

UKI Addons and extensions

safely extending UKIs kernel command line and initrd

Emanuele Giuseppe Esposito

February 03, 2024

Extremely new stuff!



- Extremely new stuff!
- Not a lot of documentation



- Extremely new stuff!
- Not a lot of documentation
- UKIs are flexible, and security is not sacrificed



- Extremely new stuff!
- Not a lot of documentation
- UKIs are flexible, and security is not sacrificed
- Attempt to advertise UKIs and their features



Storage encryption

Let's first look at Vitaly's slides...









Confidential VM provides protection from the host it runs on:

"Protection" means strong security boundary for all data in the VM. Malicious hypervisor or an actor having access (even privileged!) to the host should not be able to get access to the data.





- "Protection" means strong security boundary for all data in the VM. Malicious hypervisor or an actor having access (even privileged!) to the host should not be able to get access to the data.
- ► The host is still able to disrupt execution of the VM, e.g. it can stop it.





- "Protection" means strong security boundary for all data in the VM. Malicious hypervisor or an actor having access (even privileged!) to the host should not be able to get access to the data.
- ► The host is still able to disrupt execution of the VM, e.g. it can stop it.
- Hardware (AMD SEV-SNP, Intel TDX) is responsible for encrypting memory and CPU state.

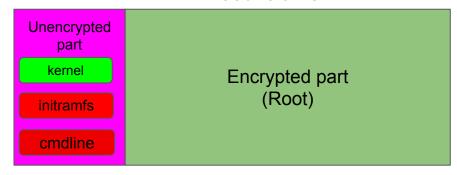




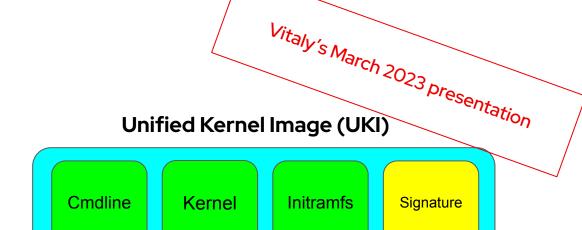
- "Protection" means strong security boundary for all data in the VM. Malicious hypervisor or an actor having access (even privileged!) to the host should not be able to get access to the data.
- The host is still able to disrupt execution of the VM, e.g. it can stop it.
- Hardware (AMD SEV-SNP, Intel TDX) is responsible for encrypting memory and CPU state.
- Storage encryption is necessary for security and must be done by the guest OS.



Root volume



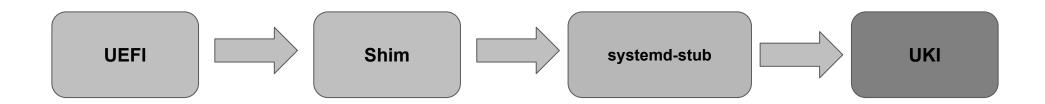
- While kernel binary is signed by Red Hat, initramfs and kernel command line are locally produced and are not signed.
- Locally produced initramfs/cmdline have unpredictable measurements.



- A single binary (UEFI application)
 produced and signed in Red Hat
 build system.
- The base for building UKI is systemd-stub.
- Contains vmlinuz, initramfs, and cmdline as PF sections.



Booting UKI



Vitaly's March 2023 presentation

Kernel cmdline is now immutable

- Systemd GPT auto generator (<u>link</u>) must be used instead of "root="
- "Limited" customization is still required:
 - · "crashkernel=" like options
 - debugging, tuning options
- A mechanism to have more than one cmdline in the UKI was requested (<u>link</u>).
- An additional "allowlist" of options which are allowed for customization is needed.
 - E.g. the basic "root=", "init=",... can't be allowed





- It cannot be static (one fits all),
 - Production and debug options
 - Some options like "crashkernel" cannot be hardcoded and are os-dependent.



- It cannot be static (one fits all),
 - Production and debug options
 - Some options like "crashkernel" cannot be hardcoded and are os-dependent.
- Secure
 - · Whoever modifies the cmdline is authenticated.
 - · By default, nobody.



- It cannot be static (one fits all),
 - Production and debug options
 - Some options like "crashkernel" cannot be hardcoded and are os-dependent.
- Secure
 - Whoever modifies the cmdline is authenticated.
 - · By default, nobody.
- Easily extensible
 - No need from RH to ship a new UKI every time cmdline changes, or have multiple UKIs with multiple cmdline



Adding kernel cmdline to an UKI







.cmdline PE section inside the UKI, created when the UKI is generated





- .cmdline PE section inside the UKI, created when the UKI is generated
- Advantages
 - · Secure, measured and shipped with UKI

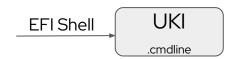




- .cmdline PE section inside the UKI, created when the UKI is generated
- Advantages
 - Secure, measured and shipped with UKI
- Disadvantages
 - Static, impossible to modify unless UKI is re-generated and shipped again

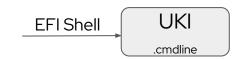






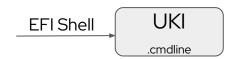


- systemd-stub looks at this only if .cmdline is missing
 - · This won't happen in RHEL UKIs



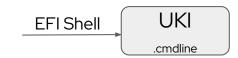


- systemd-stub looks at this only if .cmdline is missing
 - This won't happen in RHEL UKIs
- Advantages
 - Useful for type1 entries





- systemd-stub looks at this only if .cmdline is missing
 - This probably won't happen in distro UKIs
- Advantages
 - Useful for type1 entries
- Disadvantages
 - · Unsafe, an attacker can easily inject its own parameters

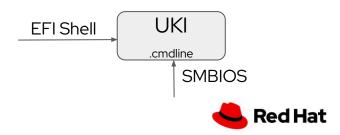




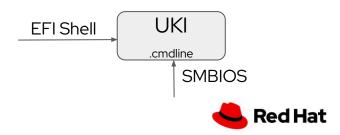
- systemd-stub looks at this only if .cmdline is missing
 - · This probably won't happen in distro UKIs
- Advantages
 - Useful for type1 entries
- Disadvantages
 - Unsafe, an attacker can easily inject its own parameters
- As a result, EFI Shell parameters are completely ignored in CVMs https://github.com/systemd/systemd/pull/28763





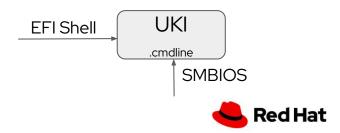


On baremetal, SMBIOS is trusted, as it is coming from firmware/BIOS.

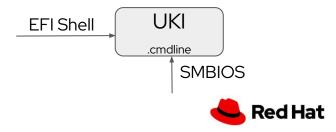


- On baremetal, SMBIOS is trusted, as it is coming from firmware/BIOS.
 - · Does not apply on CVMs, as hypervisor injects it

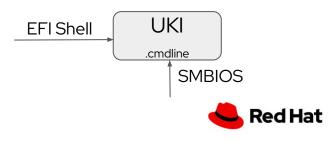
qemu-kvm [...] -smbios type=11,value=io.systemd.stub.kernel-cmdline-extra=MY_CMDLINE



- On baremetal, SMBIOS is trusted, as it is coming from firmware/BIOS.
 - Does not apply on CVMs, as hypervisor injects it
 # qemu-kvm [...] -smbios type=11,value=io.systemd.stub.kernel-cmdline-extra=MY_CMDLINE
- Advantages
 - Useful for baremetal



- On baremetal, SMBIOS is trusted, as it is coming from firmware/BIOS.
 - Does not apply on CVMs, as hypervisor injects it
 # qemu-kvm [...] -smbios type=11,value=io.systemd.stub.kernel-cmdline-extra=MY_CMDLINE
- Advantages
 - Useful for baremetal
- Disadvantages
 - Unsafe, a malicious hypervisor can easily inject its own parameters

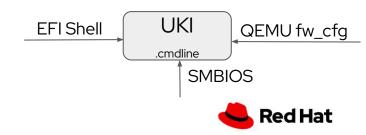


- On baremetal, SMBIOS is trusted, as it is coming from firmware/BIOS.
 - Does not apply on CVMs, as hypervisor injects it
 # qemu-kvm [...] -smbios type=11,value=io.systemd.stub.kernel-cmdline-extra=MY_CMDLINE
- Advantages
 - Useful for baremetal
- Disadvantages
 - Unsafe, a malicious hypervisor can easily inject its own parameters
- As a result, SMBIOS parameters are completely ignored in CVMs https://github.com/systemd/systemd/issues/27604



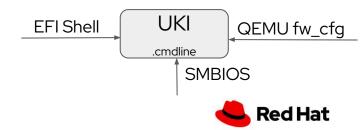
4. QEMU Firmware Configuration (fw_cfg)

- Can be used the same way as SMBIOS is
 - No baremetal, only QEMU.



4. QEMU Firmware Configuration (fw_cfg)

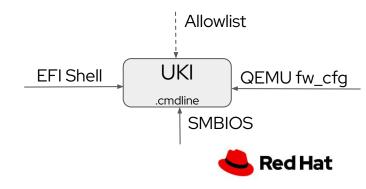
- Can be used the same way as SMBIOS is
 - No baremetal, only QEMU.
- Disadvantages
 - Unsafe, a malicious hypervisor can easily inject its own parameters



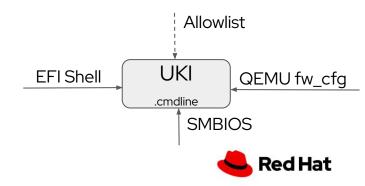
4. QEMU Firmware Configuration (fw_cfg)

- Can be used the same way as SMBIOS is
 - No baremetal, only QEMU.
- Disadvantages
 - · Unsafe, a malicious hypervisor can easily inject its own parameters
- As a result, fw_cfg parameters are completely ignored in CVMs

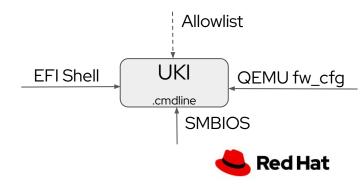




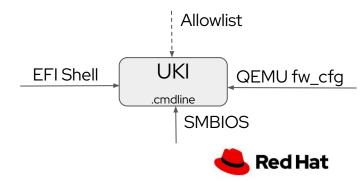
- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)



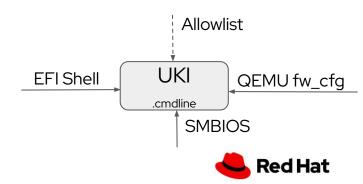
- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)
- Command line is coming from EFI Shell or SMBIOS



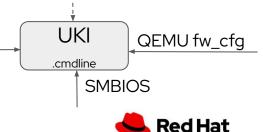
- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)
- Command line is coming from EFI Shell or SMBIOS
- Systemd-stub takes care of parsing the parameters against the allowlist



- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)
- Command line is coming from EFI Shell or SMBIOS
- Systemd-stub takes care of parsing the parameters against the allowlist
- Advantages
 - · Allowlist takes care of rejecting unwanted parameters



- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)
- Command line is coming from EFI Shell or SMBIOS
- Systemd-stub takes care of parsing the parameters against the allowlist
- Advantages
 - · Allowlist takes care of rejecting unwanted parameters
- Disadvantages
 - Regex and globbing need to be very very carefully formulated EFI Shell



Allowlist

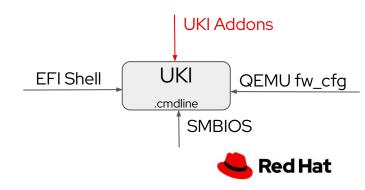
- Create an allowlist to be stored in .allowlist PE section
 - Regex/globbing of allowed parameters (ex console=ttys*)
- Command line is coming from EFI Shell or SMBIOS
- Systemd-stub takes care of parsing the parameters against the allowlist
- Advantages
 - Allowlist takes care of rejecting unwanted parameters
- Disadvantages
 - Regex and globbing need to be very very carefully formulated
- This proposal was rejected by the systemd upstream community
- https://github.com/systemd/systemd/issues/24539



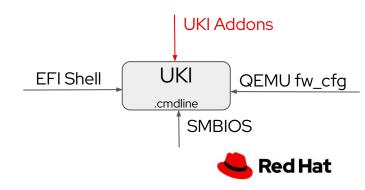


BETTER LUCK NEXT TIM

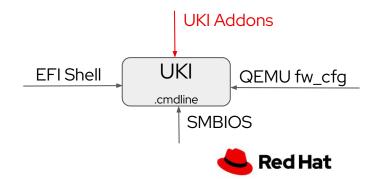




Use shim shim_validate() function to validate PE signatures

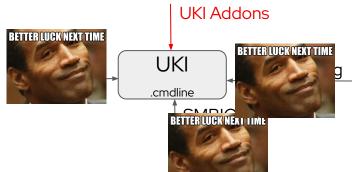


- Use shim shim_validate() function to validate PE signatures
- ukify: systemd tool that is able to build UKIs and much more https://www.freedesktop.org/software/systemd/man/ukify.html
 - Way simpler and faster than dracut and objcopy
 - · Available from systemd v253, but many features in v254 and after



- Use shim shim_validate() function to validate PE signatures
- ukify: systemd tool that is able to build UKIs and much more https://www.freedesktop.org/software/systemd/man/ukify.html
 - Way simpler and faster than dracut and objcopy
 - · Available from systemd v253, but many features in v254 and after
- ukify creates a PE file (addon) containing only .cmdline and other relevant sectior
 - It also signs the PE with a provided key

/usr/lib/systemd/ukify build --signtool=pesign --secureboot-certificate-name='UKI' --cmdline='MY_CMDLINE' --output=\$BOOT/efi/EFI/Linux/my_addon.addon.efi





UKI addons





1. ukify creates the addon



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon
- 5. Shim verifies the addon



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon
- 5. Shim verifies the addon
- 6. If validation is successful, systemd-stub reads the addon



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon
- 5. Shim verifies the addon
- 6. If validation is successful, systemd-stub reads the addon
- 7. systemd-stub gets .cmdline and appends it to UKI .cmdline section



- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon
- 5. Shim verifies the addon
- 6. If validation is successful, systemd-stub reads the addon
- 7. systemd-stub gets .cmdline and appends it to UKI .cmdline section
- 8. Provide the final cmdline to the vmlinuz contained in .linux



Global and local addons



Global and local addons

- ► Global addons: applied to **all** installed UKIs
 - \$BOOT/efi/loader/addons



Global and local addons

- Global addons: applied to all installed UKIs
 - \$BOOT/efi/loader/addons
- UKI-specific addons: applied to the specific UKI
 - Example: 'UKI_devel' installed as \$BOOT/efi/EFI/Linux/devel.efi
 - → all UKI_devel specific addons are installed in \$BOOT/efi/EFI/linux/devel.efi.extra.d/





UKIs are always located in \$BOOT/efi/EFI/Linux/



- UKIs are always located in \$BOOT/efi/EFI/Linux/
- UKIs are always ending with .efi



- UKIs are always located in \$BOOT/efi/EFI/Linux/
- UKIs are always ending with .efi
- UKI addons always end with .addon.efi



- UKIs are always located in \$BOOT/efi/EFI/Linux/
- UKIs are always ending with .efi
- UKI addons always end with .addon.efi
- UKI-specific addons are always located in \$BOOT/efi/EFI/Linux/<UKI_NAME>.efi.extra.d/



SBAT
(Secure
Boot
Advanced Targeting)



► Imagine the UKI/addon is issued and signed by a company like Red Hat



- ► Imagine the UKI/addon is issued and signed by a company like Red Hat
- Solution 1: change certificate
 - Means invalidating all TPM measurements
 - Invalidates all other UKIs and addons
 - Impractical



- ► Imagine the UKI/addon is issued and signed by a company like Red Hat
- Solution 1: change certificate
 - Means invalidating all TPM measurements
 - Invalidates all other UKIs and addons
 - Impractical
- Solution 2: add the hash of the addon to some Cloud provider blacklist



- ► Imagine the UKI/addon is issued and signed by a company like Red Hat
- Solution 1: change certificate
 - Means invalidating all TPM measurements
 - Invalidates all other UKIs and addons
 - Impractical
- Solution 2: add the hash of the addon to some Cloud provider blacklist
- Solution 3: at attestation time, check if the addon with a specific hash is being measured. If so, reject it.



- Solution 4: SBAT rules
 - Add a .sbat version "component,generation,vendor,pkg,pkg_version,url"
 - Shim checks its own sbat "component,generation" tuple with addon .sbat, if there is a match and shim generation is higher than generation, ignore the addon



UKI addons: workflow

- 1. ukify creates the addon
- 2. Addon is put in a specific location in the ESP
- 3. systemd-stub looks for addons, finds them
- 4. systemd-stub calls shim_verify() on the addon
- 5. Shim verifies the addon and checks SBAT component and generation
- 6. If validation is successful, systemd-stub reads the addon
- 7. systemd-stub gets .cmdline and appends it to UKI .cmdline section
- 8. Provide the final cmdline to the vmlinuz contained in .linux



Guest SBAT variable:

```
sbat, 1
my_addon, 2
```

```
sbat, 1, SBAT Version, sbat, 1, https://github.com/rhboot/shim/blob/main/SBAT.md my_addon, 2, My addon version, addon_product, 2.0, www.mycompany.com
```



Guest SBAT variable:

```
sbat, 1
my_addon, 2
```

```
sbat, 1, SBAT Version, sbat, 1, https://github.com/rhboot/shim/blob/main/SBAT.md my_addon, 2, My addon version, addon_product, 2.0, www.mycompany.com
```





Guest SBAT variable:

```
sbat, 1
my_addon, 2
```

```
sbat, 1, SBAT Version, sbat, 1, https://github.com/rhboot/shim/blob/main/SBAT.md my_addon, 1, My addon version, addon_product, 2.0, www.mycompany.com
```



Guest SBAT variable:

```
sbat, <mark>1</mark>
my_addon, <mark>2</mark>
```

```
sbat, 1, SBAT Version, sbat, 1, https://github.com/rhboot/shim/blob/main/SBAT.md my_addon, 1, My addon version, addon_product, 2.0, www.mycompany.com
```





Open Problem: combining addons

- What if UKI+addonA is valid, UKI+ addonB, but UKI + addonA + addonB creates security issues
 - · Couldn't come up with a concrete example yet
 - Only solution would be to use attestation and see if addonA and addonB are measured, and if so reject the verification



Systemd-sysext initrd addons



► To extend initramfs or even extend the host fs.

"A system extension image extend the base system with an overlay containing additional files."



- ► To extend initramfs or even extend the host fs.

 "A system extension image extend the base system with an overlay containing additional files."
- Same concept, different tools
 - · System extension image vs PE file
 - mkosi vs ukify to create a signed sysext (requires dm-verity loaded)
 - · Same path where to put it



- ► To extend initramfs or even extend the host fs
 "A system extension image extend the base system with an overlay containing additional files."
- Same concept, different tools
 - System extension image vs PE file
 - mkosi vs ukify to create a signed sysext (requires dm-verity loaded)
 - · Same path where to put it
- Create sysext extension



- ► To extend initramfs or even extend the host fs
 "A system extension image extend the base system with an overlay containing additional files."
- Same concept, different tools
 - System extension image vs PE file
 - mkosi vs ukify to create a signed sysext (requires dm-verity loaded)
 - · Same path where to put it
- Create sysext extension
- Put it into \$BOOT/efi/EFI/Linux/\$UKI.efi.extra.d (must be a .raw)



- ► To extend initramfs or even extend the host fs.

 "A system extension image extend the base system with an overlay containing additional files."
- Same concept, different tools
 - · System extension image vs PE file
 - mkosi vs ukify to create a signed sysext (requires dm-verity loaded)
 - · Same path where to put it
- Create sysext extension
- Put it into \$BOOT/efi/EFI/Linux/\$UKI.efi.extra.d (must be a .raw)
- systemd-stub will take care of copying it into initrd's /.extra/sysext/ folder



- To extend initramfs or even extend the host fs. "A system extension image extend the base system with an overlay containing additional files."
- Same concept, different tools
 - System extension image vs PE file
 - mkosi vs ukify to create a signed sysext (requires dm-verity loaded)
 - Same path where to put it
- Create sysext extension
- Put it into \$BOOT/efi/EFI/Linux/\$UKI.efi.extra.d (must be a .raw)

https://github.com/systemd/mkosi/commit/c42d816

- systemd-stub will take care of copying it into initrd's /.extra/sysext/ folder
- systemd-sysext will take care of taking the extension and using it before switching to root



Target users



► 3 group of users can create them:



- ► 3 group of users can create them:
- Vendors (RH) that have already their keys inserted into the machine secureboot database
 - · Add machine-specific cmdline (in the cloud, different VMs with different features), debug addons for developers to debug kernel issues, ...



- ▶ 3 group of users can create them:
- Vendors (RH) that have already their keys inserted into the machine secureboot database
 - Add machine-specific cmdline (in the cloud, different VMs with different features), debug addons for developers to debug kernel issues, ...
- Virt host admins that use host-side tools like virt-firmware to inject keys in machine OVMF variables
 - · Custom cmdline, debug addons, ...



- ▶ 3 group of users can create them:
- Vendors (RH) that have already their keys inserted into the machine secureboot database
 - Add machine-specific cmdline (in the cloud, different VMs with different features), debug addons for developers to debug kernel issues, ...
- Virt host admins that use host-side tools like virt-firmware to inject keys in machine OVMF variables
 - · Custom cmdline, debug addons, ...
- Guest admins that use guest-side tools like MOK to insert keys in the secureboot db
 - Note: usually not allowed by cloud providers, like Azure
 - · Add custom cmdline, debug addons, ...



Available tools



Systemd tools

- v253: ukify capable of creating UKIs
- v254: ukify support for UKI addons (`ukify build`)
- v255: ukify support for UKI/addons inspection (`ukify inspect`)
- Features still to merge:
 - Enable bootclt to find the addons and display for each UKI the full cmdline (default + all used addons)
- mkosi: create systemd-syext images



uki-direct (part of virt-firmware)

 kernel-bootcfg: add, update, remove UKIs (generally show and manage uefi boot entries)



uki-direct (part of virt-firmware)

- kernel-bootcfg: add, update, remove UKIs (generally show and manage uefi boot entries)
- Future releases:
 - · kernel-addon: add, update, inspect and remove UKI addons
 - Requires `ukify inspect`



Future work





- Ship a RPM with a collection of addons (debug, cloud-specific)
 - · Signed by the vendor



- Ship a RPM with a collection of addons (debug, cloud-specific)
 - · Signed by the vendor
- RPM installs them into a specific location not in ESP (/usr/lib/linux/extra.d/ for global, /usr/lib/linux/\$UNAME/\$UKI.efi.extra.d/ for UKI-specific)



- Ship a RPM with a collection of addons (debug, cloud-specific)
 - · Signed by the vendor
- RPM installs them into a specific location not in ESP (/usr/lib/linux/extra.d/ for global, /usr/lib/linux/\$UNAME/\$UKI.efi.extra.d/ for UKI-specific)
- Use kernel-addon to install them globally or to a specific UKI



- Ship a RPM with a collection of addons (debug, cloud-specific)
 - Signed by the vendor
- RPM installs them into a specific location not in ESP (/usr/lib/linux/extra.d/ for global, /usr/lib/linux/\$UNAME/\$UKI.efi.extra.d/ for UKI-specific)
- Use kernel-addon to install them globally or to a specific UKI
- Useful when customer has a bug and developer needs to debug UKI



Cloud



Cloud

- Cloud providers need to provide a way to the user to inject his own certificate into the secureboot db
 - Otherwise custom addons cannot be added



Cloud

- Cloud providers need to provide a way to the user to inject his own certificate into the secureboot db
 - · Otherwise custom addons cannot be added
- This also implies that the certificate must be measured in PCR7
 - Solution: add dummy addon at first boot, so that the cert is measured



On prem



On prem

Libvirt should do the same as what the cloud provider should offer: possibility to upload a certificate for secureboot

https://issues.redhat.com/browse/RHEL-9690



On prem

- Libvirt should do the same as what the cloud provider should offer: possibility to upload a certificate for secureboot
 - https://issues.redhat.com/browse/RHEL-9690
- Insert dummy addon for measurements with `virt-customize --upload`



Questions?



Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

- in linkedin.com/company/red-hat
- youtube.com/user/RedHatVideos
- facebook.com/redhatinc
- **y** twitter.com/RedHat

