Agenda

What we will talk about today:

1. Partition Mobility
2. What implications are there on IBM i
3. Allowing or preventing a Live Partition Migration on IBM i
4. Post Live Partition Migration considerations on IBM i
5. Implications for PASE based AIX applications
Partition Mobility on IBM i

- With IBM v7.1 TR4 IBM i started to support the capability to migrate partitions between two systems, referred to as the “Source System” and the “Destination System”

- IBM i partitions can be migrated providing they are in any of the following states:
  
  - **Active Partition Mobility**: Active Partition Mobility is the actual movement of a running partition from one physical machine to another without disrupting the operation of the operating system and applications running in that partition.

  - **Inactive Partition Mobility**: Inactive Partition Mobility transfers a partition that is “powered off” (not running) from one system to another.

  - **Suspended Partition Mobility**: A partition can be suspended and later resumed. Suspended Partition Mobility transfers a partition that is suspended from one system to another. It may then be resumed on the target system at a later time.
Partition Mobility Applicability / Benefits

- With IT organisation being required to provide improved service levels to their business with reduced planned outages, IT security and resiliency are as more critical than ever before

- Resource balancing
  - e.g. move workload onto a system that has lighter workloads. Useful for performance management and energy management

- Reduce planned CEC outages for maintenance/upgrades
  - Migrate partitions between systems to provide continued availability of an IBM i partition users

- Impending CEC outages
  - As an option to keep a partition running if hardware warnings are received
Partition Mobility for IBM i - Minimum Requirements

- IBM 7.1 TR4 or above
- Power VM Enterprise Edition on both Source and Destination Systems
- Compatible Power 7 hardware
- POWER 7 Firmware 740.40 or 730.51 or above
- HMC v7 r7.5.0 or above
- VIOS 2.2 1.4 or above on both systems
Partition Mobility for IBM i - Minimum Setup Requirements

- All resources for the partition must not have physical I/O adapters assigned

- The Source and Destination systems must have
  - the same LMB (Logical Memory Block) size
  - VIO Servers must have visibility to the disks and must be set a reserve policy of no_reserve
  - VIO Servers providing Mover capabilities
  - Time of Day should be similar on each system (could use dedicated VIOS to sync clocks)
  - The same VLAN's defined and available

- Destination System:
  - No partition with the same name as the one which is to be Migrated
  - Cannot be running on battery power
  - Must have sufficient resources (e.g. CPU & Memory) available

- Dual Virtual I/O Server and multipath I/O
  - only if the destination system is configured with two Virtual I/O Servers that can provide the same multipath setup

*Important Planning Note:* All user-defined virtual devices in VIOServer must have a Virtual Slot number higher than 10. This will be enforced by the HMC code.
Partition Mobility – Configuration Requirements

- The logical partition must have all disks backed by physical volumes.
- The logical partition must not be assigned a virtual SCSI optical or tape device.
- The logical partition cannot be activated with a partition profile which has a virtual SCSI server adapter.
- The logical partition cannot be activated with a partition profile which has a virtual SCSI client adapter that is hosted by another IBM i logical partition.
- No virtual SCSI server adapters can be dynamically added to the logical partition.
- No virtual SCSI client adapters that are hosted by another IBM i logical partition can be dynamically added to the logical partition being moved.
- An IBM i logical partition cannot be moved if it has a varied on NPIV attached tape device.
- For NPIV, you must zone both Worldwide Port Names (WWPNs) associated with a virtual fibre channel adapter.
- The logical partition must not be an alternative error logging partition.
- The logical partition cannot collect physical I/O statistics.
- The logical partition must not be an alternative error logging partition.
- The logical partition must not have a Barrier Synchronization Register (BSR).
- The logical partition must not have huge memory pages.
Partition Mobility – Setting a new partition to support mobility

**User enabled the setting at create partition**

- At create partition,
  - Checkbox in GUI and optional in parameter in CLI termed 'RestrictedIO partition'.
  - Defaults to NOT support partition mobility
  - No restrictions/changes in create/add profile and modify profile.
Partition Mobility – Changing an existing partition to support mobility

User enabled of the setting on existing partition

- On in-active partitions only
  - HMC validates if the partition is capable of being a “RestrictedIO Partition” when the profile update is attempted to be saved.
  - Validation will fail if this profile does not meet the configuration requirements
Partition Mobility

Partition

Virtual Disk
Mapped through VIO Server to a LUN

Virtual Ethernet
Mapped through SEA in VIO Server

Hypervisor

vSCSI
Virtual Ethernet
Support for migration

VIO Server

Provides the Virtual I/O
VASI interface to Hypervisor
Mover Service Partition (MSP)

HMC

Configuration of required capabilities
Validation of configuration
Orchestrates the sequence of events

VASI = Virtual Asynchronous Services Interface
Partition Mobility – Summary

Server A
- vhost0
- vsctxsi0
- fcs0
- hdisk1
- en1
- en0
- en2
- SEA

Service Processor
- VASI
- MSP

Server B
- vhost0
- vsctxsi0
- fcs0
- hdisk1
- en1
- en0
- en2
- SEA

Service Processor
- VASI
- MSP

HMC

Ethernet Network

Storage Area Network

LUN
Partition Mobility – Validating

- A new set of tasks are available for partitions that support mobility, as we can see below the task allow a user to migrate a partition or validate partitions readiness to be migrated.
Partition Mobility – Validating

- The Validation dialogue, requires you to select the Destination System then simply click on the 'Validate' button. This will validate both systems to ensure that a partition can be migrated between systems.
Partition Mobility – validation (HMC)

**HMC high level steps**

*Explicit HMC user interfaces are available for validation only.*

- Validation is also part of actual migration operation.

- Checks the Remote Monitoring and Control (RMC) connections to both VIO Servers
- Checks the RMC connection to the partition to be Migrated
- Checks the LMB sizes on Both Systems
- Checks the Partition to be Migrated:
  - No physical adapters defined as “required” in the LPAR
  - The LPAR uses only external LUNs
  - The LPAR supports active migration (OS support)
  - The LPAR is NOT a Mover Service Partition
  - The LPAR is NOT using Barrier Synchronisation Registers
  - The LPAR is NOT using Huge Pages
  - The LPAR state is active/running
  - The LPAR is NOT in a Partition workload Group
  - The LPAR MAC address is unique (across both servers)
  - The LPAR has a name which is NOT in use on the Target System

- Checks that we do not exceed the number of active migrations
Partition Mobility – Validation RMC connections
Partition Mobility – Validation
Partition Mobility – Validation
Partition Mobility – Validation

Diagram showing the connection between Server A and Server B through Ethernet and Storage Area Network. Disk units, Ethernet, vhost0, vhost0, fcs0, vtscsi0, en1, en2, SEA, en0, hdisk1, and VASI MSP are depicted on the diagram.
Partition Mobility – Validation

- Checks the target system has the required resources
  - Processor, Memory, Virtual Slots
  - SEA with matching VLAN(s) to support the migrated partition
  - Same disk UDID (universal Device ID) available
Partition Mobility – Validation

- Disk units
- Server A: vhost0, hdisk1, en1, en2, SEA
- Server B: vhost0, hdisk1, en1, en2, SEA
- Operating System and application readiness
- HMC
- Ethernet Network
- Storage Area Network
- LUN
Partition Mobility – Validation

- State Information
  - Processor configuration – dedicated/shared, count and entitlement
  - Memory configuration - Minimum/maximum/desired
  - Virtual Adapter configuration
- State Information from the Source System Hypervisor
  - Partition’s memory, hardware page table (HPT), Processor state
  - Virtual adapter state, non-volatile RAM (NVRAM), The time of day (ToD)
  - Partition configuration, state of each resource
- Source system MSP collect this through the VASI
- MSP transfers this information to the target MSP
- State information is then available to the new partition shell

Note: The State Information represents the LPAR’s current characteristics (it is NOT based on any of the LPAR profiles). The existing profile will be modified with the new Virtual Device mappings.
Partition Mobility – Validation

- After the validation has completed, the dialog information is updated to include information on how the migration would occur.
Partition Mobility – Migration

- After you have validated you can have the option to begin a migration process
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively, you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration

- Alternatively you can select a migration from the partition tasks and you will be taken through a set of screens to set up the migration parameters.
Partition Mobility – Migration
Partition Mobility – Migration

Server A
- Disk units
- vscsi
- ethernet
- vhost0
- hdisk1
- fcs0
- en1
- en2
- SEA

Server B
- Disk units
- vscsi
- ethernet
- vhost0
- hdisk1
- fcs0
- en1
- en2
- SEA

HMC
- VASI
- MSP

Ethernet Network
- MSP Setup

Storage Area Network
- LUN
Partition Mobility – Migration

Disk units

Server A

vscsi
ethernet

en1
en0
fcs0
hdisk1
vtscsi0
vhost0

Server B

Disk units

ethernet
vscsi

en1
en0
fcs0
hdisk1
vtscsi0
vhost0

Logical Memory Copy

Service Processor

Service Processor

VASI
MSP

VASI
MSP

en2
en2
en1
en1

Logical Memory Copy

Memory Copy over network

Storage Area Network

LUN
Partition Mobility – Migration
Partition Mobility – Migration
Partition Mobility for IBM i

- Clicking on the 'Migrate' button starts the migration process, on the HMC you will be given a progress indication of how the migration is progressing.
Partition Mobility for IBM i

- At the end of the migration process the progress dialog will be updated to a status of success
  - During the migration process we have seen some minor delays in ping response times of between 50ms and 1.5 seconds during the final stage.
Partition Mobility – Configuration Gotcha’s

- Configuration, management and usage for partition migration is common across both IBM i and AIX, however there is some information missing in the current online documentation

- All disk units must be visible to both systems and have the reserve policy set to no_reserve, you will receive the following error when you attempt to migrate the partition:

  HSCLA27C The operation to get the physical device location for adapter U9117.MMB.102709P-V20-C6 on the virtual I/O server partition VIOS1-Dilling has failed. The partition command is: mgmgr -f get_adapter -t vscsi -s U9117.MMB.102709P-V20-C6 -d 1 The partition standard error is: child process returned error
  HSCLA9C1 The partition suspend or resume operation has stopped unexpectedly. If necessary, perform a suspend or resume recovery operation for the partition.

- VLAN id’s must be common to both systems otherwise the validation phase will fail

- If using Dual VIOS for multipathing this must be available on the Destination system
Agenda

What we will talk about today:

1. Partition Mobility
2. What implications are there on IBM i
3. Allowing or preventing a live Partition migration on IBM i
4. Post Live Partition Migration considerations on IBM i
5. Implications for PASE based AIX applications
Partition Mobility and Power-HA

- Currently Partition Mobility and Power HA are independent capabilities and are unaware of each other, this means:
  - You need to ensure that your heartbeat intervals for failure detection will not be impacted by a live partition mobility operation
    - When you move an live partition between systems we have seen a minor delay in TCP/IP response times at the moment of the workload becomes active on second system
    - You need to ensure that your failure detection is not set so sensitively that this will cause your environment to believe it is had a failure on the source system
Live Partition Mobility and TCP/IP (or APPC) communications

- As we discussed previously we have seen a minor delay in response times during the final stage of the mobility process these should not cause any issues with applications dependant on TCP/IP or APPC. However if you use very low timeout values you should ensure they will not impacted by a mobility activity

- As with Software likely to be impacted:
  - MQ: Typically there will be no impact
  - WebSphere Application Server on IBM i: Typically there will be no impact
  - WebSphere Application Server connecting to an IBM i during mobility: Typically there will be no impact, however capabilities with timeout values, such as JDBC, may need to be revalidated
  - Sockets programs: Typically there will be no impact, you may wish wish to check you time-out values within your applications
  - APPC programs: Typically there will be no impact, you may wish wish to check you time-out values within your applications
  - FTP: Typically there will be no impact
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which is matches what we see when using ping.

- If you use interactive IBM i PASE within your environment then when thinking about Live Partition Mobility we consider this to be a 5250 session and the same behaviour is seen as described previously.

Note: Our investigation was undertaken with IBM Personal Communications as the terminal emulator, the findings may differ when utilizing other manufactures emulators.
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which is matches what we see when using ping.
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which is matches what we see when using ping.
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which matches what we see when using ping.
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which is matches what we see when using ping.
Live partition Mobility and Personal Communications 5250 sessions

- There should be no impact when use live partition mobility to move a partition that has active 5250 sessions. We have seen minor delays in 5250 response times of between 50ms and 1.5 seconds during the final stage which is matches what we see when using ping.
What we will talk about today:

1. Partition Mobility
2. What implications are there on IBM i
3. Allowing or preventing a Live Partition Migration on IBM i
4. Post Live Partition migration considerations on IBM i
5. Implications for PASE based AIX applications
Allowing or preventing a Live Partition Migration on IBM i

- During a partition migration, or a standalone partition migration validation, from the hardware management console (HMC) there is a communication between the hypervisor and the IBM i partition to ensure that partition is in a state that it can be migrated.

- What does this mean? In a state that can be migrated? What are you on about…
  
  - All systems at some point will be undertaking some business critical function that should not, or can not, delayed at that point in time. If we were to migrate the partition at this time then there would be a dire set of business consequences that we wish to avoid at all times.

- The partition migration validation utilizes a capability on the IBM i partition that can be programmatically exploited to ensure that the business critical functionality will be completed in a time window, has been quiesced or can block the migration of the IBM i partition.
Allowing or preventing a Live Partition Migration on IBM i
Allowing or preventing a Live Partition Migration on IBM i

- This is capability is surfaced on the IBM i by a Work Management exit added with PTF SI42815 called QIBM_QWC_SUSPEND
- There is no restriction on the number of programs register against this exit point
Allowing or preventing a Live Partition Migration on IBM i

- This simplified representation of the interaction between the HMC and the IBM i exit point during a partition migration shows that the exit program will get called a minimum of two times (Validate and Prepare) and potential a third time if the migration is cancelled from the HMC during the process.
Allowing or preventing a Live Partition Migration on IBM i

- When the exit point program is called during a validate or check phase it has the option to allow or deny the migration request as we can see here
Allowing or preventing a Live Partition Migration on IBM i

- When the exit point program is called due to a standalone validation request or as part of the validation phase of a migration request the data passed to the program will have a value of 1 in the phase parameter and will receive exit information in the format of BMGF0100

**BMGF0100 Format**

The following table describes the information that is located in the Exit information parameter for the BMGF0100 format. For detailed descriptions of the fields, see Field Descriptions.

<table>
<thead>
<tr>
<th>Offset</th>
<th>Dec</th>
<th>Hex</th>
<th>Type</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>BINARY(4)</td>
<td>Exit information length</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>BINARY(4)</td>
<td>Status flag</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>BINARY(4)</td>
<td>Wait time</td>
</tr>
<tr>
<td>12</td>
<td>C</td>
<td>C</td>
<td>BINARY(4)</td>
<td>Action</td>
</tr>
</tbody>
</table>

- To block the request the exit program must return zero for the value of the status flags
- To allow the request to proceed the exit program should return a one for the value of the status flag and a value between 1 and 900 representing the number of seconds that will be required by the system to prepare itself for migration
- Note: If you specify an invalid value for the status parameter the migration will proceed and if you specify an invalid value for the wait time the it will be ignored and the wait time will be 1 second
Allowing or preventing a Live Partition Migration on IBM i

- When the exit point program is called to allow the system to prepare itself for a migration, it should ensure that any environment or application changes are made before the migration starts.
Allowing or preventing a Live Partition Migration on IBM i

- When the exit point program is called during the prepare phase (the migration is about to start) the data passed to the program will have a value of 2 in the phase parameter and will receive exit information in the format of BMGF0200

**BMGF0200 Format**

The following table describes the information that is located in the Exit information parameter for the BMGF0200 format. For detailed descriptions of the fields, see Field Descriptions.

<table>
<thead>
<tr>
<th>Offset</th>
<th>Hex</th>
<th>Type</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>BINARY(4)</td>
<td>Exit information length</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>BINARY(4)</td>
<td>Prepare time allowed</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>BINARY(4)</td>
<td>Action</td>
</tr>
</tbody>
</table>

- The value in the returned in the Prepare time allowed parameter will be the maximum of all wait time returned by the exit point programs registered for this exit point
- If the exit program does not complete within the wait time, terminates abnormally the migration will proceed regardless.
Allowing or preventing a Live Partition Migration on IBM i

- When the exit point program is called due to a migration being cancelled by the user, it is supposed to undertake any cleanup necessary to return the source system to its original state.

Diagram:
- **Check/validate**
  - Exit program allows migration request and specifies time to wait in prepare phase
- **Prepare**
  - Migration is cancelled
  - Exit program is notified to allow cleanup
- **Reject**
  - Exit program rejects migration request
- **Cancel**
Allowing or preventing a Live Partition Migration on IBM i

- At any time during an LPM process the migration may be cancelled from the HMC. If this were to occur then the exit program will be called again, in these situations will have a value of 2 in the phase parameter and will receive exit information in the format of BMGF0200

**BMGF0200 Format**

The following table describes the information that is located in the Exit information parameter for the BMGF0200 format. For detailed descriptions of the fields, see Field Descriptions.

<table>
<thead>
<tr>
<th>Offset</th>
<th>Dec</th>
<th>Hex</th>
<th>Type</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>BINARY(4)</td>
<td>Exit information length</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>BINARY(4)</td>
<td>Prepares time allowed</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>BINARY(4)</td>
<td>Action</td>
</tr>
</tbody>
</table>

- This is to allow the exit program to notify other applications that the system will not be migrated or undertake any 'clean-up' activity that may be required.
Allowing or preventing a Live Partition Migration on IBM i

- Whilst at first glance this seems like a pretty straightforward exit point, there are several things to be aware of:
  - The exit point calls a standalone program
  - This program runs in a standalone job

- What are the implications of this?
  - The prime implication is that it is responsibility of the exit point implementer to ensure that the system (or application's that are suspend/resume and mobility aware) are in a state to allow the partition to be suspended or migrated, however as the program runs in it’s own job there is a burden on the exit point developer to have some type of inter process communication, or interaction, with other jobs and processes on the system and to respond to back to the operating system in a timely manner

- This could potentially mean retro-fitting the required Suspend/Resume awareness into existing applications and environments
Allowing or preventing a Live Partition Migration on IBM i

- What mechanism are available to an Exit Point developer to do this, here are some with a discussion to follow:
  - Interrupt Job Exit - Invoking a Program in Another Job
  - Inter process communications via signals and semaphores
  - Inter process communications via queues
  - Inter process communications via locks
  - Inter process communications via data area/system area

- Ideally you would develop capabilities suitable to your organization or applications
Allowing or preventing a Live Partition Migration on IBM i

- Interrupt Job Exit
  - This allows you to invoke a program in another job, allowing you therefore to validate some job specific environment setting such as a QTEMP object or environment variable
  - To use this capability you need to know information about each job on the system you need to interrogate, this would need to be designed into your environment
  - You would also potentially be stopping a critical piece of functionality whilst this check occurs, which may be problematic for time sensitive communications activities
Allowing or preventing a Live Partition Migration on IBM i

- Inter-process communications (POSIX style semaphores and signals)
  - There are many options here

When some critical activity is occurring a semaphore could be create, the exit program checks for the semaphore and if found blocks the suspend or migration operation. If no such semaphore exists then the exit program creates one which indicates that the system is about to be suspended or migrated and no critical activity should occur. On resume another exit program could clear the semaphore

When the exit program starts execution it sends out a signal and waits for a response (semaphore or other IPC) to indicate whether there is critical system activity at that time

- Be aware of the implications of user activities

- Ideal you would implement this via a set of library functions in service programs

- Downside is awareness of POSIX capabilities on IBM i
Allowing or preventing a Live Partition Migration on IBM i

- Inter-process communications (Queues)
  - Same concepts as for the previous example, however using queues and non-destructive reads for the Inter-process communication between the environment and the exit program

- Inter-process communications (Locks)
  - Utilizing locking on objects to indicate that critical system activity is occurring or that a partition is being suspended or migrated

- Inter process communications via data area/system area
  - Utilize a global data area(s) or system area(s) in a common library to indicate critical system activity or being suspended or migrated

- These options are better understood by the broader IBM i community
- Ideally implemented via a set of functions with a service library
## Agenda

**What we will talk about today:**

1. **Partition Mobility**
2. **What implications are there on IBM i**
3. **Allowing or preventing a Live Partition Migration on IBM i**
4. **Post Live Partition Migration consideration on IBM i**
5. **Implications for PASE based AIX applications**
Post Live partition Migration considerations for an IBM i partition

- After a live partition has been migrated onto the destination system, there may be some cleanup or activity that needs to occur on the partition to bring the system up to an active state
  - Such as activating reconnecting to an NPIV tape library

- If this is required then it is possible to utilize a capability on the IBM i to programmatically achieve this
Post Live partition Migration considerations for an IBM i partition

- This is capability is surfaced on the IBM i by a Work Management exit added with PTF SI42815 called QIBM_QWC_RESUME
- There is no restriction on the number of programs register against this exit point
Post Live partition Migration considerations for an IBM i partition

- Programs registered against the QIBM_QWC_RESUME exit point need to utilise the RSMS0100 exit format as shown below. Note the phase parameter will always have a value of 4 to indicate that the migration has been completed.

**Resume System Exit Programs**

<table>
<thead>
<tr>
<th>Required Parameter Group:</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exit information</td>
<td>Input</td>
<td>Char(*)</td>
<td></td>
</tr>
<tr>
<td>Format of exit information</td>
<td>Input</td>
<td>Char(8)</td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Input</td>
<td>Char(1)</td>
<td></td>
</tr>
</tbody>
</table>

Exit Point Name: QIBM_QWC_RESUME
Exit Point Format Name: RSMS0100

The Resume System exit program is called when the system becomes available again after the partition was migrated to another machine or the partition resumed from hibernation. See [Format RSMS0100 Required Parameter Group](#) for parameter information or [RSMS0100 Format Usage Notes](#) for more information.

- Information on the formats used by the exit point can be found on the IBM i v7.1 Infocenter [here](#)
Post Live partition Migration considerations for an IBM i partition

- When the exit point program is called to indicate completion, the exit information will always be in the format of AMGF0100

**AMGF0100 Format**

The following table describes the information that is located in the Exit information parameter for the AMGF0100 format. For detailed descriptions of the fields, see Field Descriptions.

<table>
<thead>
<tr>
<th>Offset</th>
<th>Dec</th>
<th>Hex</th>
<th>Type</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>BINARY(4)</td>
<td>Exit information length</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>BINARY(4)</td>
<td>System state</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>BINARY(4)</td>
<td>Action</td>
</tr>
</tbody>
</table>

- For a migration the value of the Action attribute will always be a one
- The value of the system state attribute is used to indicate the state of the system (restricted or normal) and will be:
  - 0 The system has resumed during normal operation.
  - 1 The system has resumed while in restricted state.
Post Live partition Migration considerations for an IBM i partition

- Whilst at first glance this seems like a pretty straightforward exit point, there are several things to be aware of:
  - The exit point calls a standalone program
  - This program runs in a standalone job

- What are the implications of this?

- The prime implication is that it is responsibility of the exit point implementer to ensure that the system, or application's that are Suspend/Resume and Migration aware, that the systems is being resumed. However as the program runs in it’s own job there is a burden on the exit point developer to have some type of inter process communication, or interaction, with other jobs and processes on the system and to respond to back in a timely manner

- This could potentially mean retro-fitting the required Suspend/Resume and Migration awareness into existing applications and environments
Post Live partition Migration considerations for an IBM i partition

- What mechanism are available to an Exit Point developer to do this, here are some with a discussion to follow:
  - Interrupt Job Exit - Invoking a Program in Another Job
  - Inter process communications via signals and semaphores
  - Inter process communications via queues
  - Inter process communications via locks
  - Inter process communications via data area/system area

- Ideally you would develop capabilities suitable to your organisation or applications
Post Live partition Migration considerations for an IBM i partition

- **Interrupt Job Exit**
  - This allows you to invoke a program in another job, allowing you therefore to validate some job specific environment setting such as a QTEMP object or environment variable.
  - To use this capability you need to know information about each job on the system you need to interrogate, this would need to be designed into your environment.
  - You would also potentially be stopping a critical piece of functionality whilst this check occurs, which may be problematic for time sensitive communications activities.
Allowing or preventing a Live Partition Migration on IBM i

- Inter-process communications (POSIX style semaphores and signals)
  - There are many options here

When some critical activity is occurring a semaphore could be create, the exit program checks for the semaphore and if found blocks the suspend or migration operation. If no such semaphore exists then the exit program creates one which indicates that the system is about to be suspended or migrated and no critical activity should occur. On resume another exit program could clear the semaphore

When the exit program starts execution it sends out a signal and waits for a response (semaphore or other IPC) to indicate whether there is critical system activity at that time

- Be aware of the implications of user activities

- Ideal you would implement this via a set of library functions in service programs

- Downside is awareness of POSIX capabilities on IBM i
Allowing or preventing a Live Partition Migration on IBM i

- **Inter-process communications (Queues)**
  - Same concepts as for the previous example, however using queues and non-destructive reads for the Inter-process communication between the environment and the exit program

- **Inter-process communications (Locks)**
  - Utilizing locking on objects to indicate that critical system activity is occurring or that a partition is being suspended or migrated

- **Inter-process communications via data area/system area**
  - Utilize a global data area(s) or system area(s) in a common library to indicate critical system activity or being suspended or migrated

- These options are better understood by the broader IBM i community
- Ideally implemented via a set of functions with a service library
Agenda

What we will talk about today:

1. Partition Mobility
2. What implications are there on IBM i
3. Allowing or preventing an LPM request on IBM i
4. Post Live Partition migration considerations for IBM i
5. Implications for PASE based AIX applications
Implications for PASE based AIX DLPAR aware applications

- AIX uses a different mechanism to inform applications that there is a Suspend/Resume or Live Partition Mobility activity occurring for a partition.

- The AIX kernel sends a SIGRECONFIG signal out to all processes.

- The process can then use the `dr_reconfig()` api to gain insight into what is occurring within the partition and can use the same api to veto the activity.
Implications for PASE based AIX DLPAR aware applications

- The IBMi PASE environment has the signal and the API defined within the open system include files.
Implications for PASE based AIX DLPAR aware applications

- Running the same AIX LPAR aware application inside the IBMi PASE QP2TERM environment, we see that the IBMi does not send a SIGRECONFIG signal to listening AIX DLPAR aware applications.
Implications for PASE based AIX DLPAR aware applications

- What does this mean

- Any application vendor who has an existing IBMi PASE application or is in the process of porting an application which is DLPAR aware needs to recode their application(s)
- They need to utilize the existing IBMi OS/400 exit points as described in the previous section
- Multiple choices on how to achieve this:
  - Redesign the application to utilize the IBMi exit point
    - or
  - Custom IPC mechanism such as
    - Have an exit point program that sends out the SIGRECONFIG signal to the relevant applications (there may be additional work to have this list of interested applications available to the exit program)
    - Provide a dr_reconfig api wrapper that handles the investigatory and veto process via the external exit program
Further information

- IBM i exit programs
  - Suspend exit point
    http://publib.boulder.ibm.com/infocenter/iseries/v7r1m0/topic/apis/xsuspend.htm
  - Resume exit point
    http://publib.boulder.ibm.com/infocenter/iseries/v7r1m0/topic/apis/xresume.htm

- Preparing for partition mobility

- Redbooks
  - “IBM PowerVM Virtualization Introduction and Configuration” Redbook.
IBM i zone on developerWorks

- Resource for the IBM i community
- Wide variety of technical information

ibm.com/developerworks/ibmi