Architected for Performance

PCI Express® (PCIe®) Infrastructure for Live Migration

Sponsored by NVM Express™ organization, the owner of NVMe® Family of Specifications
Speaker

Mike Allison

SAMSUNG
Agenda

• Benefits
• High Level Operational View
• Other Items of Consideration
Benefits

• NVM Express is adding capabilities to allow host to manage the migrating VM from one NVM subsystem to a different NVM subsystem by supporting the migration of the controller being used by the VM which includes the attached namespaces and the controller state.

• Pre-Copy Phase Host Actions
  • Requests the controller to track LBA changes (dirty LBAs) of the attached namespaces
  • Migrate the allocated LBAs of the attached namespaces
  • Migrate the dirty LBAs
  • Host may use a new mechanism to throttle commands processing by migrating controller to slow down changes

• Stop-and-Copy Phase Host Actions
  • Requests the controller to pause causing all fetched commands to be completed
  • Migrate any remaining dirty LBAs

• Post-Copy Phase
  • Migrate controller state
  • Resume the migrated controller
Benefits

• Investigating allowing the tracking the memory changes due to the migrating controller processing NVM Express® (NVMe®) commands

• Pre-Copy Phase Host Actions
  • Request the controller to track these memory changes (dirty memory)
  • Migrate the dirty memory

• Stop-and-Copy Phase
  • Migrate remaining dirty memory
Building the Pieces

• TP4165 Tracking LBA Allocation with Granularity
  • Reporting of allocated LBAs within a namespace for migrating a namespace
  • Usable in Snapshot use cases
• TP4159 PCIe® Infrastructure for Live Migration
  • Developing the theory of operation
• A TPAR to:
  • Support limit the BW and IOPS of a controller to allow slowing down of command processing on a migrating controller
Pre-Copy Phase Start

- Source Admin Host initiates a migration of a controller by requesting to log LBA changes (dirty LBAs)
- A Migration Queue is established
Pre-Copy Phase Start

- Source Admin Host initiates a migration of a controller by requesting to log LBA changes (dirty LBAs)
- A Migration Queue is established
- The memory associated with the migrating VM can be moved anytime by the Source Admin Host
Pre-Copy Phase – Initial Namespace Migration

Source Admin Host issues Get LBA status command to obtain the allocated LBAs.

Source Host

VM

Get LBA Status

Controller Y

Controller X

NVM Subsystem

Migration Queue

Namespace Allocation Map

Namespace

Target Host

VM

Get LBA Status

Controller G

Controller H

NVM Subsystem

Namespace
Source Admin Host issues Get LBA status command to obtain the allocated LBAs
• Controller returns a list of descriptors. Each descriptor indicates an LBA range
Source Admin Host issues Get LBA status command to obtain the allocated LBAs:

- Controller returns a list of descriptors. Each descriptor indicates an LBA range.
- The Source Admin Host uses these LBA ranges to issue read commands to copy the allocated LBAs to the destination.
NVMe® commands that cause LBA changes to the namespace are logged in the Migration Queue

- Write commands
- LBA deallocation due to the Dataset Management command
Stop-and-Copy Phase – Pause Migrating Controller

After copying the allocated LBAs to the destination, the Source Admin Host may migrate the dirty LBAs.
Stop-and-Copy Phase – Pause Migrating Controller

At some point the Source Admin Host pauses the VM
Issues a command to Pause the migrating controller to have the controller:

- Stop fetching commands
- Complete all previously fetched commands
Stop-and-Copy Phase – Finish Migrating

Source Host
- Completes migration of VM
Stop-and-Copy Phase – Finish Migrating

Source Host

- Completes migration of VM
- Completes Migration of namespace dirty LBAs
Post-copy Phase – Migrate Controller State

Source Admin Host

- Issuing command to get the migrating controller state and put that state into the destination controller
Post-copy Phase – Resuming Migrated Controller

Source Host

- VM
- VM Manager
- Controller X
- Controller Y
- Namespace
- NVM Subsystem

Target Host

- VM
- VM Manager
- Controller G
- Controller H
- Namespace
- NVM Subsystem

Target Admin Host

- Resume VM
- Issues a command to resume controller that was migrated
Post-copy Phase – Resuming Migrated Controller

Target Admin Host
- Resume VM
- Issues a command to resume controller that was migrated

Source Admin Host
- Remove VM
- Reset the migrated controller
Questions?