(complex) Remote Restart

- Originally RR needs Active Memory Sharing (AMS) memory & backing paging devices
- This was used to store the VIOS & LPAR details to allow a remotely rebuild LPAR/VM

- AMS use was not widely implemented
- This (complex) Remote Restart rarely used
Simplified Remote Restart

• Much simpler!!
  → VM details captured to the HMC

• Need to tell HMC to do this
  & reboot the VM 😞 sorry about that!

• Worth doing at VM (LPAR) creation time
  or set it for next reboot

As simple as 1-2-3

1. Get Live Partition Mobility working
   – If LPM does not work then RR never will!
2. Set Remote Restart - ASAP
3. LPM Validate + SRR Validate
4. Test it works
   – HMC CLI
   – PowerVC GUI

5. Prepare for the worse: pre-decide VM priority & targets
6. BANG!! Don’t Panic – do it for real!!! (unlikely)
Live Partition Mobility Reminder

- **Live Partition Mobility** = while running
  - Many years experience now
- **Static Partition Mobility** = while shutdown
  - Quick as no memory to move
- **Dead Partition Mobility** = from the grave!
  - Simple Remote Restart

Live Partition Mobility - Reminder

- Requires PowerVM Enterprise Edition
- Requires “spare” capacity on the target server
- Keep HMC & VIOS’s up to date
- Pure virtual network (SEA) & disks (vSCSI or vFC)
- Source + Target need same subnet & disks/LUN access
- Gotcha!!
  1. Virtual optical media from VIOS → can just delete it
  2. Logical Memory Block size → ASMI + Server reboot
  3. Processor Mode to older box (can’t move P8 VM to a P7 box)
  4. Linux on POWER OK but missing the IBM RPMs
Live Partition Mobility - Best Practice

- 10G to 10G dedicated Ethernet connection for LPM (if possible)
  1. Keep up to date using Fix Level Recommendation Tool (FLRT)
  2. LPM setup checklist for first time
       s1184.html
  3. LPM prep checklist if its been a while since LPM for a partition
       s1185.html
  4. Follow VIOS performance guidelines for LPM
     - http://www-01.ibm.com/support/knowledgecenter/9119-
       MHE/p8hc3/p8hc3_viostune_lpmperf.htm?cp=9119-MHE&lang=en

2 Set Simplified Remote Restart flag

- Assuming LPM is a “go”
2 Set Simplified Remote Restart flag

• Machine Properties Capable

2 Set Simplified Remote Restart flag

• Can’t set Remote Restart flag at the
  – VM / LPAR OS
  – HMC GUI LPAR profile
  – HMC Enhanced+ LPAR profile

• Current only via the HMC command line
2 Set Simplified Remote Restart flag

• Assuming LPM is a “go” ✓

• Machine Properties Capable is “go” ✓

• Access to the HMC command line is “go” ✓

HMC CLI Check (if capable)

Issyscfg -r lpar -m machine --filter lpar_names="LPARname"

```bash
hmc> Issyscfg -r lpar -m P8-lime-8284-22A-SN215296V --filter lpar_names="vm61"
name=vm61,lpar_id=7,lpar_env=aixlinux,state=Running,
    resource_config=1,os_version=AIX 7.1 7100-03-05-1524,
    logical_serial_num=215296V7,default_profile=default_profile,
    curr_profile=default_profile,work_group_id=none,
    shared_proc_pool_util_auth=1,allow_perf_collection=1,
    power_ctrl_lpar_ids=none,boot_mode=norm,lpar_keylock=norm,
    auto_start=0,redundant_err_path_reporting=0,rmc_state=active,
    rmc_ipaddr=9.137.62.61,time_ref=0,lpar_avail_priority=127,
    desired_lpar_proc_compat_mode=POWER8,
    curr_lpar_proc_compat_mode=POWER8,suspend_capable=0,
    remote_restart_capable=0,remote_restart_status=Remote Restartable,
    simplified_remote_restart_capable=1,NOT this option
    sync_curr_profile=0,affinity_group_id=none,vtpm_enabled=0
```
**HMC CLI Check (if current flags)**

Check whole machine:
```
hmc> lssyscfg -r lpar -m P8-lime-8284-22A-SN215296V \
    -F simplified_remote_restart_capable,name
0,limevios1
0,limevios2
0,vm36_Ubuntu1504
0,vm26-ubuntu1504
0,vm35_SLES12
0,vm20-SLES-11.3
0,vm22-RHEL7-GA
0,vm112-64d8b471
1,vm61
hmc>
```

---

**HMC CLI set**

- To enable the remote restart feature:
  ```
  chsyscfg -r lpar -m server \
  -i "name=partition name, simplified_remote_restart_capable=1"
  ```

- Example:
  ```
  chsyscfg -r lpar -m P8-lime \
  -i "name=vm61,simplified_remote_restart_capable=1"
  ```
  - Takes a couple of seconds

- To disable similar but “=1” → “=0”
Warning:

- I am told this adds some CPU overhead on the HMC
- It regularly collects VIOS config details for SRR VM’s
- Not seen how much CPU%
  - It is not large = occasional

HMC Classic GUI (once set on the CLI)

View only - Not settable here
HMC Enhanced+ *(once set on the CLI)*

3 Validate for LPM & SRR

- Live Partition Mobility – you should know this!
3 Validate for LPM

- Live Partition Mobility – you should know this!
3 Validate for LPM

This is OK & LPM/SRR ready

3 Validate for Simplified RR

Back on the HMC CLI

hmc> rrstartlpar -o validate
    -m Source-box
    -p LPAR-name
    -t Target-box
3 Validate for Simplified RR

hmc> rrstartlpar -o validate -m P8-lime-8284-22A-SN215296V -p vm61 -t P8-emerald-8286-42A-SN100EC7V

Warnings:
HSCLB504 The migrating partition cannot use hardware-accelerated Active Memory Expansion on the destination managed system because the destination managed system does not support hardware-accelerated Active Memory Expansion.
HSCLA4CC The management console cannot maintain the source Virtual I/O Server (VIOS) slot number 14 for virtual SCSI adapter 5 on the destination VIOS partition 2*8286-42A*100EC7V.
HSCLA4CC The management console cannot maintain the source Virtual I/O Server (VIOS) slot number 3 for virtual SCSI adapter 4 on the destination VIOS partition 3*8286-42A*100EC7V.

Does rrstartlpar validate = LPM validate ?

• No as it also checks Remote Restart flag is set
• If Flag not set, you get:
  – HSCLA9B9 Partition LPARNAME is not remote restart capable.

• It might check other things too.
rrstartlpar Complete Syntax

- rrstartlpar
  -o { restart | validate | cancel | cleanup | recover }
  -m managed-system [-t target-managed-system]
  {-p partition-name | --id partition-ID}
  [--redundantvios {0 | 1 | 2}] ← old
  [--mpio {1 | 2}]
  [--vlanbridge {1 | 2}]
  [--retaindev]
  [--usecurrdata]
  [-w wait-time] ← default 3 minutes
  [-d detail-level] ← amount of output
  [--force] ← cleanup/recover
  [-v] ← verbose

SRR Official Pre-Requisites Briefly

Machine Level
1. LPM pre-reqs = access to same external storage & sub-net
2. The HMC 820 SP1 or later (with latest PTF) + 820 firmware
3. Machines are simplified remote restart capable
4. Both hosts must be managed by the same HMC
5. HMC ↔ FSP connection (needs to definitely know the box is off)
6. The source host must be in Error, Power Off, or Error - dump in progress state on the HMC. (NOTE: Power off from the HMC is OK)

LPAR / VM Level
1. VM must be Simplified Remote Restart capability enabled
2. Remote restart state of the VM must be "Remote restartable"

Note:
• SSP not officially supported ("Mine works fine!!" says Nigel)
Lets pretend the machine crashed

Example from my machines

hmc> rrstartlpar -o restart
    -m P8-lime-8284-22A-SN215296V
    -p vm61
    -t P8-emerald-8286-42A-SN100EC7V
Example using Shell variables

SOURCE=P8-lime-8284-22A-SN215296V
VM=vm61
TARGET=P8-emerald-8286-42A-SN100EC7V

hmc>
> rrstartlpar -o restart -m $SOURCE -p $VM -t $TARGET
...
>

It just works fine

• Like LPM you can watch it on the HMC
  go through various phases of creating the VM

• Restarts the VM automatically at the end for you
Clean up the debris!

VM / LPAR Definition on the source machine is still there → unlike LPM

Why?
• Source Machine was powered-off so Not possible to remove the LPAR
• VIOS was shutdown = no virtual I/O deconfig

Clean up later - after the PANIC!

SOURCE=P8-lime-8284-22A-SN215296V
VM=vm61
hmc> rrstartlpar -o cleanup -m $SOURCE -p $VM
HSCLA9CE The managed system is not in a valid state to support partition remote restart operations.
[[ Need to power-on the machine ]]

hmc> rrstartlpar -o cleanup -m $SOURCE -p $VM
HSCLA928 The Virtual I/O Server (VIOS) partition limevios1 is not in the Running state. This operation is only allowed when the VIOS partition is running.
[[ Need to start the VIOS(s) ]]
Clean up later - after the PANIC!

After machine power up, VIOS(s) started and settled down

hmc> rrstartlpar -o cleanup -m $SOURCE -p $VM
    [[for my small VM about 10 second]]

hmc>

All Done!
So how to test RR?

Test Prep setup

• Get set up (reminder)
  – Set up the Virtual Machine with SRR flag
  – HMC GUI: Operations ➔ Mobility ➔ Validate
  – HMC CLI: rrrstartlpar -o verify . . .
Test Prep setup

Get to a machine state were SSR works
   = **Power Off**

1. Shutdown Virtual Machine
2. Shutdown VIO Server(s)
3. Power Off the machine
   • **Power Off** state → ready for SRR

Note: This is slightly cheating!!
• In a real crash the VM + VIOS don’t cleanly stop & flush disks

Test Prep setup

Get to a machine state where SSR works
   = **Power Off**

• **Ugly Alternatives:**
• 1) Yank the power cords! * **
• 2) HMC Power-off Server - VIOS & VM running*

Note:
* Now VM + VIOS have to crash recover filesystems
** Need the FSP running afterwards, so power up to PowerOff
Testing via HMC CLI

- HMC CLI: `rrstartlpar -o restart` ...
  - Should return fairly quickly (after validate phase)

- Watch the HMC status to see it working on the target machine
- The VM will start automatically

- If it fails: `rrstartlpar -o recover` ...

Post Testing

Clean up source machine

1. Restart machine
2. Restart VIOS
3. Wait 5 minutes
4. `rrstartlpar -o cleanup` ...
But there is another way ...

PowerVC can do Simplified Remote Restart

Same as process but don’t use HMC rrstartlpar
Don’t worry that you can’t find any SRR clues in PowerVC

- SRR button only appears when you click a server in the right state (like Power-off) and the Host machine is SRR ready

- Warning PowerVC pre-reqs for SSR
  - HMC 830 + Firmware 830 – IMHO recommend latest update
  - VIOS 2.2.3.4+
  - Current PowerVC is 1.2.3.2
Click to select this VM

Then Click
First Host is determined by PowerVC as the best target or you select a target machine from the list.

Then Click

It starts a background job for that task.

On the screen it shows the status.

Later is just shows the VM running on the target machine.

Rather an anticlimax!!

It will clean up the original VM once the machine + VIOS started.

→ It checks every 2 minutes.
Further Information

HMC RR – Knowledge Center → Manual
HMC Community Files – DevelopWorks → Link
Whitepaper & rrMonitor script → Simplified Remote Restart Info
PowerVC Overview of Remote Restart → Link
PowerVC Deep Dive by Christine Wang → Link
Mr chmod666 Articles → PowerVCHints UsingSRR

But ...

We have only covered the Trivial Case

• Manually
  – Setup the machine / VM(s) for SRR
  – Later: Noticed a machine has stop unexpectedly
  – Selected a high priority VM to restart
  – Selected a target with enough resources
  – Kick off the recovery

• Back in the real world
  – 10’s to 100’s of machines some with
    100’s to 1000’s of LPARs
Isn’t SRR PowerHA on the cheap?

No it is not like PowerHA

**PowerHA**
- Standby VM constantly monitoring
- One per LPAR
- A second running hot OS
- Already to take-over / fail-over
- Disks already configured
- Works without HMC, FSP etc.
- Completely automatic

**Simplified Remote Restart**
- Only possible when machine = Error/PowerOff
- Can’t demine if OS failed / stuck / halted
- HMC monitoring FSP
- No HMC or no FSP = no SRR possible
- Nothing ready to take-over=has to be built/LPM
- Needs common HMC +cold reboot
- Default is manual

---

**SRR scaled up**

- Automate
  - Setup → HMC CLI = fairly easy then LPM = BAU
  - Noticed crash → HMC will notice or other Alert Tool
  - Selected priority → Your job!
  - Selected target → Your job! (also what can be sacrificed)
  - Kick off SRR → a script or tool
Manual SRR ?
Operator Driven

HMC reported HW Events (email)
1. HMC informs “operations team” of machine failure
2. They run procedure to recover production VMs ASAP
3. Check the Server is still down
4. Runs rstartlpard for
   selected VM’s to
   selected targets

• IMHO scaling to a few dozen high priority VMs - OK

Automatic SRR ?
Script(s) Driven

Via remote ssh commands to the HMC
1. Script polls for the machine state
2. Check it says down for, say, 5 minutes
   — i.e. not a reboot
3. Run a remote ssh rstartlpard for
   selected VM’s to
   selected targets
Automatic SRR ?

Script(s) Driven

Via remote ssh commands to the HMC
1. Script polls for the machine state
2. Check it says down for, say, 5 minutes
   — i.e. not a reboot
3. Run run remote ssh rrstartlpar for
   selected VM’s to
   selected targets

• Complications:
  — How to update this, if using LPM regularly!
  — Needs some way to switch off for maintenance
  — Multiple Server failures? Might be cause a right mess
  — IMHO but not scalable to dozens of Servers with 1000’s of VMs

SRR Script !

rrMonitor – not officially supported

HMC Community Files

https://www.ibm.com/developerworks/community/groups/service/html/communityview?communityUuid=0196fd8d-7287-4dff-8526-102b5fc0df5#fullpageWidgetId=W395818bd593b_487f_a7ec_79c3c27093f8
– rrMonitor_readme.txt
– rrMonitor Perl Script Download ~140 lines

• It does a number of things on a regular basis but not the cleanup function
• Parameters: source-server, target-server,
  LPAR-name, time interval (seconds) between steps
• Limitation: 1 script for 1 LPAR
• I am told that this has to be run on HMC, which is normally BANNED!!
SRR Script !

rrMonitor fixed

- POWER guru Benoit
  
  benoit.creau@chmod666.org

- Modified the original script = 217 lines of Perl

- To run as a ssh remote command from another machine

- Find it here


---

Simplified Remote Restart (RR)
via HMC and/or PowerVC

Simple to configure – assumes pure virtual & LPM

Requires POWER8 & HMC/FW 820+ [PowerVC 830]

Excellent “get out of jail free” card

No further costs

Effort required for large scale deployment

Not a PowerHA replacement
Previous Sessions:
- Linux on POWER Field Exp
- POWER8 E850 from exper
- PowerSC Tools for IBM i
- HMC 8.20 Tech Preview
- The "Key" to IBM i Licensing
- POWER8 E870 from exper
- Linux on Power: Best Pract
- Linux for AIX/IBM i guys
- PowerKVM Deep Dive
- More Tricks Power Masters
- Powerfit from hands-on
- Power up your Linux
- And more.....

Future Sessions ➔
- Dec 20th - Boost IBM i performance with IBM FlashSystem
- Suggestions Welcome

Webinar wiki: [http://tinyurl.com/IBMPowerVUGYoutubeChannel](http://tinyurl.com/IBMPowerVUGYoutubeChannel)
Webinar wiki: [http://tinyurl.com/IBMPowerVUGTwitter](http://tinyurl.com/IBMPowerVUGTwitter)

Twitter:
- @power_gaz
- @mr_mnon
- @JyotiDodhia
- @MandieQuarty

More.....

Nov 18th - VIOS Shared Storage Pool phase 5 and SSP Update

Power up your Linux

HMC 8.20 Tech Preview
- Nov 18th - VIOS Shared Storage Pool phase 5