SwapFile handoff in FSR

- by Anupama Chandwani

Swap file transfer modes

- FSR – same host
  - Shared: Close on source and open on destination. Only metadata transfer.
  - Unshared: Create and copy swap file on destination. Transfer metadata, reservations and pDir during switchover.

FSRInfo

- Swap specific variables

  - hasSwapCopy;

  - swapCopyWorldID; to destroy mirror node from FSRTransferSwapFile()

  - swapCopyState; used by FSRSwapCopyHelper

    FSR_SWAP_COPY_STARTING,
    FSR_SWAP_COPY_IN_PROGRESS,
    FSR_SWAP_COPY_DONE,
    FSR_SWAP_COPY_SLEEPING,
    FSR_SWAP_COPY_STUNNING,
    FSR_SWAP_COPY_FAILED,
    FSR_SWAP_COPY_EXITED.

  - resumeState; Atomic variable to decide the Victor

    FSR_RESUME_NONE,
    FSR_RESUME_SOURCE,
    FSR_RESUME_DEST,
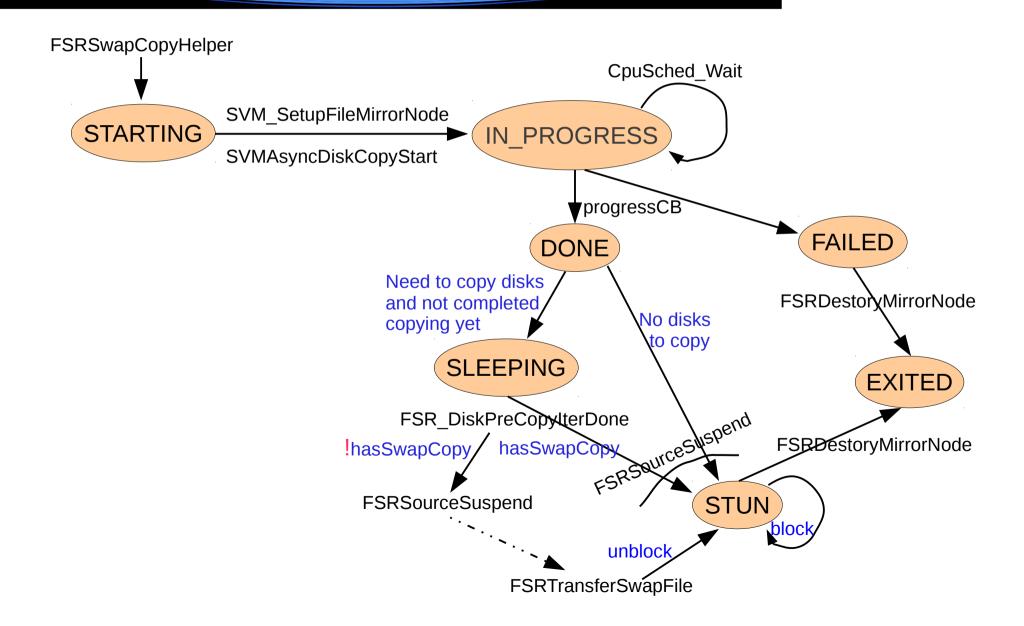   FSR_RESUME_DESTVERIFY.

Copy unshared swap file

- Since MN.next, uses svmmirror to copy unshared swap file.

FSR_InitMigration()

FSRSwapCopyHelper() *assistant world*

FSRSetupSwapMirror() *manually calls SVM_MakeDev()*

SVMAsyncDiskCopyStart() *with it's own findExtentCB*

FSRSwapCopy statemachine

Checkpoint functions

- FSR_CheckpointStarted()

- FSR_WriteCptData()

- FSR_ReadCptData()

- FSR_CheckpointFinished()

  Refer Ricardo's checkpoint slides

RestorePhase

- FSR_CheckpointFinished on source

    source: MIGRATE_PAGEIN

    dest: MIGRATE_CPTLOAD

- FSR_CheckProgress on destination

    FSRTransferMiscState

    1. transfer DVSState, DVSFilter, loadHistory.

    2. Mark all pages as remote & copy src lpool pointers to dest and share them. See VmMemMigrate_TransferMemory for actual mem transfer

    Set resumeState; *to decide the Victor*

    dest: MIGRATE_PAGEIN

FSR_ResumeDone

- Invoked by monitor at the tail end of RESTORE_SYNC for final handshakes & cleanups.

    FSRPrepareForMemoryTransfer()

    VmMemMigrate_PrepareForMemoryTransfer(): *disable swapping on source and waits for all active swapouts to complete*

    FSRTransferSwapFile()

    if hasSwapCopy: wakeup FSRSwapCopyHelper to destroy mirror node

    FSRTransferSwapState(); *details in next slide*

    else: Close swap file on source

    FSRTransferSwapState() then open swap file on dest

    Swap_EnableExclusive() on destination

    FSRTransferMemoryAndReservations

    FSRTransferReservations()

    VmMemMigrate_TransferMemory(); *details in next slide*

FSRTransferSwapState

- Compare swap files in terms of pages and not bitmap length.

- Support dynamically growing swap files.

- SwapExtend_Files() to adjust to source swap file size

- SwapMigrate_SetSwapBitmap() to copy source bitmap to destination bitmap.

VmMemMigrate_TransferMemory

VmMemMigrateMemoryHandoffStarted: *Sets MemoryHandoff Bools and ensures all PCPUs have read this value.*

VmMemMigrateTransferMemory: *Lock each pDir and copy all pages within a pDir while updating the accounting info*

VmMemZip_TransferMemZipInfo: *Compression cache info header*

Mark all pages as unmapped and transfer ownership of large page pool

VmMemMigrateMemoryHandoffFinished: *Clears the MemoryHandoff Bools*