Mt Wilson

**Recovery from Trust-Related Launch Failure**

# Requirements

## Automatic migration of VM after trust-related VM launch failure in OpenStack

The OpenStack scheduler with trust filter attempts to schedule VMs with a trust policy to launch on trusted hosts. The scheduler is not a trusted component itself and does not guarantee that VMs land on trusted hosts. Furthermore, an administrator can bypass the scheduler and launch a VM on a specific host.

# Analysis

## No trust-related launch failures with “measure only” policy

When the VM trust policy indicates “measure only”, the VM will launch successfully even if it does not match its trust policy. The VM launch will not fail due to trust when using this policy.

## VM trust failure with “measure and enforce” policy

When the VM trust policy indicates “measure and enforce”, the VM will launch successfully only if it matches its trust policy. The host must be trusted to enforce the policy. When the host is trusted and enforces the trust policy, the enforcement point creates a potential for a trust-related VM launch failure called a “VM trust failure”. When a VM trust failure happens, moving the VM to another host would not help because it is the VM itself that does not match its own trust policy.

The recovery plan in this scenario is to shut down the non-compliant VM and launch it again from its original image. If the original image does not match the policy, the VM is not recoverable, and either the image or its associated trust policy must be corrected to proceed.

## Network and trust-based key retrieval failures with “encrypted” policy

When the VM is encrypted, the customer can enforce that the host is trusted to enforce the VM policy because the VM decryption key will only be released to a trusted host. When the host is trusted, it receives a key to decrypt the VM. The key retrieval process creates a potential for a network failure called a “network key retrieval failure” and for a trust-related VM launch failure called a “trust-based key retrieval failure”.

When a network key retrieval failure happens, it’s because the key server was not accessible. Automatic recovery from this scenario may be possible by periodically retrying the request until the key server or the network issue is fixed.

When a trust-based key retrieval failure happens, it’s because the host was not trusted to receive the key, or because the host was trusted to receive the key in the past and cached it but is no longer trusted to access the key and was denied by the TPM. Automatic recovery from this scenario is possible by selecting a different trusted host on which to launch the VM image. After a trusted host receives the decryption key, it attempts to decrypt the VM. This process creates a potential for a failure called a “VM image decryption failure”.

When the host fails to decrypt the VM image with the available decryption key, if the error is determined to be a mismatch between the VM encryption key and the key specified in the policy, automated recovery is not possible and the customer must either re-encrypt the VM image with the correct key and re-launch or update the trust policy to reference the correct key and re-launch. Otherwise, if the error is because the system is out of disk space, the VM can be automatically launched on a server with greater capacity.

# Architecture

## Reporting errors in OpenStack

OpenStack has a message queue which is used to communicate errors from the compute nodes to the controller.

### VM trust failure

A trusted host must report via the message queue a “VM trust failure” condition to associate with a VM instance in the database.

The orchestrator must detect the “VM trust failure” condition in the VM instance data when handling errors. If the VM instance was a new launch, the problem is either in the original image or the associated trust policy and automated recovery is not possible - the orchestrator must report an error as it would normally do. If the VM instance was a migration, the orchestrator can assume the condition may have been caused by changes to the VM prior to or during migration, and it may choose to either re-launch a new instance on the destination host or to report an error as it would normally do.

### Network key retrieval failure

A trusted host must report via the message queue a “network key retrieval failure” condition to associate with a VM instance in the database.

The orchestrator must detect the “network key retrieval failure” condition in the VM instance data when handling errors. The orchestrator may automatically attempt to re-launch the VM on the same host after a period of time. The number of times to try and the delay between attempts must be configurable by the OpenStack administrator.

### VM image decryption failure

A trusted host must report via the message queue a “VM image decryption failure” condition to associate with a VM instance in the database.

The orchestrator must detect the “VM image decryption failure” and determine if the host has enough disk space to decrypt the image. If the host does not have enough disk space, the orchestrator must attempt to launch the VM on a different host. If the host does have enough disk space, the error may be caused by an incorrect key and is not automatically recoverable.

# Implementation

## OpenStack Integration

Policy agent in the compute node must report “VM trust failure” condition via OpenStack message queue.

Orchestrator must be patched to include recovery logic for “VM trust failure”, “network key retrieval failure”, and “VM image decryption failure”. The patch can be included in the Mt Wilson OpenStack Controller Extensions package.

The Trust Director must include a new setting in the trust policy to allow the user to specify whether they need an automatic recovery from a “VM trust failure” condition or not. The default should be “no” to conserve resources because a migration and re-launch are not equivalent operations from the end-user perspective and may not be what end-users want to do.

Scheduler must be patched to read the recovery setting from the trust policy and automatically set it as a VM instance attribute for the scheduler to use later.