Running Linux in a Shielded VM

Michael Kelley
Principal Software Engineer
Microsoft Corporation

Microsoft

Introduction

Linux Shielded VMs

Capability of the Hyper-V hypervisor

Originally introduced for Windows guests, but now works for Linux guests as well

A mix of Windows/Hyper-V technologies and open source Linux technologies

This session focuses on the Linux aspects

Quick overview of the problem (and what's out of scope)

End-to-end view of the components that enable Linux Shielded VMs (LSVMs)

Drilldown into 6 technical areas

Wrap-up with current status of Linux Shielded VMs

What is the problem?

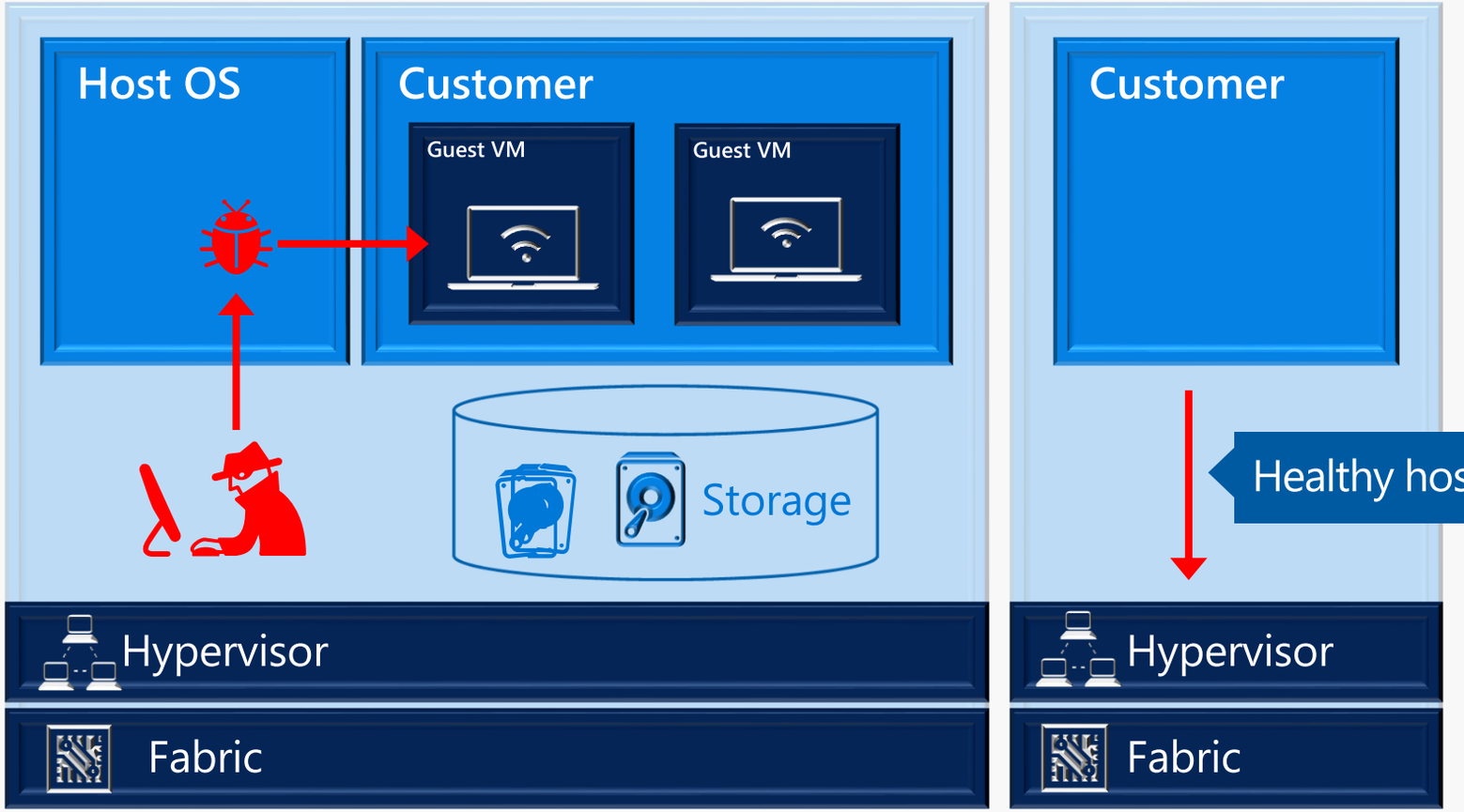
Virtual machines introduce new security risks

Compromised or malicious fabric admins can access guest virtual machines

Health of hosts not taken into account before running VMs

Tenant's VMs are exposed to storage and network attacks

VMs don't benefit from hardware-rooted security such as TPMs



Shielded VMs: Security Assurance Goals

Encryption of data, both at-rest & in-flight

Virtual TPM enables the use of disk encryption within a VM (e.g. dm-crypt, BitLocker)

Both Live Migration and VM state are encrypted

Fabric admins locked out

Host administrators cannot access guest VM secrets (e.g. can't see disks, video, etc.)

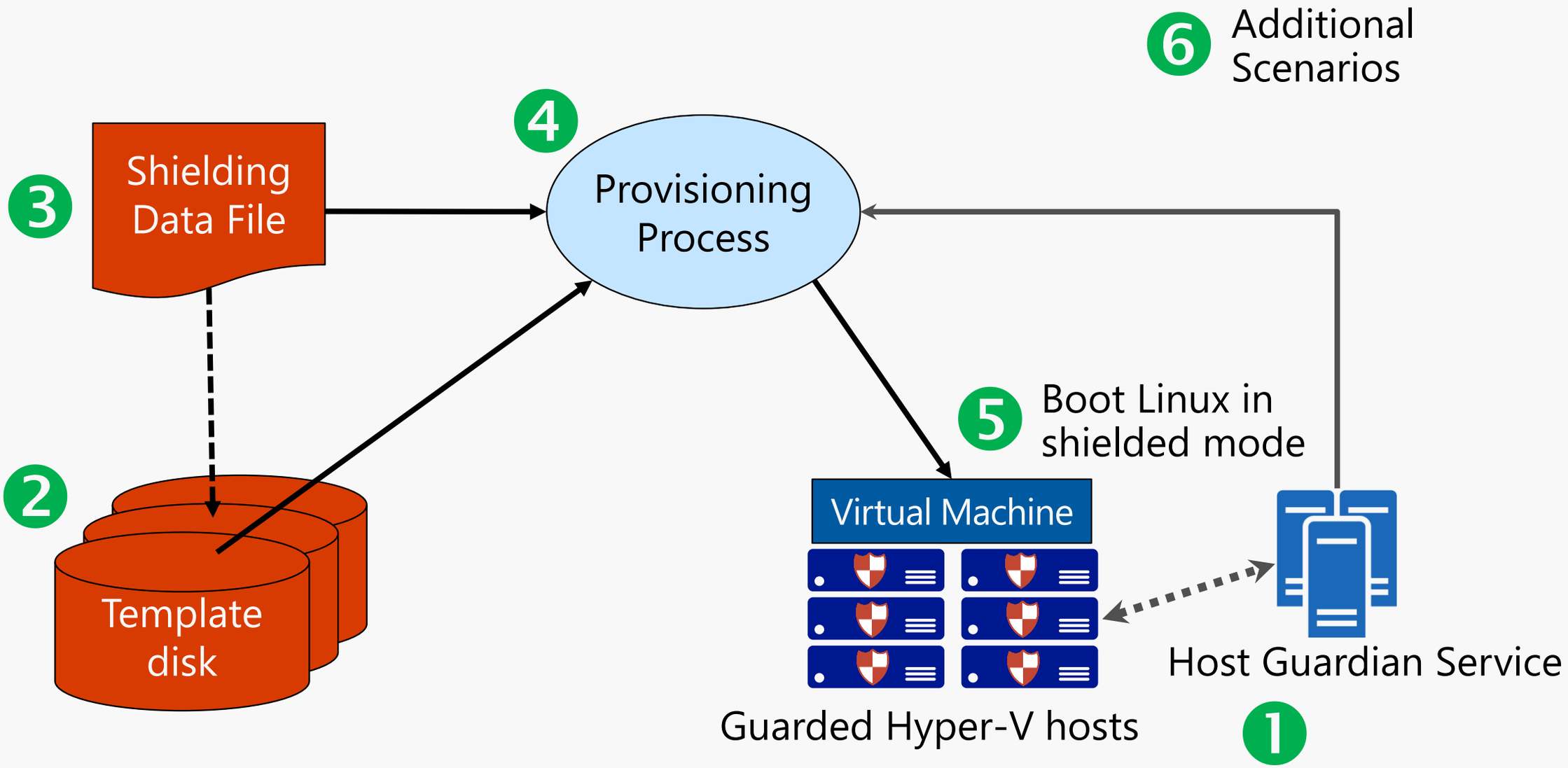
Host administrators cannot run arbitrary kernel-mode code

Attestation of host health required

VMs can only run on "healthy" hosts designated by the VM owner

NOTE: Shielding is not intended as a defense against DoS attacks

End-to-End Flow



Drilldown: Guarded Fabric (1 of 2)

Key Goals

Know and trust the hypervisor software you are running on
Prevent the introduction of malware or guest VM observing tools into the hypervisor
Protect VM secrets as VMs are created and initially booted

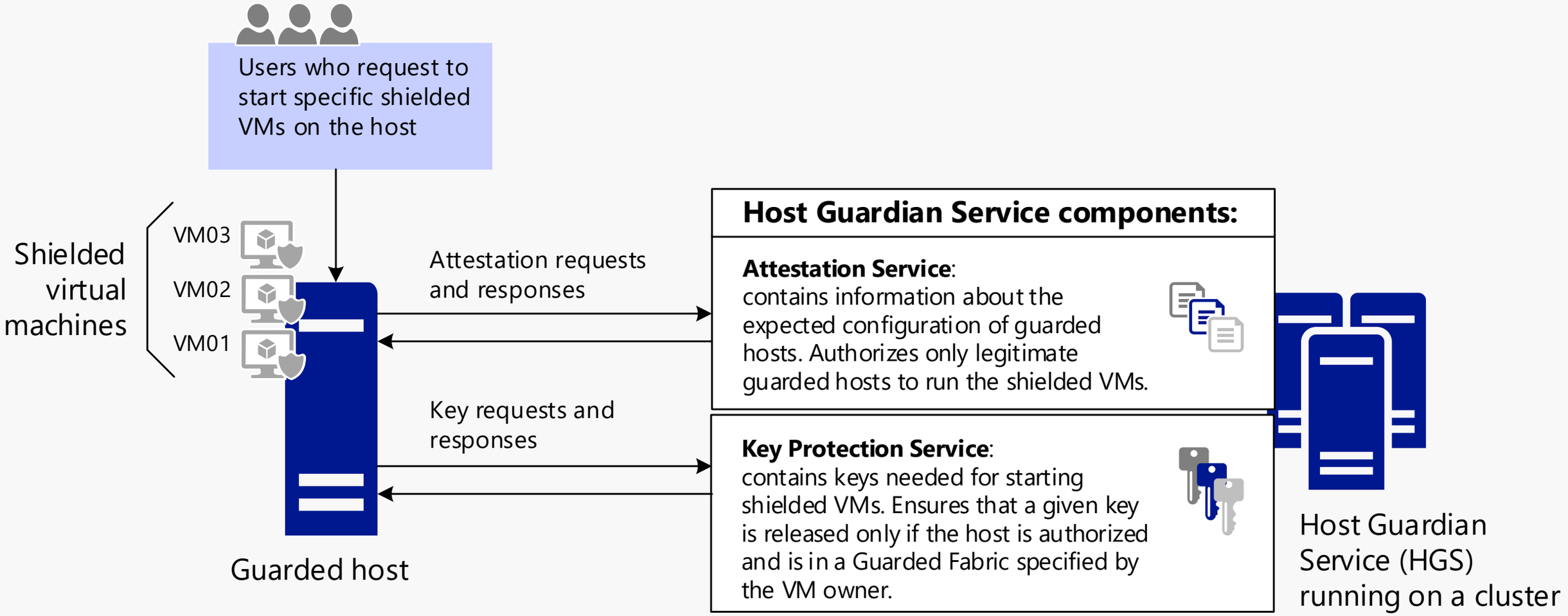
Host Guardian Service

Attestation of guarded hosts based on hardware TPM
Release keys to run a VM

Guarded Host running Hyper-V

Boot sequence is validated via physical TPM
Code Integrity policy is enabled – only authorized executables can run
Disabled functionality: debug of VM worker process, guest VM console, most host/guest data exchange
Supplies a “virtual” TPM (vTPM) to the guest VM
Supports Virtualization-Based Security (VBS) for protecting vTPM data and other secrets

Drilldown: Guarded Fabric (2 of 2)



Drilldown: Linux Template Disk (1 of 4)

Key Goals

Base level Linux OS installation from which to create a shielded VM
Should be cloneable as the base for multiple VMs
Should not contain any secrets

Creating a Linux Template Disk

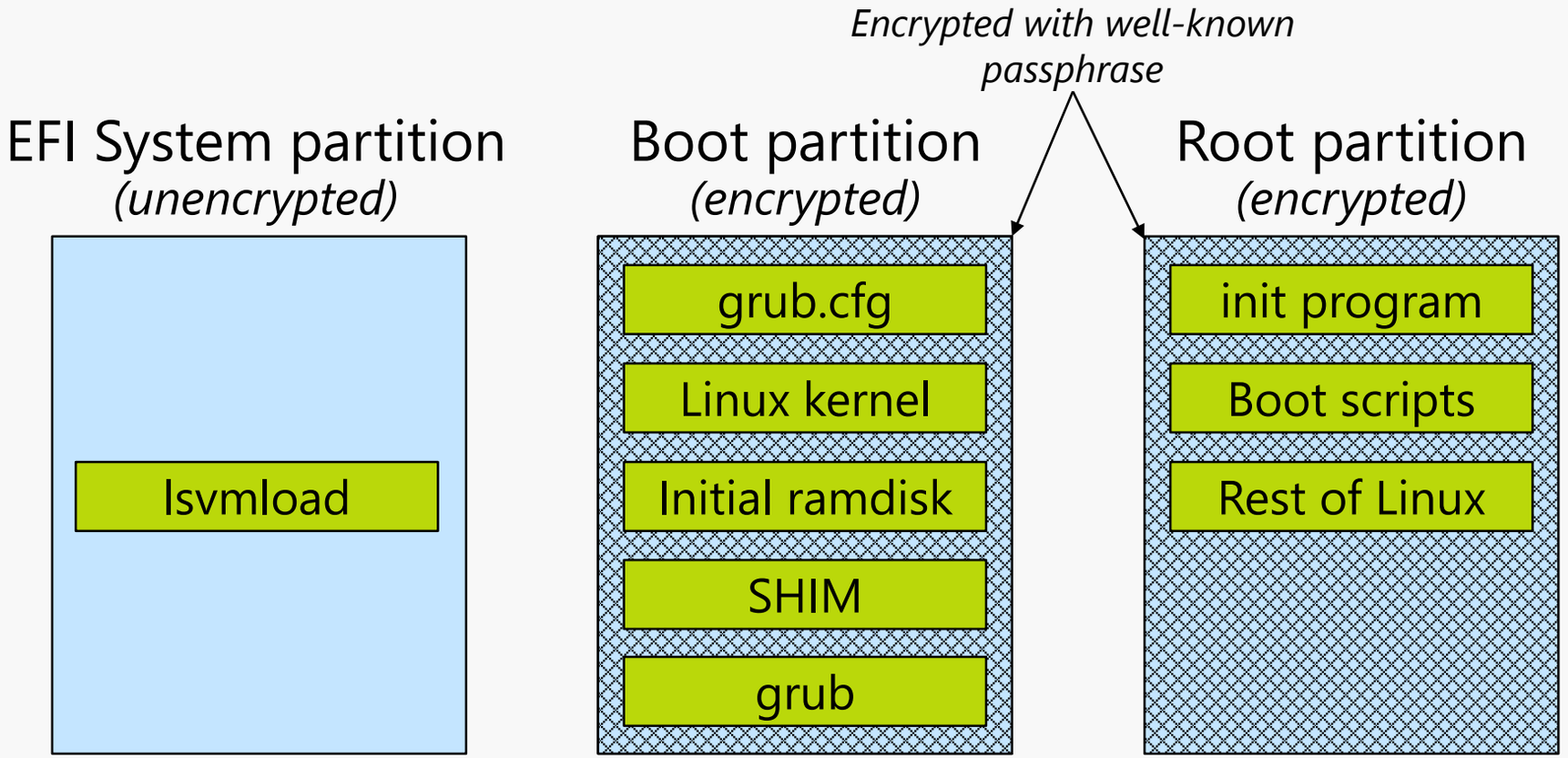
Do a standard Linux installation into the VM (typically on a standalone Hyper-V)
Must be in VHDX format as a “Generation 2” VM so can boot using UEFI Secure Boot
Set up root partition as encrypted using LUKS/dm-crypt with a well-known pass phrase
Install VMM agent for Linux to handled specialization

Drilldown: Linux Template Disk (2 of 4)

Run “lsvmprep” tool in the Linux VM

Makes transformations to help protect the boot path

Updates initramfs to get dm-crypt passphrase from a “file” instead of console



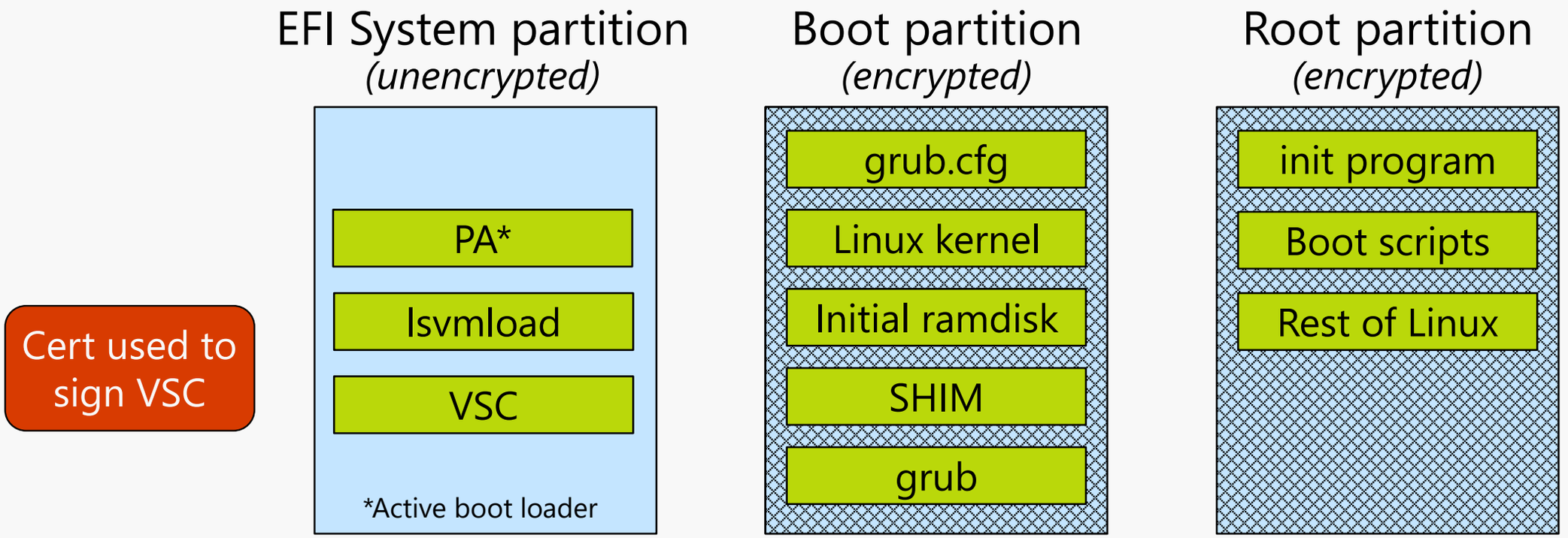
Drilldown: Linux Template Disk (3 of 4)

Run "Template Disk Creation" wizard in Windows

Creates Volume Signature Catalog – hash of all disk blocks in the boot and root partitions

Puts the VSC in the EFI System Partition, and signs it with a certificate you provide

Puts the Provisioning Agent (PA) into EFI System Partition as the active boot loader



Drilldown: Linux Template Disk (4 of 4)

Linux Template Disk Summary

❶ Hyper-V virtual disk image (VHDX) that can be used to create multiple VMs

❷ Standard Linux install with LUKS/dm-crypt encryption on boot and root partitions, using a well-known passphrase

❸ Microsoft-signed early boot loader (lsvmload) is installed

❹ Creator is asserting that the template disk is “good” when the VSC is created

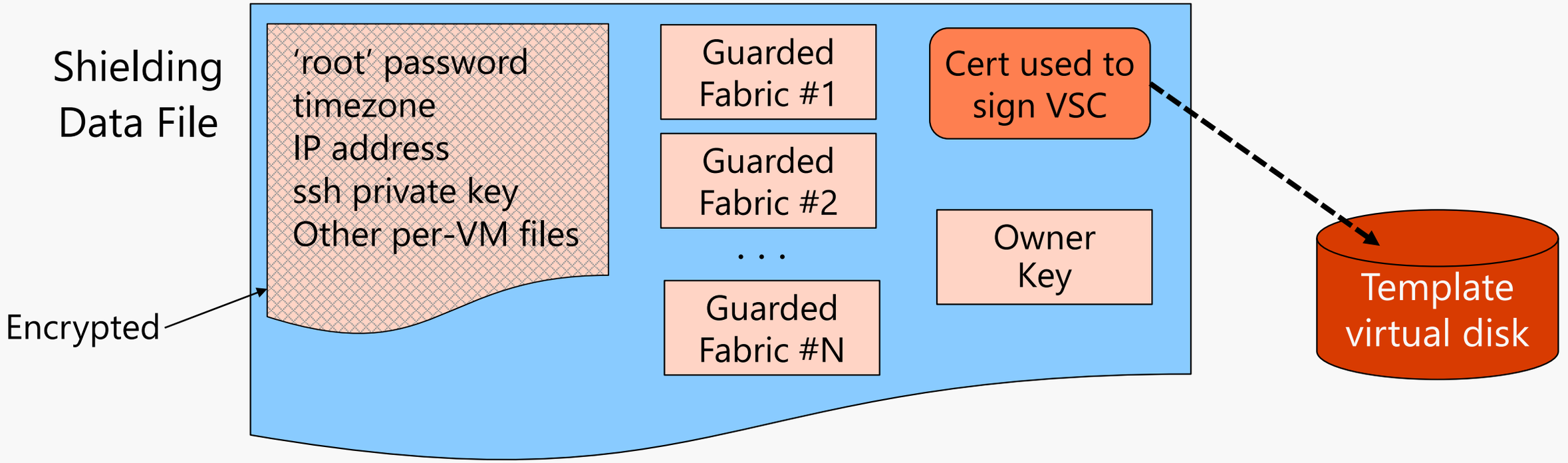
❺ VSC is used later to detect (and reject) any subsequent modifications to the disk

Drilldown: Shielding Data File

Contains per-VM data and secrets

Links together components for secure deployment

Created using Shielding Data File Wizard in Windows



Drilldown: Provisioning (1 of 3)

Key Goals

Create the shielded Linux VM, using a clone of the template disk

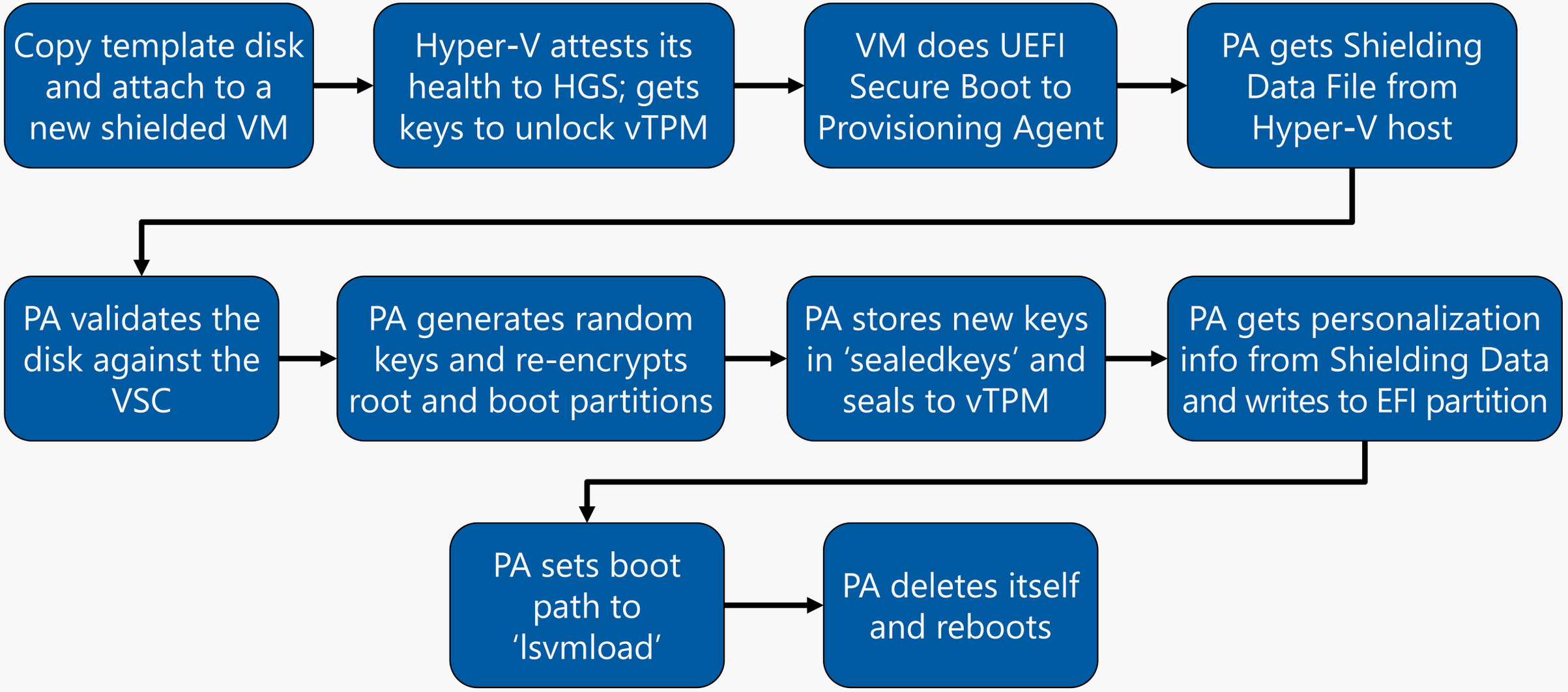
Set up VM root & boot partitions with unique dm-crypt encryption keys

Make per-VM settings and secrets available to the VM

Set up everything for normal boot process

Don’t let hypervisor admin have access to the encryption keys or per-VM secrets

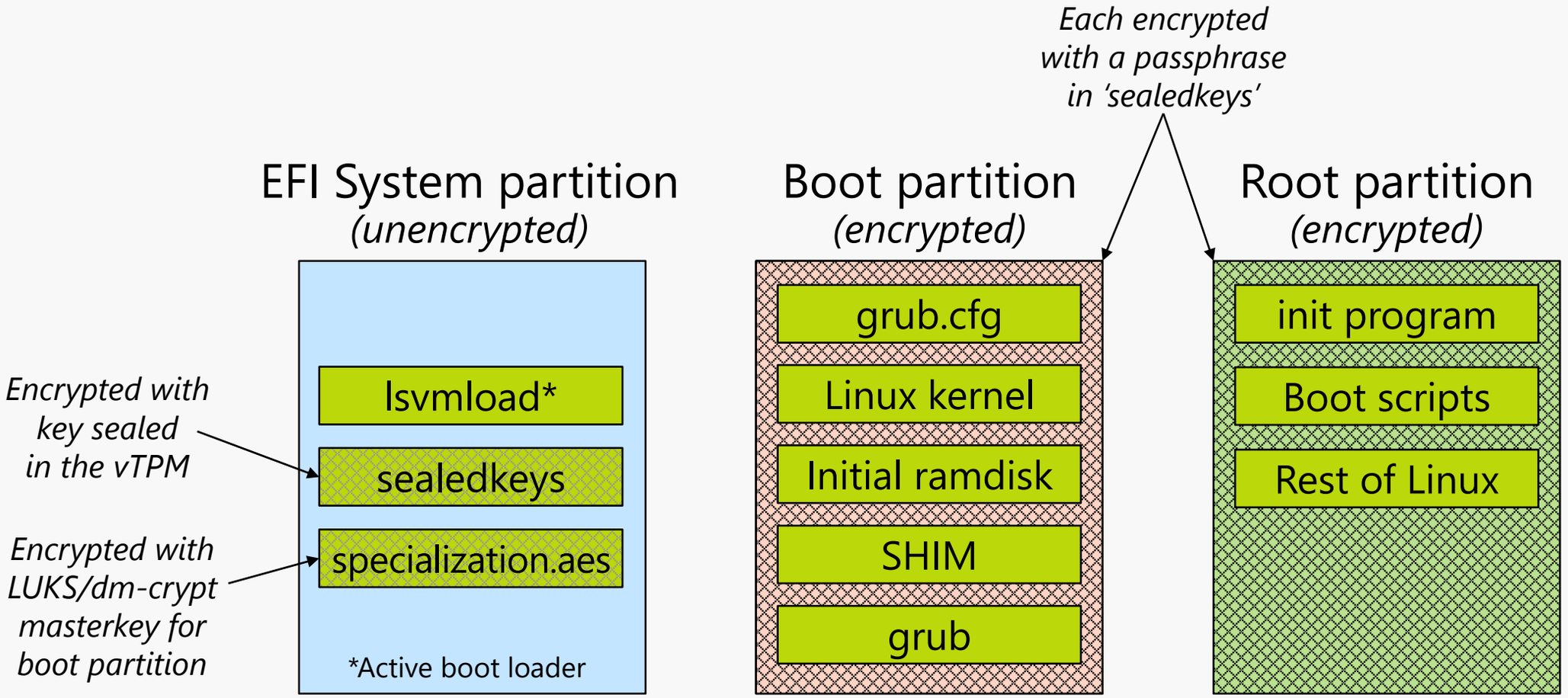
Drilldown: Provisioning (2 of 3)



Drilldown: Provisioning (3 of 3)

VM disk layout after provisioning is complete

Ready to do normal Linux boot



Drilldown: Normal Boot Path (1 of 3)

Key Goals

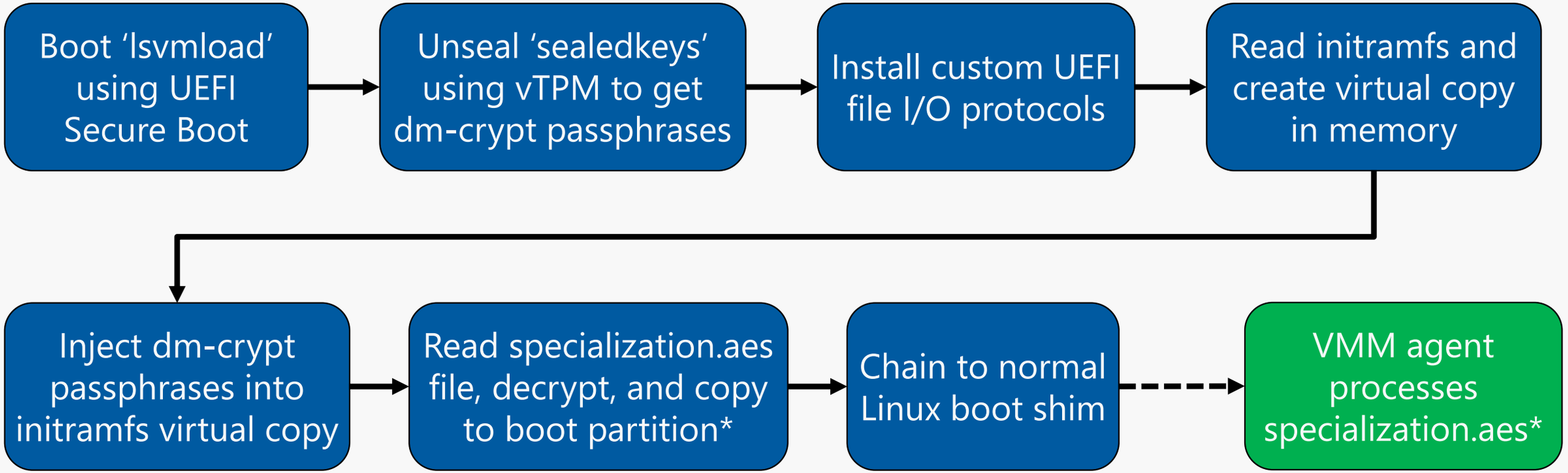
Prevent tampering with boot path that could expose secrets

Automate supplying dm-crypt passphrase from the vTPM

Don't modify the code in the normal Linux boot path (shim, grub, kernel)

What is modified	What is not modified
• initramfs updated to get dm-crypt passphrase from a file • lsmvload used as a precursor to the normal Linux boot shim • lsmvload inject disk passphrases as a file into virtualized copy of initramfs	• Linux shim • grub • Linux kernel

Drilldown: Normal Boot Path (2 of 3)

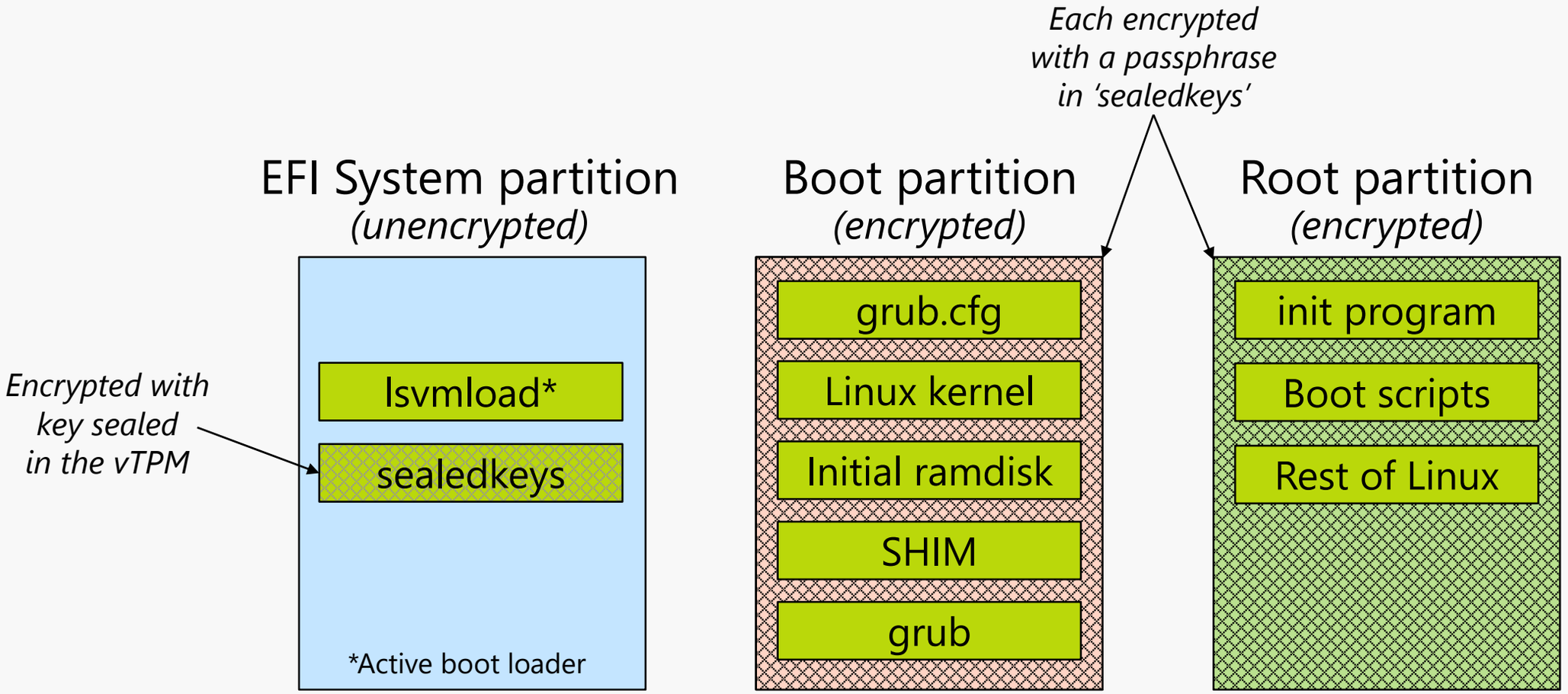


- I/O to encrypted boot partition is mediated by custom UEFI file I/O protocols
- initramfs gets dm-crypt passphrases from injected file

*First boot only

Drilldown: Normal Boot Path (3 of 3)

VM disk layout for ongoing operation



Drilldown: Additional Scenarios

Live migration

Can migrate a Linux VM to other Hyper-V hosts in same Guarded Fabric
Disks and VM memory are encrypted in transit
vTPM moves with the VM

Linux updates

Kernel and initramfs updates work normally – no special actions needed
grubx64.efi and shimx64.efi updates must be manually recopied to new location
lsvmload updates work normally – no special actions needed

Non-root disks

Managing the encryption of additional disks is outside the scope of Shielded VMs
Use normal dm-crypt techniques

Wrap-Up

Current Status

Hyper-V is updated to enable Linux Shielded VMs

Updated Hyper-V release ships later this fall

Works with Linux distros that can UEFI Secure Boot

We're collaborating with commercial distro vendors to ensure official support
Red Hat, SUSE, and Canonical

Windows Shielded VMs commercially available today

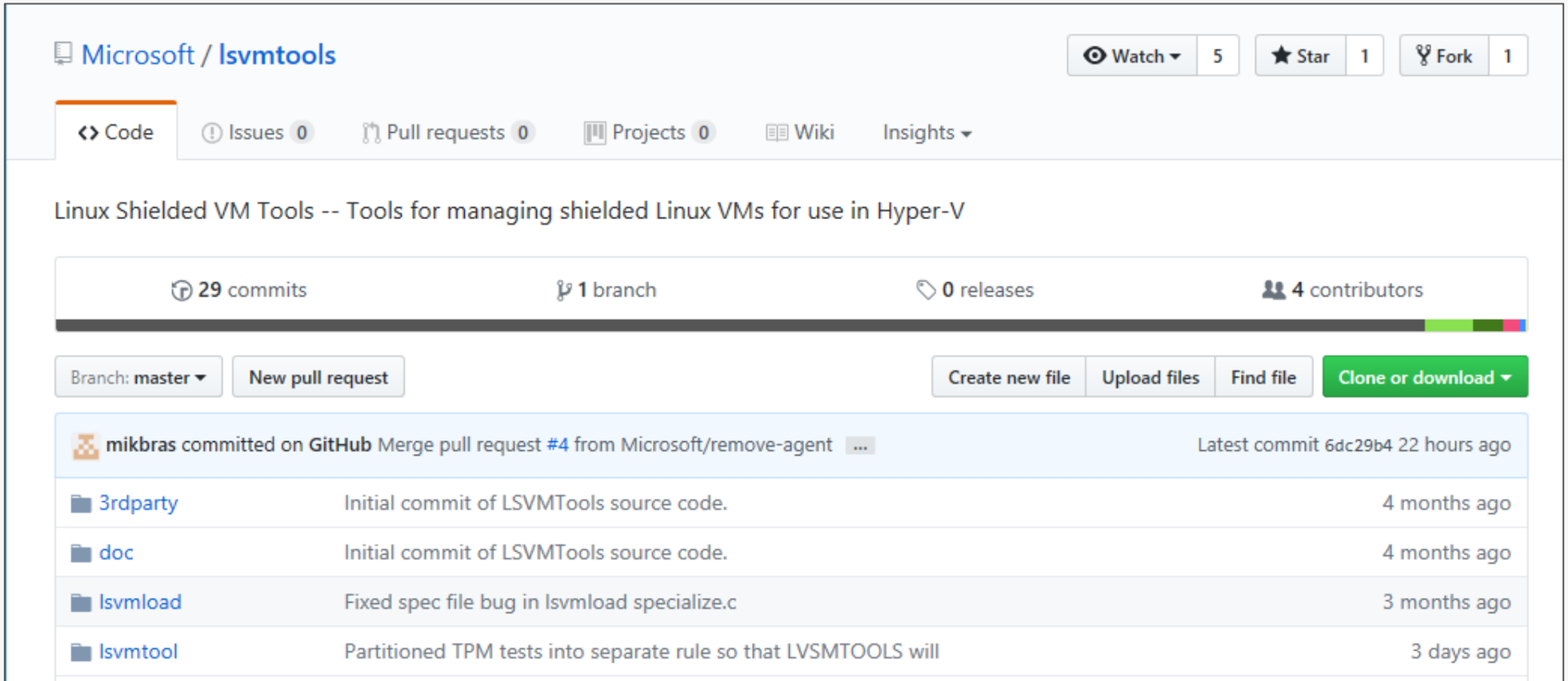
Rackspace, brightsolid (UK)
Expect to see Linux Shielded VMs commercially available after Hyper-V update ships

Open Source

github.com/Microsoft/lsvmtools

lsvmload – precursor boot loader

lsvmprep – script and tools for creating a Linux template disk



Wrap-Up

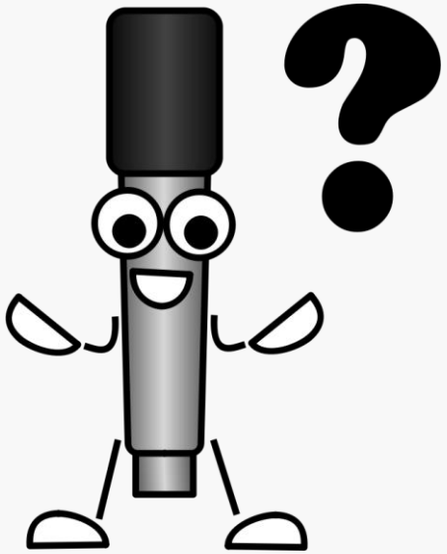
Linux Shielded VMs are working end-to-end

Great input from the community over the past year

Made significant changes based on that feedback

Ongoing feedback is always welcome

Questions?



Microsoft