPasswordProtected = true
```

The first line contains the URL for your Gitea (or equivalent) instance. Replace the values for sslCert and sslKey with your URI collected via the `p11tool` command above.

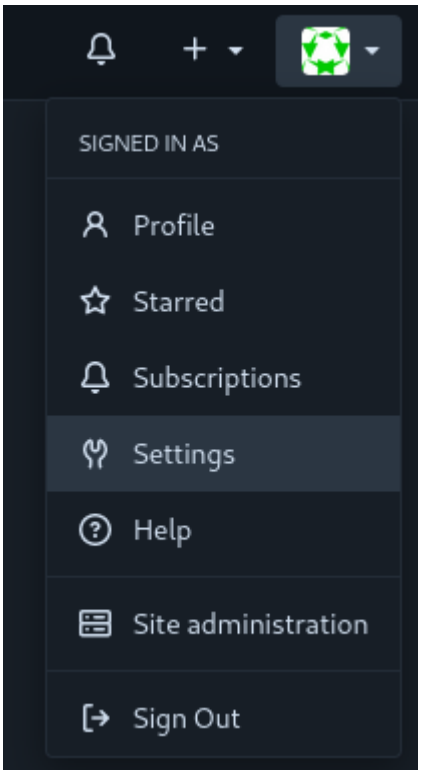
Gitea Configuration

These steps are specific to Gitea, but similar features are available for popular Git-based source code repository services like Bitbucket, GitLab, or GitHub Enterprise.

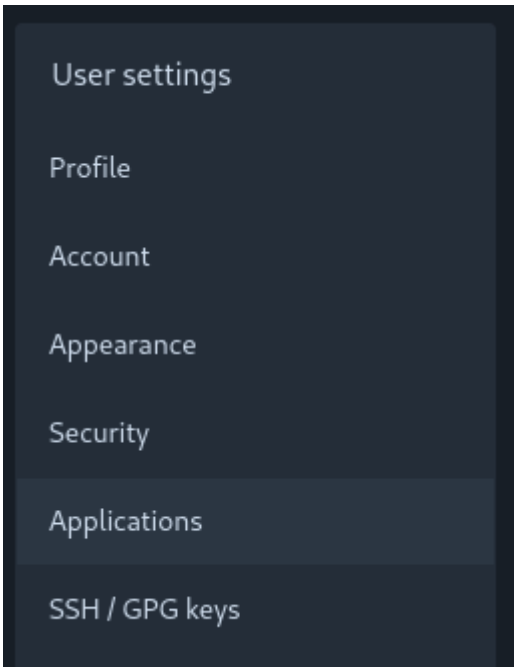
First, login and then press on your profile picture in the top right:



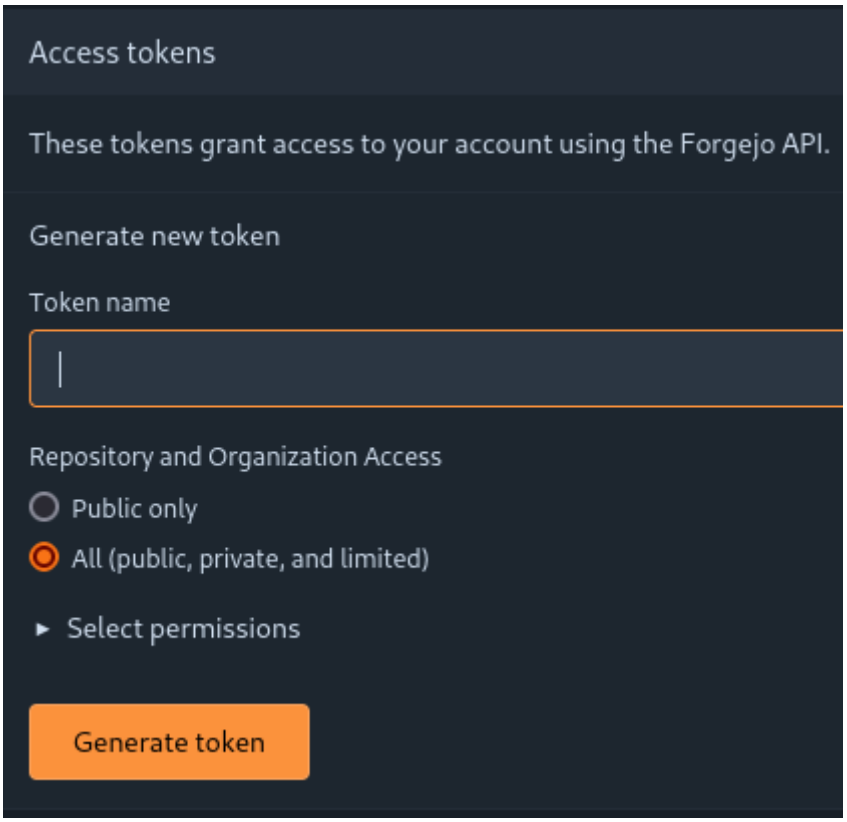
Select "Settings":



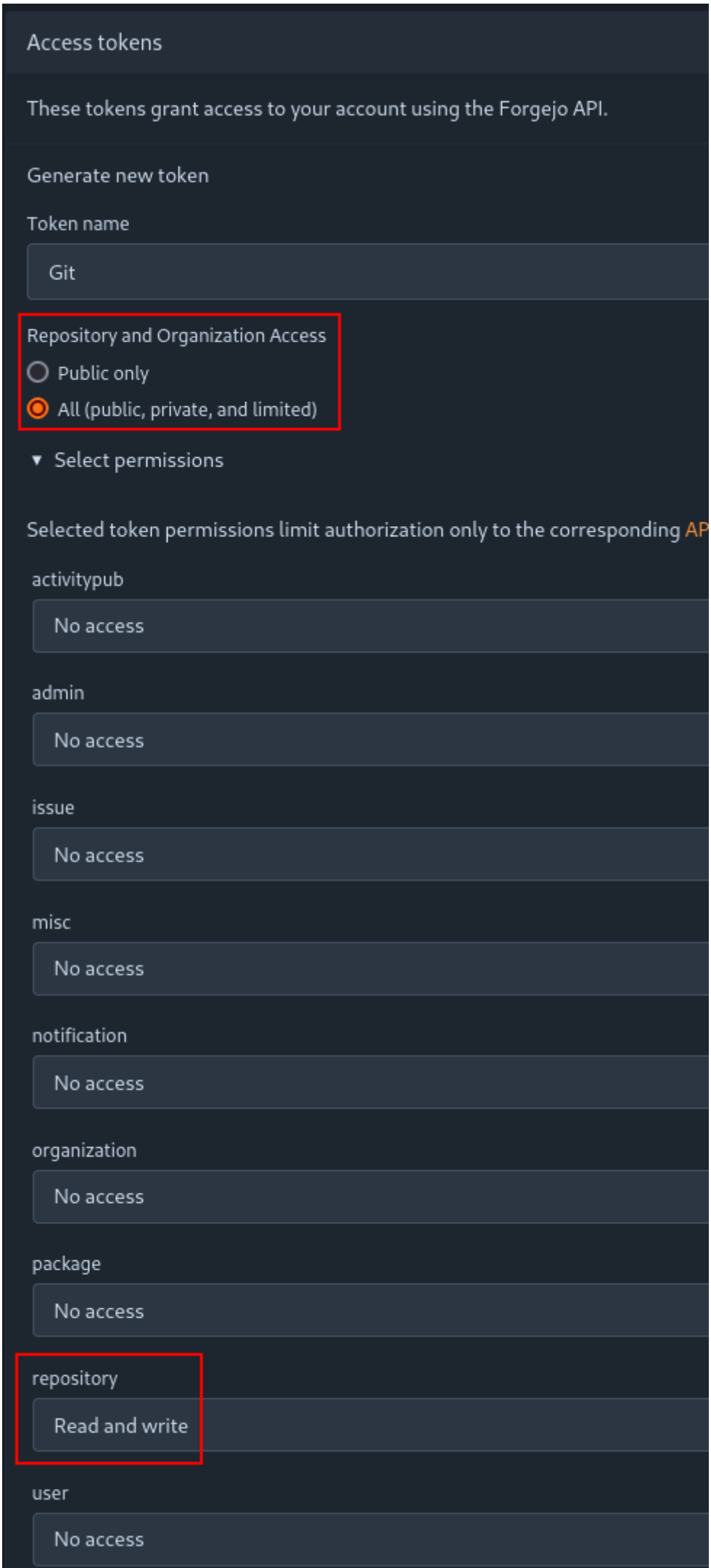
Then, select "Applications":



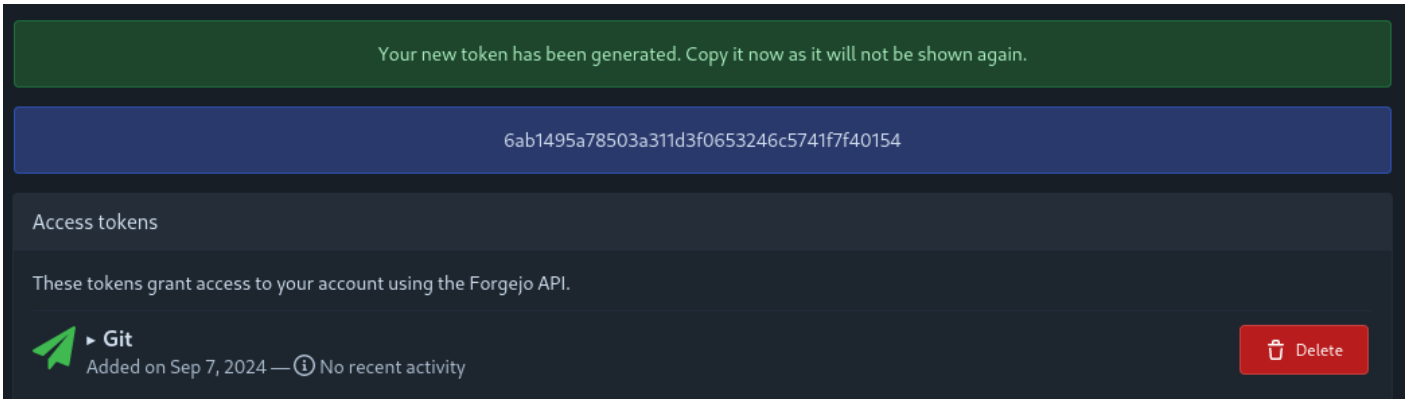
You will be brought to the following screen. Here, you can define a token with granular permissions. This token will be used after the mTLS authentication to the web server. The menu initially looks like so:



In our case, we set the token to have access to all public, private and limited repositories and organizations; and we allow read and write access to repositories:


Access tokens
These tokens grant access to your account using the Forgejo API.
Generate new token
Token name
Git
Repository and Organization Access
Public only
All (public, private, and limited)
Select permissions
Selected token permissions limit authorization only to the corresponding AP
activitypub
No access
admin
No access
issue
No access
misc
No access
notification
No access
organization
No access
package
No access
repository
Read and write
user
No access

Then, press "Generate token". You will be brought to the following screen, where the token secret is display:



> You must save this secret and use it every time you wish to do a Git operation on the remote instance

This is how a git pull looks like with this configuration:

```
$ git pull origin master
Password for
'cert:///pkcs11:model=PKCS%2315%20emulated;manufacturer=piv_II;serial=0011223344556677;token=someU
sername':
Username for 'https://git.example.com': someUsername
Password for 'https://someUsername@git.example.com':
From https://git.example.com/someUsername/someRepository
 * branch            master     -> FETCH_HEAD
Already up to date.
```

Git will first prompt you for your PIV PIN. Then, it will prompt for the username and the token password.

OpenZFS Issue #15526 Patch and Mitigation for Older Versions

> SUBSTANTIAL REVISION: Issue #15526 has been patched in ZFS versions 2.2.2 and 2.1.14, according to open source reporting. The mitigation below only applies to older versions of ZFS.

Summary

A silent data corruption bug exists in ZFS versions 2.1.4 up to 2.1.13, 2.2.0 and 2.2.1, per this GitHub issue. Versions 2.2.2 and 2.1.14 have patched the issue, according to open source reporting .

> A ZFS scrub will not identify any data corrupted by this bug. The only high-assurance method to check if any files were corrupted is to compare files within ZFS to their copies stored outside of ZFS.

If you are still using an older version of ZFS that is impacted by this issue, you may significantly lower the chance that this issue impacts your file system by setting the ZFS parameter zfs_dmu_offset_next_sync to 0. Note that this does not prevent the issue from occurring, per this GitHub comment, but it does lower the chance that silent data corruption occurs.

Details

Mitigation

Linux

Runtime

To apply the mitigation in runtime, run the following command as the root user:

```
echo 0 > /sys/module/zfs/parameters/zfs_dmu_offset_next_sync
```

Permanent

To apply the mitigation permanently, create a file in `/etc/modprobe.d/` such as:

```
/etc/modprobe.d/mitigation.conf
```

Containing the following:

```
options zfs zfs_dmu_offset_next_sync=0
```

FreeBSD

Runtime

To apply the mitigation in runtime, run the following command as the root user:

```
sysctl -w vfs.zfs.dmu_offset_next_sync=0
```

Permanent

To apply the mitigation permanently, append the following line to `/etc/sysctl.conf`:

```
vfs.zfs.dmu_offset_next_sync=0
```

Reproducing the Bug

Linux

To reproduce the bug in Linux, use the script below (copy pasted from the following gist):

```
#!/bin/bash
#
# Run this script multiple times in parallel inside your pool's mount
```

```
# to reproduce https://github.com/openzfs/zfs/issues/15526.  Like:
#
# ./reproducer.sh & ./reproducer.sh & ./reproducer.sh & ./reproducer.sh & wait
#

#if [ $(cat /sys/module/zfs/parameters/zfs_bclone_enabled) != "1" ] ; then
#	echo "please set /sys/module/zfs/parameters/zfs_bclone_enabled = 1"
#	exit
#fi

prefix="reproducer_${BASHPID}_"
dd if=/dev/urandom of=${prefix}0 bs=1M count=1 status=none

echo "writing files"
end=1000
h=0
for i in `seq 1 2 $end` ; do
	let "j=$i+1"
	cp  ${prefix}$h ${prefix}$i
	cp --reflink=never ${prefix}$i ${prefix}$j
	let "h++"
done

echo "checking files"
for i in `seq 1 $end` ; do
	diff ${prefix}0 ${prefix}$i
done
```

FreeBSD

To reproduce the bug in FreeBSD, use the script below (copy pasted from the following post):

```
#!/bin/bash
#
# Run this script multiple times in parallel inside your pool's mount
# to reproduce https://github.com/openzfs/zfs/issues/15526.  Like:
#
# ./reproducer.sh & ./reproducer.sh & ./reproducer.sh & ./reproducer.sh & wait
#
```

```
#if [ $(cat /sys/module/zfs/parameters/zfs_bclone_enabled) != "1" ] ; then
#       echo "please set /sys/module/zfs/parameters/zfs_bclone_enabled = 1"
#       exit
#fi

prefix="reproducer_${BASHPID}_"
dd if=/dev/urandom of=${prefix}0 bs=1M count=1 status=none

echo "writing files"
end=1000
h=0
for i in `seq 1 2 $end` ; do
        let "j=$i+1"
        cp  ${prefix}$h ${prefix}$i
        cp  ${prefix}$i ${prefix}$j
        let "h++"
done

echo "checking files"
for i in `seq 1 $end` ; do
        diff ${prefix}0 ${prefix}$i
done
```

Commentary

I was unable to reproduce this issue in TrueNAS Core 13.0-U5.3 (FreeBSD) but I was able to reproduce it in Proxmox 8.0.4 (Debian).

Source Description Block

> Multiple sources:
>
> Issue tracking in OpenZFS: https://github.com/openzfs/zfs/issues/15526
>
> Mitigation: https://github.com/openzfs/zfs/issues/15526#issuecomment-1823737998
> Linux reproducer script:
>
> https://gist.github.com/tonyhutter/d69f305508ae3b7ff6e9263b22031a84
>
> FreeBSD reproducer script: https://www.truenas.com/community/threads/truenas-13-0-u6-is-now-available.114337/page-3

> TrueNAS Core (FreeBSD) issue forum thread: https://www.truenas.com/community/threads/silent-corruption-with-openzfs-ongoing-discussion-and-testing.114390/
>
> Documentation on dmu_offset_next_sync: https://openzfs.github.io/openzfs-docs/Performance%20and%20Tuning/Module%20Parameters.html#zfs-dmu-offset-next-sync
> Data corruption bug occurs even with zfs_dmu_offset_next_sync set to 0: https://github.com/openzfs/zfs/issues/15526#issuecomment-1826348986
> Reddit thread on the bug: https://old.reddit.com/r/DataHoarder/comments/1821mpr/heads_up_for_a_data_corruption_bug_in_zfs_few/
> Reddit thread on the bug: https://old.reddit.com/r/zfs/comments/1826lgs/psa_its_not_block_cloning_its_a_data_corruption/
>
> Issue fixed in versions 2.2.2 and 2.1.14: https://www.phoronix.com/news/OpenZFS-2.2.2-Released

Licensing

Docker and Docker Compose v2 in Fedora CoreOS

Summary

If you prefer to use Docker over Podman in Fedora CoreOS, use the Butane file below to add the latest version of Docker and Docker Compose v2 to your system.

Details

Butane

```
variant: fcos
version: 1.4.0
passwd:
  users:
    - name: core
      ssh_authorized_keys:
        - ssh-[Your SSH key]
storage:
  files:
    - path: /etc/yum.repos.d/docker-ce.repo
      overwrite: true
      contents:
        inline: |
          [docker-ce-stable]
          name=Docker CE Stable - $basearch
          baseurl=https://download.docker.com/linux/fedora/$releasever/$basearch/stable
          enabled=1
          gpgcheck=1
```

```
      gpgkey=https://download.docker.com/linux/fedora/gpg
systemd:
  units:
    # Removing unofficial copies of docker and related packages
    - name: rpm-ostree-uninstall.service
      enabled: true
      contents: |
        [Unit]
        Description=Docker rpm-ostree install
        Wants=network-online.target
        After=network-online.target
        # We run before `zincati.service` to avoid conflicting rpm-ostree
        # transactions.
        Before=zincati.service
        ConditionPathExists=!/var/lib/%N.stamp

        [Service]
        Type=oneshot
        RemainAfterExit=yes
        ExecStart=/usr/bin/rpm-ostree override remove docker containerd runc
        ExecStart=/bin/touch /var/lib/%N.stamp

        [Install]
        WantedBy=multi-user.target
    # Installing Docker as a layered package with rpm-ostree
    - name: rpm-ostree-install.service
      enabled: true
      contents: |
        [Unit]
        Description=Docker rpm-ostree install
        Wants=network-online.target
        Requires=rpm-ostree-uninstall.service
        After=rpm-ostree-uninstall.service
        # We run before `zincati.service` to avoid conflicting rpm-ostree
        # transactions.
        Before=zincati.service
        ConditionPathExists=!/var/lib/%N.stamp

        [Service]
```

```
        Type=oneshot
        RemainAfterExit=yes
        ExecStart=/usr/bin/rpm-ostree install docker-ce docker-ce-cli containerd.io docker-buildx-plugin docker-
compose-plugin
        ExecStart=/bin/touch /var/lib/%N.stamp

        [Install]
        WantedBy=multi-user.target
```

Butane - Explanation

On line 7, add your SSH public key to be able to sign into your Fedore CoreOS machine. We add the Docker repository as a file. Then, we use some systemd trickery to remove docker, runc and containerd. These are installed by default in Fedora CoreOS, but conflict with the up-to-date versions of Docker, so we remove them. The next service waits for the uninstall service to complete, and installs docker per the Fedora installation guide here.

Your Fedora CoreOS system will reboot in 10 minutes after running these systemd services. It's unfortunately impossible to apply software removals live, so a restart is required. If you wish to restart sooner, you can run `systemctl reboot` manually.

Why?

Podman doesn't have the equivalent of Docker Compose. Per the suggestion of the Podman development team, we can simply use Docker Compose with a Podman backend. There needs to be some trickery done to support building images with a Podman backend, which can be seen here.

Overall, I found Podman to be more trouble than it's worth. As I worked with Podman for nearly a year, I ran into constant incompatibilities and oddities that had me searching for workarounds for things that should *just work.* Simply running the latest version of Docker and Docker Compose not only needs my needs, but is stable—I have yet to have any breaking changes due to automatic updates with Docker and Docker Compose v2.

Licensing

Importing VMs from TrueNAS Core (Bhyve) to Proxmox

Summary

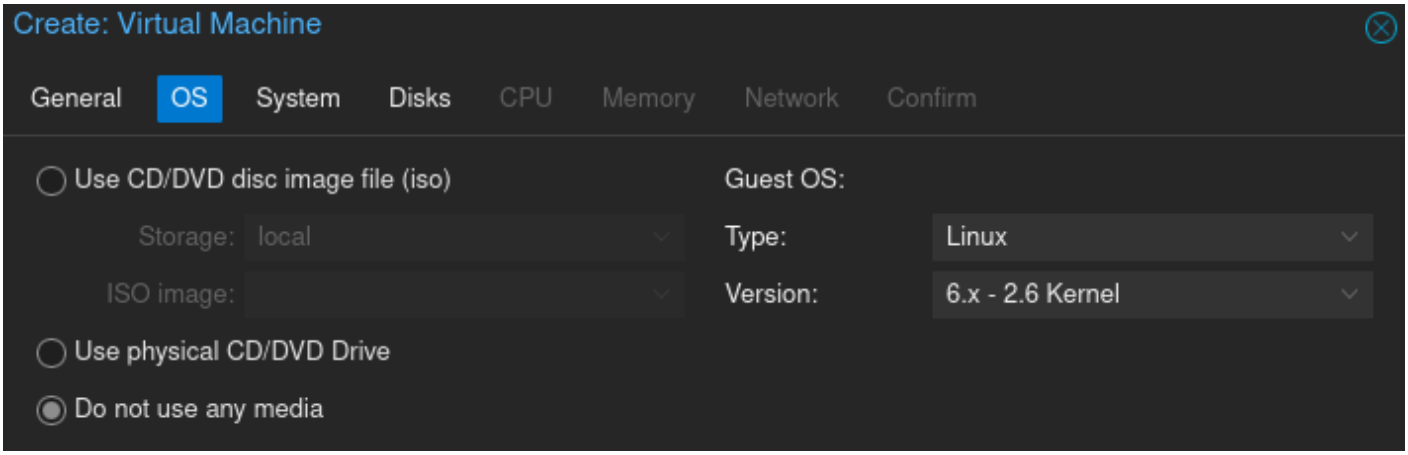
This page explains the process of importing VMs from TrueNAS Core, which uses FreeBSD's Bhyve for virtualization, to Proxmox.

Details

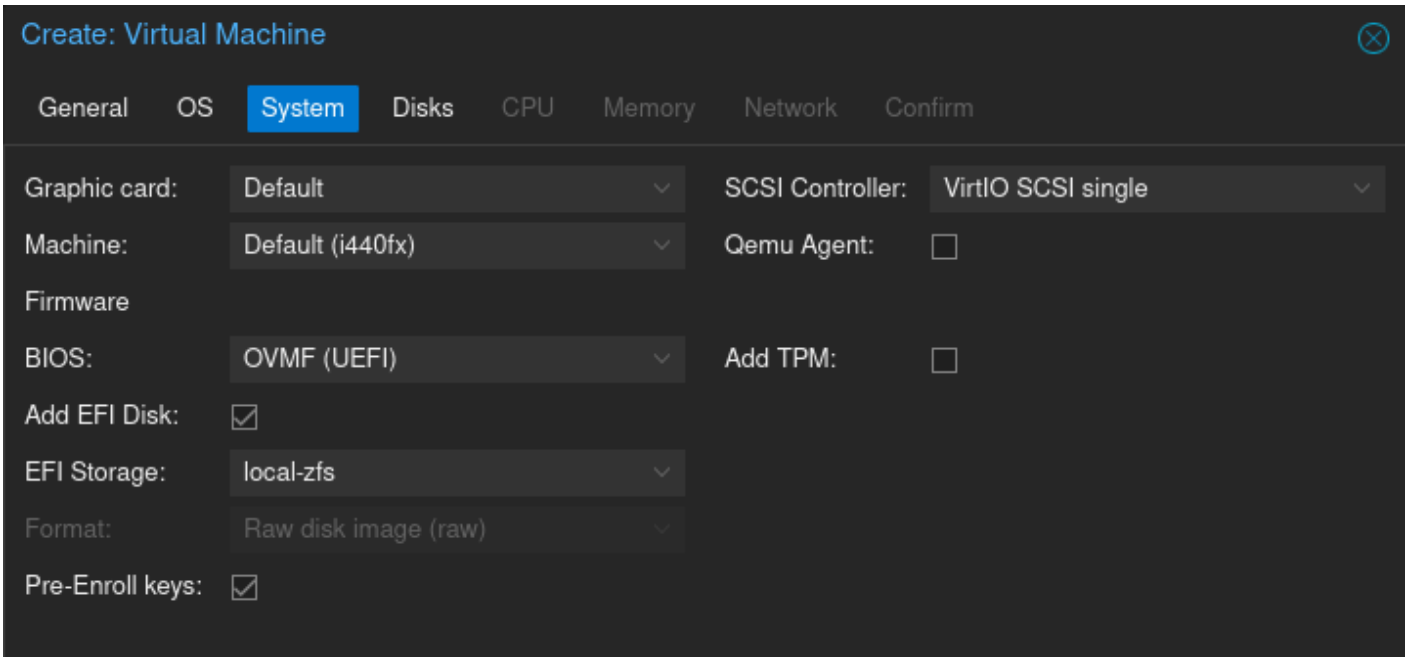
Proxmox

In Proxmox, create a new VM and note its VM number. When creating the VM, follow these guidelines:

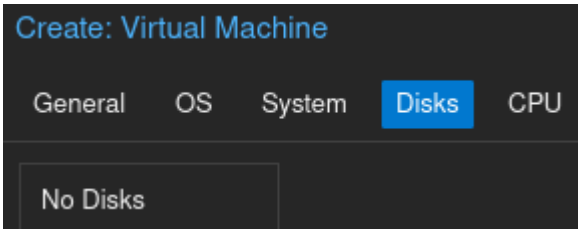
In the OS section, select "Do not use any media"



In the System section, select "OVMF (UEFI)" for BIOS. Also select EFI Storage to be the same dataset as where you would like your VM's disk to be. We chose the default local-zfs dataset, but you may choose any other dataset, such as an encrypted dataset if you want your VMs to be encrypted.



In the disk section, remove the default disk and do not set a disk.



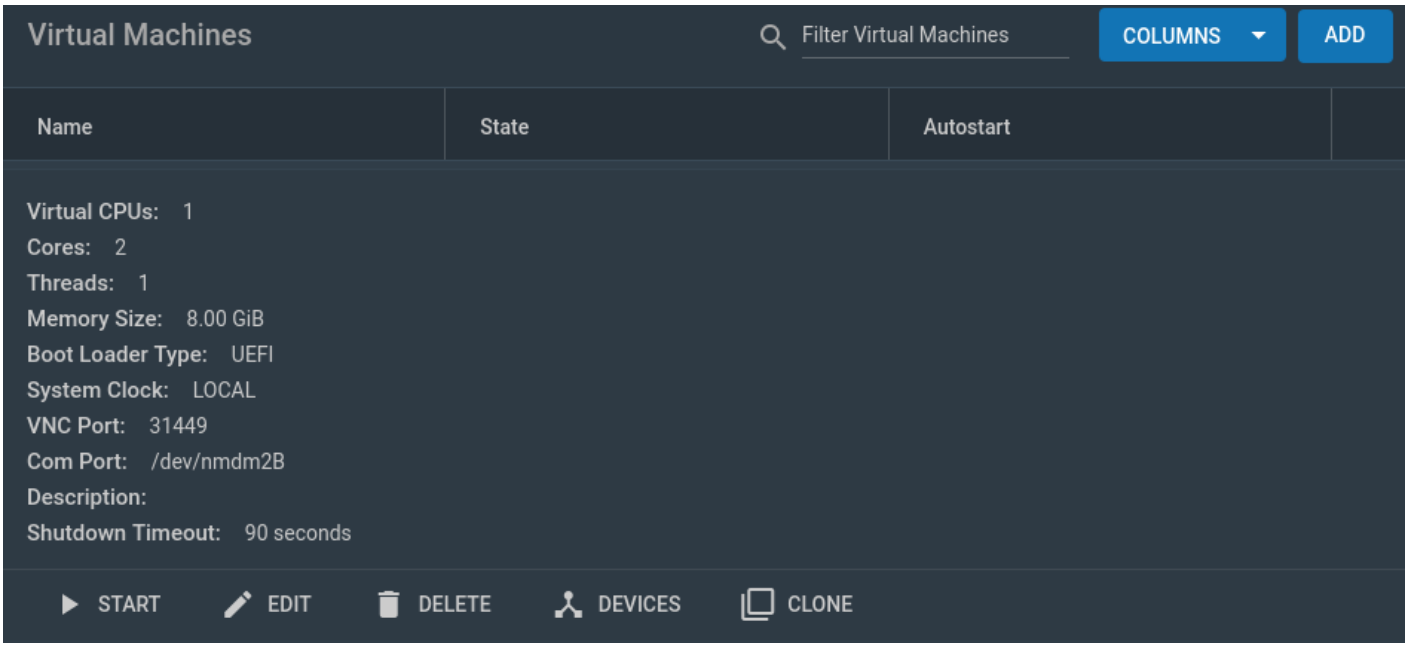
Continue with the rest of the sections per your own personal requirements.

TrueNAS

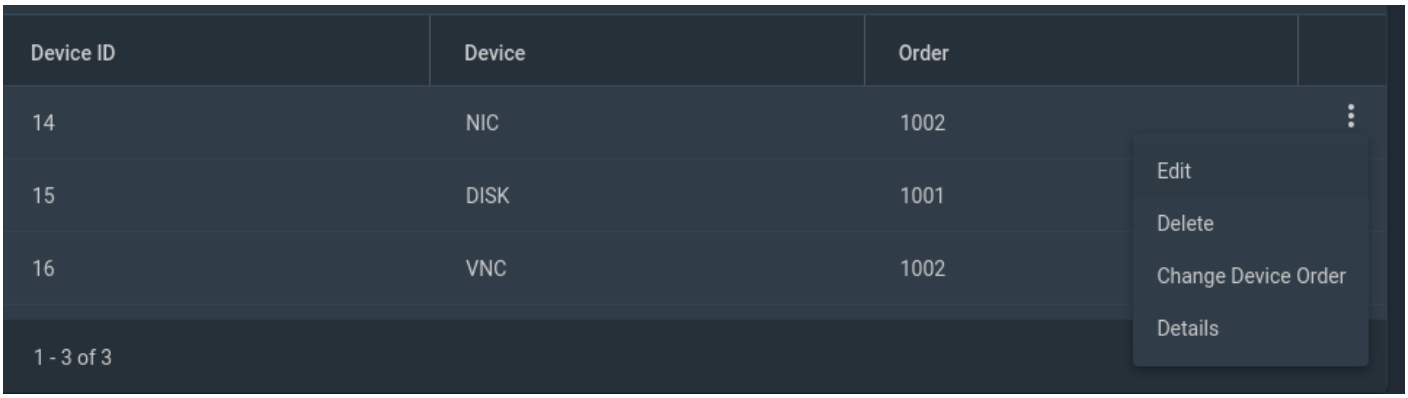
Shutdown the VM in TrueNAS and make a snapshot of the VM dataset in TrueNAS. Login to SSH in TrueNAS as the root user and run the following command to send the dataset using SSH to Proxmox:

```
zfs send [VM_Dataset]@[snapshot_name] | ssh root@proxmox 'zfs receive rpool/[any dataset here]/vm-[num]-disk-1'
```

If you use DHCP in your network and you would like your IP address for the VM to be identical after migration, press the devices button in your virtual machine menu:



Then, press the three dots over the "NIC" device and press Edit.



A new menu should show up which displays the MAC address. Copy this MAC address. In Proxmox, you may edit your Network Device and paste the MAC address there.

Back to Proxmox

Back in Proxmox, login to the root shell and run the `qm rescan` command. Then, go into your VM's hardware menu. The disk should show up as an unattached device. You may now attach it.

Congratulations, you have successfully migrated a virtual machine from TrueNAS Core to Proxmox!

Source Description Block

> Multiple Sources:
> https://forum.proxmox.com/threads/adding-existing-disk-from-storage-to-vm.108645/
> https://www.youtube.com/watch?v=yKZ_JJaQHDk

Licensing

Image Credit - Book Cover Art

Photo by Nadin Sh from Pexels.

The photo represents my feelings about DevOps and related tooling.