

## HPS Service Pack 12 FLASH/Readme Second Addendum

Service Pack 12 is issued to provide support for the initial release of "RDMA transport on HPS."

Please see [Section 2: Software Updates/Performance Enhancements Item #7 RDMA transport on HPS.](#)

Service Pack 12 Addendum 2 is issued to provide support for new [HMC release 3.3.5](#)

**If you are upgrading to SP12 from a service pack that is earlier than SP9, then installing CSM1.4 is required.**

**The recommended AIX service level is AIX 5L version 5.2 is ML5.**

**The Maintenance package contains CSM 1.4 which requires RPM update openCIMOM 0.8(5.2).**

IBM pSeries High Performance Switch (HPS) Service Pack 12 requires updates to the Hardware Management Console (HMC), Switch Network Manager (SNM), HPS/SNI LPP Software (CSS), Global Firmware (GFW), Power Subsystem Microcode (ptcode), AIX 5L Version 5.2 and various supporting AIX LPPs.

This document contains general guidelines for upgrading the components listed in Section 3: Component Update/Download Information.

These guidelines are intended to be a supplement to the other IBM documents referred to in this document. We strongly advise that you have the referenced documents available before you begin the upgrade process. Section 1 contains a list of referenced documents.

Only the levels listed in "Section 3" are supported with HPS Service Pack 12. Therefore, if you are migrating/installing Service Pack 12, all components should be at the indicated service levels when you are finished. No mixing of Service Pack maintenance levels is supported at this time. Users must follow the procedure in the order outlined in Section 4: "Recommended Installation Sequence (Overview)". Failure to do so may cause problems with configuring the HPS SNI devices.

Because '/var' is a system data repository, system administrators should check periodically to maintain /var such that there is at least 30 Mb free [use 'df -k']. If it is more than 75% full, look for the directories that contain the most data. [use 'du /var | sort -n' ]

Table of Contents

- ↓ [Supporting Documentation](#)
- ↓ [Software Updates/Performance Enhancements](#)
- ↓ [Component update/download information](#)
- ↓ [Recommended installation sequence \(Overview\)](#)
- ↓ [Installation Guidelines](#)
- ↓ [Known problems/workarounds/Restrictions](#)
- ↓ [Service Pack fix list \(by component\)](#)

## Section 1: Supporting\_Documentation

### Cluster Systems Management for Linux on xSeries

1. [Hardware Management Console](#)
2. [Hardware Management Console for pSeries Installation and Operations Guide \(SA38-0590-07\)](#)
3. [Hardware Management Console for pSeries Maintenance Guide \(SA38-0603-05\)](#)
4. [pSeries High Performance Switch \(HPS\) Planning, Installation and Service Guide \(GA22-7951-01\) \(HPS Guide\)](#)
5. [IBM eServer pSeries 690](#)
6. [Switch Network Interface for eServer pSeries High Performance Switch Guide and Reference \(SC23-4869-01\)](#)
7. [AIX 5.2 documentation](#)
8. [AIX 5L Version 5.2 Installation Guide and Reference](#)
9. [AIX 5L Version 5.2 Performance Management Guide](#)
10. [Reliable Scalable Cluster Technology \(RSCT\) Library](#)
11. [General Parallel File System \(GPFS\) Library](#)
12. [Cluster System Management \(CSM\) Library](#)
13. [IBM Parallel Environment for AIX 5L Installation Version 4 Release 1.1](#)
14. [IBM LoadLeveler for AIX 5L and Linux Using and Administering Version 3 Release2](#)
15. [LoadLeveler 3.2 documentation updates](#)

## Section 2: Software Updates/Performance Enhancements

1. [New SNM/FNM Features](#)
2. [MPI and LAPI Performance Enhancements](#)
3. [Support for Node Switch Board \(NSB\) and Intermediate Switch Board \(ISB\) failure](#)
4. [Application striping and fail-over of packets across multiple interfaces for fault resilience](#)
5. [Application Checkpoint/Restart and preemption](#)
6. [The Web-based System Manager Remote Client a.k.a WebSM PC CLIENT may need to be reinstalled](#)
7. [RDMA \(remote direct-memory access \) transport on HPS](#)

---

## [1. New SNM Features:](#)

### **New features and functions (on The Switch Network Management Panel ):**

Introduced in HMC 1.3.1.0 - Service Pack 9

For details of operation please see pSeries High Performance Switch (HPS) Planning, Installation and Service Guide (GA22-7951-02) (HPS Guide) Chapter 4. System management components and Appendix E. Switch Network Manager (SNM)

#### **Features:**

**HMC Fail-over:** FNM/SNM daemon ("Enable SNM Software in Normal Mode") runs on up to four HMCs in the cluster.

**Verification mode:** "Enable SNM Software for Switch Network Verification" used to check out the system.

#### **Functions :**

There are three(3) new tasks and one(1) renamed task on the Switch Network Management Panel :

NEW: "Select Logical Topology", "Display Cluster Components", "Enable SNM Software for Switch Network Verification" RENAMED: "Enable SNM Software for Normal Mode" used to be "Enable SNM Software"

- \*\*\* Select Logical Topology this introduces a requirement to set the logical topology of your system before enabling SNM software. If the logical topology has not been set, neither of the tasks to Enable SNM Software will be selectable. This is required. See note 1.
  - Note: If you are using the SNM Fail-over support introduced in the Software Update release, the logical topology must be set on each HMC where you enable SNM software.
- Display Cluster Components shows the frames and cages of the cluster components with which SNM can communicate. This task can only be used if SNM software is **NOT** enabled.
  - The new lsswcomp command provides the same function as the "Display Cluster Components" task.
- Enable SNM Software for Switch Network Verification task is used to check out the system by

NOT removing bad links or routes. The main purpose is to discover cable mis-wires and defective hardware. No Routing is modified or loaded. It is NOT intended to be used in a working environment.

- \*\*\* Enable SNM Software for Normal Mode. select this to run SNM in a normal, working environment. This is required. See note 1.

#### **Notes:**

**1) \*\*\* If you are upgrading from Service Pack 8 or less, then these steps are REQUIRED for an initial installation of FNM/SNM:**

**First Select Logical Topology then  
Enable SNM Software for Normal Mode.**

See "Enable SNM Software" in the HPS Guide.

---

#### **2. MPI and LAPI Performance Enhancements:**

LAPI and MPI host communication stacks have been tuned, based on experience, for lower latency. These changes complement the communication performance improvements released in part 1. They are being released as one package with the LAPI and MPI striping function.

Note: Beginning with the Service Pack 9:

- a single MPI user-space job which wants to use HPS adapters in both of two switch planes **must set MP\_EUDEVICE=sn\_all**(or csss). In previous releases, a single MPI/us job with multiple tasks per node could use adapters from two planes by setting MP\_EUDEVICE=sn\_single (or css0).
  - A job which sets MP\_EUDEVICE=sn\_single will only be able to use half of the HPS adapters on the nodes where it runs.
  - A job must set MP\_EUDEVICE=sn\_all in order to use all the HPS adapters in the nodes it is using.
- 

#### **3. Support for Node Switch Board (NSB) and Intermediate Switch Board (NSB) failure:**

The SNM daemon has been enhanced to ensure that the failure of an NSB and an ISB will be handled appropriately. The main enhancement is that the Switch Network Manager daemon gathers and downloads multiple path table updates into a single transaction instead of one update per transaction. An entire switch board failure or recovery will generate multiple of path entries changes. Therefore handling these updates in groups involves many fewer transactions. These can be processed in a more timely fashion and place a smaller load on SNM and the service network.

---

#### **4. Application striping and fail-over of packets across multiple interfaces for fault resilience:**

Striping provides a method for a single task of a parallel application to utilize multiple SNI links. This allows a single link to fail without the application using the link being terminated because there is an alternate path to all other tasks of the application. The striping method implemented is designed to provide resilience to switch or interface/link failure and is not designed to provide an increase in the aggregate network bandwidth as seen by tasks.

This link failure resilience function requires support in LoadLeveler, Parallel Environment and the host communication stack for MPI and LAPI. When an interface/link recovers, the striping function will ensure that the link is used again. Another goal of this striping design is to evenly distribute traffic over multiple switch networks to maintain a balance in the communications load.

#### **KNOWN ISSUES/CONCERNS**

- In non-striping mode you only get half the links on a p655 system using a 2 plane configuration when setting MP\_EUDEVICE=sn\_single. You must use MP\_EUDEVICE=csss or sn\_all to get all links.
- Performance degradation for single task per link of 3% for unidirectional and 5% for exchange bandwidth at large messages. However, multiple tasks per link get full link bandwidth.
- Striping currently limited to two links per task. Multiple tasks can use all links as long as there are more tasks than half the links available.

**Note that when using RDMA, there is no large message performance penalty for striping. In fact striping greatly improves large message RDMA bandwidth. The above performance degradation occurs only in packet mode.**

---

#### **5. Checkpoint/Restart and Preemption:**

Checkpoint/restart provides a way to stop and resume applications at some later time. It is a very useful tool in managing the clusters workload since applications can be check-pointed to allow other applications to run or the system made inactive for a maintenance window.

Preemption is also provided so a running application can be suspended to allow another higher priority application to execute right away.

---

**6. The Web-based System Manager Remote Client a.k.a WebSM PC CLIENT** may need to be reinstalled on your remote server or PC in order for WebSM to work with this HMC Build.

To install the PC Client software on your remote server or PC go to: [http://<hmc-hostname>/remote\\_client.html](http://<hmc-hostname>/remote_client.html) .

The two versions you see there are the "legacy" websm client - "install shield" and - webstart. Both versions can exist simultaneously on your client workstation

### **Install shield:**

If there is a currently installed install shield version of WebSM, it **must** be **UNinstalled** before the new version is installed.

Uninstall : From Start -> Control Panel -> Add/Remove Programs -> Web-based System Manager Remote Client click "Change/Remove" and follow instructions.

**Failure to uninstall will result in undefined behavior of the WebSM PC Client.**

Install : Go to: [http://<hmc-hostname>/remote\\_client.html](http://<hmc-hostname>/remote_client.html) -> "Install Shield" -> select the Operating System where the program will be running.

Select "Open" to immediately install on the machine on which the browser is running.

Selecting "Save" will allow you to store the install program for installation at a later time.

NOTE: This is a 100 Mb file so either process will require about 15 minutes.

Once saved it is a self extracting file that can be used to upgrade multiple PC's.

### **Webstart:**

The webstart installation is a two part installation

first - install the 1.4.2 JVM that is has the webstart code.

second - invokes the launch code and gets the initial classes downloaded from the HMC.

A feature of the webstart client is that when the client is launched, it will check for new jar files on the HMC and download them, so subsequent update of the HMC server code do not require the webstart client code to be constantly updated .

There may be an issue of having multiple jvms on the client workstation and the default jvm path not pointing to the IBM 1.4.2 jvm. Webstart does not work well with the sun 1.4.2 jvm, so some cleanup of multiple jvms on your workstation may be needed.

The Java Web Start version requires multiple logins on launching. You have to login to the original HMC whenever you try to connect even if you're just downloading the webstart .jsp files.

If you are getting the .jsp files from a regular AIX box, you will have one less login.

You dont experience this in the non-webstart websm because you are not connecting to the webserver to download the .jsp files. Trying to 'cancel out' will result in multiple (up to 16) re-presentations of the login verification box.

#### 1) Install Java Web Start:

Java Web Start for Linux    Install Java Web Start on a Linux platform. Once installed, return to this page to download the Remote Client. Java Web Start for Windows    Install Java Web Start on a Windows

platform. Once installed, return to this page to download the Remote Client.

2) Download Web-based System Manager Remote Client for Java Web Start:

Remote Client Download Web-based System Manager Remote Client for Java Web Start on Linux and Windows systems.

The images are quite large, so it may take some time for the download to complete.

---

## [7. RDMA transport on HPS](#)

### **Summary:**

The IBM eServer pSeries High Performance Switch (HPS) with Remote Direct Memory Access (RDMA) enabled delivers 1.8GB/s of unidirectional bandwidth across a single link and over 3.5GB/s over 2 links to a single MPI task for some applications. Bidirectional bandwidth measures almost 3.0GB/s over a single link and over 5.9GB/s for a single MPI task striping over 2 links.

These measurements were achieved on 2 nodes containing 1.9GHz, Power 4+ based processors with 32 CPUs and 2 HPS adapters, and links on each of the 2 nodes. Large pages and cache-line aligned buffers were used for the MPI measurements. These results indicate nearly linear scaling of large message striping performance. A single link is able to deliver over 1.65GB/s of bandwidth over TCP/IP.

---

### **Advantages of RDMA transport:**

- Decouples the CPU from the movement of data allowing for better overlap of computation and communication.
- Reduces the computational load of the CPU by off-loading segmentation and reassembly of messages to the network adapter, thereby reducing the number of packet arrival interrupts.
- Reduces the stress on the memory subsystems by reducing the number of bus crossings - one I/O bus crossing versus the traditional 2 memory bus and 1 I/O bus crossing when not using RDMA.
- Enables the protocols to efficiently stripe a message (or different messages) from a single task across multiple network interfaces to exploit the available communication bandwidth in parallel without engaging multiple CPUs.
- Provides improved raw transport performance. In cases where the transport bottleneck is the copy rate (memory bandwidth), RDMA helps eliminate that bottleneck.

### **How can applications better exploit RDMA?**



To exploit the overlap potential of RDMA, applications should make use of non-blocking calls and schedule communication as early as possible to extract the full overlap potential.

For MPI/LAPI applications the RDMA transport kicks in for large messages (to justify the setup costs of RDMA).

Applications that reuse the same communication buffers will benefit further from RDMA since the setup cost is amortized over larger number of transfers.

For applications that send large messages asynchronously, enabling striping should help the applications realize superior transport bandwidth.

RDMA transport is **disabled** by default in this code release in order to insure that customers running production workloads that do not need RDMA operate without change.

**Note :** When RDMA is enabled by changing the device attribute **rdma\_xlat\_limit** to a value greater than 0, it is possible for an RDMA enabled job to surpass the AIX pin limit which may lead to a hung system. If this pin limit is exceeded slowly enough the parallel application will catch a SIGDANGER signal and kill the job. However, if this pin limit is exceeded too rapidly, AIX may resort to killing processes, including systems processes, which may cause system to appear to be hung. This situation may be prevented by setting the **rdma\_xlat\_limit** to a number which is smaller than 4096 times the sum of the number of large and small pages on the system.

## **RDMA Usage:**

By default, the RDMA capability of the SNI adapters is turned off.

Use the **chdev** (or optionally **/usr/sni/aix52/chgsni**) command to turn on RDMA by setting **rdma\_xlat\_limit** to a value greater than 0x0.

The **rdma\_xlat\_limit** attribute controls amount of memory allowed for RDMA, per LPAR.

The default value of the **rdma\_xlat\_limit** attribute is 0x0, effectively turning RDMA off.

When using RDMA it is recommended that this attribute value be set to 0x8000000000000000.

This allows the SNI device driver to optimally manage RDMA memory usage.

**PLEASE NOTE: This odm attr ( rdma\_xlat\_limit ) is changed for ALL sni's in an LPAR with just the single command.**

**For the change to take effect, the LPAR must be rebooted.**

**The RDMA option MUST be turned ON OR OFF across the entire cluster.**

**If it is not, SNI adapters will not communicate due to version mismatch failures!**

## **NOTES:**

When RDMA is turned on, the IP protocol automatically uses the RDMA capability.

For user space jobs, an additional LoadLeveler keyword (**bulkxfer=yes**) must be set in the LoadLeveler job control file to indicate that this job is requesting RDMA.

This allows users to run both RDMA and non-RDMA jobs on the system.

[For the most current details see: LoadLeveler 3.2 documentation updates; Addition of support for Bulk Data Transfer ...](#)

[For more information, please see the "LoadLeveler Using and Administration Guide".](#)



To enable RDMA: /usr/sbin/chdev -l sniX -a rdma\_xlat\_limit=0x8000000000000000

To disable RDMA: /usr/sbin/chdev -l sniX -a rdma\_xlat\_limit=0x0000000000000000

To check the RDMA status: -- run lsattr -El sni0 ; lsattr -El sni1 ;

```
Ex: for i in `lsdev|grep sni.*Avail|awk '{print $1}'`; do lsattr -El $i|grep rdma; done
chdev Usage: chdev -l Name [-a Attribute=Value]...[-p ParentName][[-P|-T]
chdev -l Name -a rdma_xlat_limit=size
```

The 'True' seen in the lsattr output is whether or not the value is user modifiable.

The rdma\_xlat\_limit attribute is one of 5 that are not valid on just 1 sni -- so the change method updates the odm for all sni's

This attribute is modified for all existing devices. Any user supplied -l option is ignored.

Attribute values for rdma\_xlat\_limit :

Default value: 0x0 (Implies RDMA is turned off)

MINimum value: 0x0

MAXimum value: 0xFFFFFFFFFFFFFFFF ( Maximum amount of memory allowed for RDMA (in bytes). )

Recommended value: 0x8000000000000000

---

### Section 3: Component update/download information

Component	Service pack updates	Download sites
<a href="#">HMC</a>	<p>This service pack requires that the HMC is updated to HMC V3R3.5. Choose HMC_Update_V3R3.5</p> <p>Service Pack 12 requires updates to this component.</p> <p>Note: CD images <b>cannot</b> be downloaded directly from this site. Contact your IBM Sales Representative or Business Partner, and order Hardware Feature Code (MES) 0960 for the initial upgrade CDs.</p>	<p><a href="#">HMC corrective service</a> or <a href="http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/v3r30.html">http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/v3r30.html</a></p>
<a href="#">SNM/FNM</a>	<p>This service pack requires that the SNM is updated to SNM Version: 1.3.3.0 , Release 1 APAR IY67734 PTF U802490 Choose "SNM_Serv_Pack12.zip"</p> <p>Service Pack 12 requires updates to this component.</p>	<p><a href="#">HMC corrective service</a></p>

---

## [GFW](#)

3H050215.img (p690) or 3J050215.img (p655)

[Microcode download \(System\)](#)

This code is machine specific.

Go to the "System" section.

For 7040-681 Version 3 p690 models, choose "version 3H050215".

For 7039-651 p655 models, choose "version 3J050215".

Service Pack 12 requires updates to this component.

Note that GFW is available in IBM CORE 3-4 days earlier than the above mentioned website. Please contact your IBM CE for the GFW in IBM CORE if not available on website. Refer to this website for detailed download and unpacking procedures:

<http://techsupport.services.ibm.com/server/mdownload/downproc.html>

---

## [Power Subsystem Microcode](#)

Go to "Power Subsystem for 7039-651 (p655) and servers containing the 7045-SW4 (High Performance Switch)"

[Download microcode \(Other\)](#)

Code is the same for both p690 and p655. (ptcode-1.74.26a7-1.i386.rpm)

Note that Power Subsystem Microcode is available in IBM CORE upto 1 week earlier than the above mentioned website. Please contact your IBM CE for the Power Subsystem Microcode in IBM CORE if not available on website.

Service Pack 12 requires updates to this component.

---

## [AIX](#)

The recommended AIX service level for AIX 5L version 5.2 is the 5200-05 Recommended Maintenance package.

Specific Fixes APAR IY66260

Service Pack 12 requires updates to this component.

[pSeries support fixes](#)

## openCIMOM

CSM 1.4 requires RPM update [openCIMOM update](#)

If you are upgrading from Service Pack 9 or higher, then you do not need to upgrade this component.

If you are upgrading from anything earlier than Service Pack 9, then Service Pack 12 requires updates to this component.

## AIX LPPs:

If you have any of the following AIX LPPs installed, Service Pack 12 requires updates to these LPP components.

If you have any of the following AIX LPPs installed, apply the listed APARs which are needed for HPS Service Pack 12. They are available from the website: [pSeries support fixes](#)

To verify that the service levels for your LPP's are current for Service Pack 12, on each logical partition issue:

```
lspp -Lc | egrep "vsd|LAPI|HPS|sni|ppe|LoadL|mmfs|rsct|csm|essl|pessl" | cut -d : -f 2,3 | sed 's:/ /'
```

Depending on the LPP's you have, your output should match these LPP fileset levels:

### Detailed LPP Level Check

LPP	Release	Component ID	APAR	Level check
<a href="#">VSD</a>	410	5765G2602	IY67437	rsct.vsd.cmds 4.1.0.11 rsct.vsd.rvsd 4.1.0.9 rsct.vsd.vsdd 4.1.0.9 rsct.vsd.vsdrm 4.1.0.3
<a href="#">LAPI</a>	232	5765G2601	IY67438	rsct.lapi.rte 2.3.2.4 rsct.lapi.nam 2.3.2.4 rsct.lapi.samp 2.3.2.0
<a href="#">HPS/SNI</a>	112	5765G2400	IY67439	devices.chrp.IBM.HPS.hpsfu 1.1.2.0 devices.chrp.IBM.HPS.rte 1.1.2.2 devices.common.IBM.sni.ml 1.1.2.1 devices.common.IBM.sni.ntbl 1.1.2.0 devices.common.IBM.sni.rte 1.1.2.2
<a href="#">PPE</a>	411	5765F8300	IY67440	ppe.poe 4.1.1.6 ppe.pvt 4.1.1.1 ppe.dpcl 3.3.4.0 ppe.perf 4.1.1.3 ppe.man 4.1.1.1 ppe.loc.license 4.1.0.0

<a href="#">LoadL</a>	320	5765E6900	IY67441	LoadL.full	3.2.0.13
				LoadL.loc.license	3.2.0.0
				LoadL.so	3.2.0.13
				LoadL.tguides	3.2.0.1
<a href="#">GPFS</a>	210	5765F6400	IY67442	mmfs.gpfs.rte	2.1.0.19
				mmfs.base.cmds	3.5.0.11
				mmfs.base.rte	3.5.0.17
<a href="#">GPFS</a>	221	5765F6400	IY67443	mmfs.gpfs.rte 2.2.1.4	
				mmfs.base.cmds 3.6.1.4	
				mmfs.base.rte 3.6.1.4	
				mmfs.gpfsdocs.data 3.6.1.1	
<a href="#">GPFS</a>	230	5765F6400	IY66081	gpfs.base	2.3.0.1
				gpfs.docs.data	2.3.0.1
<a href="#">CSM</a>	140	5765E88AP	IY67444	csm.client 1.4.0.11	
				csm.core 1.4.0.12	
				csm.diagnostics 1.4.0.12	
				csm.dsh 1.4.0.12	
				csm.gui.dcem 1.4.0.10	
				csm.gui.websm 1.4.0.10	
				csm.server 1.4.0.12	
				pssp.pssp_to_csm 1.1.1.0	
<a href="#">RSCT</a>	235	5765F07AP	IY67445	rsct.basic.hacmp	2.3.5.0
				rsct.basic.rte	2.3.5.2
				rsct.basic.sp	2.3.5.0
				rsct.compat.basic.hacmp	2.3.5.0
				rsct.compat.basic.rte	2.3.5.0
				rsct.compat.basic.sp	2.3.5.0
				rsct.compat.clients.hacmp	2.3.5.0
				rsct.compat.clients.rte	2.3.5.0
				rsct.compat.clients.sp	2.3.5.0
				rsct.core.auditrm	2.3.5.1
				rsct.core.errm	2.3.5.1
				rsct.core.fsrn	2.3.5.1
				rsct.core.gui	2.3.5.1
				rsct.core.hostrm	2.3.5.0
				rsct.core.rmc	2.3.5.2
				rsct.core.sec	2.3.5.0
				rsct.core.sensorrm	2.3.5.0

rsct.core.sr	2.3.5.0
rsct.core.utils	2.3.5.2
rsct.sdk	2.3.3.0

Parallel ESSL	310	5765F8400	<a href="#">PQ96735</a> <a href="#">PQ99640</a>	pessl.rte.common 3.1.0.2 pessl.rte.mp 3.1.0.0 pessl.rte.rs1 3.1.0.2 pessl.rte.rs2 3.1.0.0 pessl.rte.smp 3.1.0.2 pessl.rte.up 3.1.0.0 pessl.loc.license 3.1.0.0 pessl.man.en_US 3.1.0.0
------------------	-----	-----------	--	---

ESSL	410	5765F8200	<a href="#">PQ90826</a> <a href="#">PQ90823</a>	essl.rte.common 4.1.0.1 essl.rte.smp 4.1.0.1 essl.rte.up 4.1.0.0 essl.rte.mp 4.1.0.0 essl.rte.rs1 4.1.0.1 essl.rte.rs2 4.1.0.0 essl.man.en_US 4.1.0.0 essl.loc.license 4.1.0.0
------	-----	-----------	--	---

Section 4: Recommended installation sequence (overview):

1 [Install HMC Software](#)

Important Preliminary steps:

- Check if HMC is at HMC V3R3.0 or higher before you can install this update (required).
- Check if BIOS update is required on HMC.
- Check if BIOS hyperthreading is to be disabled on HMC
- Disable the SNM/FNM software from the Switch Network Management Panel.

Perform one of the following installation tasks:

- New Install HMC V3.3.0 from Recovery CD OR
- Install Upgrade from Recovery CD OR
- Update from .zip file (web)

Level Check: <>Version: 3; Release: 3.5; HMC Build Level 20050320.1

2 [Install SNM Software](#)

Level Check: IBMhsc.SNM.1.3.3.0

3	<p><a href="#">Install HPS/SNI LPP Software</a></p> <p><b>IMPORTANT:</b>          If you are upgrading from Service Pack 6 or below - do not reboot logical partitions (LPARs) until after Step 4 Install GFW is complete!          See Problem #1 in the "Known Problems" section of this document for more information.</p> <p>Level Check: devices.chrp.IBM.HPS.rte 1.1.2.2</p>
4	<p><a href="#">Install GFW</a></p> <p>(using the recommended AIX command line (update_flash) method with a locally available image)</p> <p>Level Check: ROM Level (alterable).....3H050215 (P690)          ROM Level (alterable).....3J050215 (P655)</p>
5	<p><a href="#">Install Power Subsystem Microcode on each frame</a></p> <p>Level Check: 26a7</p>
6	<p><a href="#">Install AIX base updates on each node</a></p> <p>Level Check: bos.mp64 5.2.0.54</p>
7	<p><a href="#">Install AIX LPP updates on each node</a></p> <p>Level Check: see "Detailed LPP Level Check"</p> <p>Note: CSM LPP's need to download and install openCIMOM-0.8-1 RPM update.</p>

Section 5: Installation Guidelines

[HMC Installation Guidelines:](#)

1	<p>Install HMC Software</p> <p>Important Preliminary steps:</p> <ul style="list-style-type: none"> <li>1a - Check the HMC version level</li> <li>1b - Check if BIOS update is required on HMC.</li> <li>1c - Check if BIOS hyperthreading is to be disabled on HMC</li> <li>1d - Customer Decision regarding install, upgrade or update.</li> <li>1e - Disable the SNM/FNM software from the Switch Network Management Panel.</li> </ul> <p>Perform one of the following installation tasks:</p> <ul style="list-style-type: none"> <li>1f - New Install HMC V3.3.5 from Recovery CD OR              Upgrade to HMC V3.3.5 from Recovery CD OR</li> <li>1g - Update to HMC V3.3.5 from .zip file</li> </ul> <p>Service Pack 12 includes a new version of WebSM PC CLIENT</p> <ul style="list-style-type: none"> <li>1h - If you wish to install this version of WebSM PC CLIENT              Uninstall - reinstall HMC WebSM PC CLIENT (Install Shield version)</li> </ul>
---	--

Level Check:

1i - Verify: Release 3, Version 3.5 HMC Build Level 20050320.1

**Important Preliminary steps:**

1a **If you are upgrading from Service Pack 6 (HMC 3.2.6 ) or earlier, you must perform a new install or upgrade to your HMC using the latest recovery CD's .**

Please read [Version 3.3 machine code Updates](#) , select "for POWER4 servers" tab, for important information about acquiring HMC V3.3.0 Recovery CD's:

Order the Hardware Feature Code (MES) 0960 for the initial upgrade CDs from your IBM Sales Representative or Business Partner.

For more information see: [https://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/ptf\\_v3r30mc.html](https://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/ptf_v3r30mc.html)

Contact your IBM Sales Representative or Business Partner, to order the initial upgrade CDs

1b **Check if BIOS update is required on HMC.**

There is a mandatory Bios upgrade required for all Bradley logan HMC PC's. 7315-C03, 7310-C03, 8187-KUH.

If your HMC model is not a 7315-C03, skip this step.

If you are updating the HMC on HMC model 7315-C03 hardware, then you must first update the BIOS of that HMC model. The BIOS and install instructions can be obtained from [BIOS updates](#) on this website.

This BIOS will also ship as part of Feature Code 0960

DownLoad: <http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/v3r30.html>

1c **Check if BIOS hyperthreading is to be disabled on HMC.**

Many of the rack mounted HMC's (8187-KUH, 7315-C03) have a BIOS option to enable hyperthreads.

The imbedded kernel will not run well when this option is enabled.

You must disable this setting before upgrading to HMC3.3.0.



## 1d Customer Decision regarding install, upgrade or update.

Beginning with HMC Version 3 Release 3.2, (Service Pack 10) the Ext3 (JFS) filesystem will be enabled if customers [perform a New Install/ Upgrade](#) to this level of code by using the HMC Recovery CDs. The Ext3 filesystem is a journaled filesystem and is more reliable and less prone to corruption in case of unexpected loss of power on the HMC.

### Please note the difference between Upgrade and Update:

- **Upgrade** is done via cd load as described in step - [Perform an Install New / Upgrade](#)
- **Update** is done via downloading a zip file as described in step - [Perform an Update](#)

### NOTES:

- **Updating** to HMC Version 3 Release 3.n using the Install Corrective Service will **not** enable this feature.
  - **For Upgrade Installation**
    - the following steps must be taken, prior to rebooting the HMC for the Upgrade process:
      - Ensure that the user's home directories are not filled up with debug data. ( The Upgrade partition only has 2GB in free space to preserve the upgrade data.) -
      - Ensure that HMC debug is turned off. This can be achieved by running the command pedbg -d off. ( With debug enabled, certain log files will be locked for writing and will prevent the Save Upgrade Data task to complete successfully. )
      - Perform a Save Upgrade Data task from the Software Maintenance Panel on the HMC console. The Perform a Save Upgrade Data task should be run immediately before doing the reboot of the HMC with volume 1 of the recovery CD. If the HMC reboot does not go to the install menu of volume 1 of the recovery CD, you should repeat the save upgrade task.
- ◇
- The procedure for both Installation and Upgrade is identical except:
    - For New Installation: When asked to perform an Install/Recovery or Upgrade, select Install/RecoveryF8.
    - For Upgrade Installation: When asked to perform an Install/Recovery or Upgrade, select Upgrade F1.

## Installation steps:

- 1e **Disable the SNM/FNM software** from the Switch Network Management Panel.  
Do this on all HMC's attached to the cluster.
- 1f **Perform an Install New / Upgrade :**  
**Install HMC Recovery CD 1 of 2, RG\_BASE\_050320.1.iso on all HMCs in the system - Required**  
**Install HMC Recovery CD 2 of 2, HMC\_REC\_335.1.iso on all HMCs in the system - Required**
- Reboot the HMC with volume 1 of the recovery CD inserted in the DVD Ram drive.
- NOTE: If the HMC fails to boot volume 1 of the recovery CD, the boot sequence in the HMC BIOS may need to be changed so that the DVD/CDROM is before the hard disk in the startup sequence.
- If you have run the save upgrade data task before the startup sequence was set correctly, then you should rerun the save upgrade data task before installing the HMC with volume 1 of the recovery CD.
- Select F8 for New Installation  
OR
  - Select F1 for an Upgrade installation. **(NOTE: This is NOT the same as an UPDATE)**
  - On the next screen to confirm your selection. Select F1.
  - The Install/Upgrade process will proceed until -
  - At some point you will be prompted to insert the second CD.
  - Remove the CD from the DVD Ram drive and hit enter when the install is completed.
- 1g **Perform an Update**  
**If you are UPDATING from Service Pack 7 or higher and choose to do the UPDATE:**  
**Install HMC PTF HMC\_UPD\_335.1.zip**  
Install the HMC PTF **HMC\_UPD\_335.1.zip** from the HMC support link ONLY if the HMC is at a Release 3 Version 3.X level (3.3.X):  
<http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/v3r30.html>
- Select HMC PTF Update for Service Pack 12
  - Use the HMC --> Install Corrective Service option to install.
  - Reboot HMC after successful installation.
- You may install this UPDATE directly from the web: via "HMC Install Corrective Service"**  
**See step 3 'Install Update', in HMC Readme -**  
**[http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/ptf\\_v3r33mc.html](http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/ptf_v3r33mc.html)**
- NOTE: For this release you may see a number of informational messages similar to this:  
'removal of /tmp/mcp-update/xxprt-4.2.0-234.i386.rpm failed: No such file or directory'

When we install the new MCP base, it tries to delete files that don't exist on the HMC.  
This is not an error nor a warning.

1h

**IF** the Install Shield version of the Web-based System Manager Remote Client a.k.a WebSM PC CLIENT is to be reinstalled on your remote server or PC :

After the HMC is installed :

**Uninstall, if present, the** existing Websm client:

Control Panel "Add/Remove Programs"

"Web-Based System Manager Remote Client"

**Install:**

Using your favorite browser, go to this URL:

http: web browser://<hmc-hostname>/remote\_client.html

install the PC Client software on your remote server or PC. The down load and install may take 10 to 20 minutes.

For complete details see: [Hardware Management Console for pSeries Installation and Operations Guide \(SA38-0590-07\)](#) - Chapter 9. Installing and Using the Remote Client

1i

HMC Level Check:

Command Line :

lshmc -V shows:

Version: 3

Release: 3.5

HMC Build level 20050320.1

GUI:

Menu "Help" > "About Hardware Management Console" shows:

Version: 3

Release: 3.5

HMC Build level 20050320.1

Remote Client (AKA Web-based System Manager AKA WebSM):

Menu "Help" > "About Web-based System Manager" shows:

Build: Sun Mar 20 14:22:42 CST 2005

Level: development

## HMC Important Notes:

- Install the HMC code by following the instructions in the HPS Guide.
- Have your IBM CE download the most recent copy of the HPS guide from IBM CORE to get updated HPS install information.
- Also you should review the HMC information on the web page where you downloaded the images.
- Service Pack 12 REQUIRES:
  - HMC V3.3.n (required since Service Pack 6 release) and
  - HMC V3.3.5 PTF Update.
- HMC V3.3.0 is a NEW BASE release of the HMC introduced in Service Pack 6 that uses a new imbedded kernel.
- This version MAY OVERWRITE root directories (for /, /home/root and /home/hscroot) deleting any scripts that may be there.
- This version MAY delete the Power Subsystem Microcode RPM images on your HMC and you will have to re acquire it for future installs.
- The HMC is installed using 2 CD's:  
Recovery CD filenames: (HMC 3.3.5)
  - CD 1 of 2 RG\_BASE\_050320.1.iso
  - CD 2 of 2 HMC\_REC\_335.1.iso
- The login available at virtual console 0 (via the CTRL-ALT-F1 key sequence) is no longer available.
- New Installation and PTF update installation are supported for this release.
- Upgrade installation is only supported when upgrading from HMC 3.2.X or greater.

As part of any system change - it is recommended to have a hard copy of network connections, 8 port/ran box

configurations and Switch Group IP's.

- Network connections are on the GUI: HMC Maintenance Panel => System Configuration => Customize Network

Settings:

IP Address and Netmask for Ethernet0 and Ethernet1, Default Gateway, Nameserver, Domain

- 8 port RAN box configurations are on the GUI: HMC Maintenance Panel => System Configuration => Configure

Serial Adapter:

Option 2 shows the current configuration

- Switch Group IP's are on the GUI: Switch Management => Switch Utilities => Switch Group Configuration

Known problems and issues with the HMC V3 R3.0 Environment:

- wu-ftp will be removed from the HMC distribution. One will be able to ftp out of the HMC but not into the HMC. The 'scp' command is available if you enable secure shell (ssh).
- The websm PC client has a performance decline when downloading the plugin classes from the server. The first time an operation is performed using the client, the task may be slow to launch. Subsequent use of the task, will respond as normal.

Retain Tip on how to use pesh:

To give IBM support personnel the ability to retrieve certain trace/debug information on the HMC, the customer should create a user "hscpe" and assign a password. IBM support can contact the customer to get the password, and then remotely connect to the HMC (with customer consent).

This allows IBM support to perform additional functions, such as viewing logs or starting trace to diagnose problems on the HMC. This user has access similar to the "hscroot" user on HMC. When accessing the HMC remotely via ssh, the "hscpe" user is put into the restricted shell environment. To bypass the restricted shell, pesh command is provided. pesh command can only be run by the "hscpe" user, allowing this user to pass in the serial number of the HMC. If the serial number is correct, the user is required to enter a password obtained from IBM Support. If the password is correct, then the user is then put into the un-restricted shell as user "hscpe".

Example:

```
pesh 23A345K          ( enter the serial number in upper case letters)
```

You will be prompted for a password. Enter password that was provided by IBM Support in lower case letters.

The HMC serial number can be queried using the command, "lshmc -v | grep SE" or read from the label that is on the front of the HMC.

Use the command "date" to verify that the date of the HMC is for the day you intend to use the pesh command.

Starting with HMC Version 3 Release 3.0 and Version 4 Release 1.0, user can also access the restricted shell terminal on the local HMC, by right mouse click on the desktop and selecting the Terminal--rshterm task. If one login at the HMC as user hscpe, the pesh command can also be run from the restricted shell terminal.

For HMC Version 3 Release 3.0 and below, the "hscpe" user id can be created with any role, however, in order to use some of the High Performance Switch (HPS) debug commands, the Service Rep role needs to be selected.

---

For new HMC installation(s) follow the instructions as described in [Hardware Management Console for pSeries Installation and Operations Guide \(SA38-0590-07\)](#).

To understand how to connect the rs422/rs232 cables see the HPS Guide: Chapter 6; Step 6, "Install the Hardware Management Console (HMC)" thru

Step 16. "Verify Installation is Complete"

For "Code load requirements for existing server frames" see Chapter 6

For p655 "Code load requirements for existing p655 server frames"

For p690 "Code load requirements for existing p690 server frames"

[FNM/SNM Installation Guidelines](#)

2 Install SNM Software (APAR IY67734 , PTF U802490)  
Level Check: Version:.1.3.3.0

2a The SNM/FNM software should already be disabled from procedure 1.  
If not, disable it now.

2b Installation of SNM\_Serv\_Pack12.zip is required on all of the HMCs in the cluster.  
( The version of SNM distributed with HMC 3.3.4 is not the most current version. )  
Follow the install instructions by choosing PTF U802490 on the website: <http://techsupport.services.ibm.com/server/hmc/power4/fixes/mcode/v3r30.html>  
You can Install the corrective service file directly from the Internet. For detailed instructions see the Readme "view" link.

2c **Reboot all HMCs to complete SNM software update.**  
Note: Do not reboot HMCs until the SNM\_Serv\_Pack12.zip has been successfully installed on all HMCs.

2d After the HMC is rebooted, follow these steps to verify that the corrective service update was successful:  
Select Switch Network Management from the Switch Management folder in the Navigation area.  
The Status line, the last line on the Switch Network Management panel, should show:  
STATUS RPM Version: 1.3.3.0  
Management Properties > Management > 'SNM Version' column will show:  
IBMhsc.SNM-1.3.3.0-1

NOTES:

Refer to the HPS Guide for more details on the Switch Network Manager:

Review - Chapter 4. System management components; Step 1: Enable SNM

Review - Appendix E. Switch Network Manager (SNM) - The SNM Graphical User Interface

The SNM GUI does NOT update its view automatically. You MUST refresh the display via the GUI menu "Reload" button - the 5th icon from the left.

OR

from the "Menu">"View">"Reload" function

OR

simply press the 'F5' key

[HPS/SNI LPP Installation Guidelines](#)

3	Install HPS/SNI LPP Software. Level Check: devices.chrp.IBM.HPS.rte                      devices.chrp.IBM.HPS.rte                      1.1.2.2
---	---

3a Apply HPS/SNI LPP Base 1.1.2.2 fileset images to LPARs.  
(see "AIX LPPs" in Section 3: Update/Download Information)  
Leave SNM Software disabled.

To verify that the service levels for HPS/SNI LPP's are current for Service Pack 12, on each logical partition issue:

```
lslpp -Lc | egrep "HPS|sni" | cut -d : -f 2,3 | sed 's:/ /'  
or  
dsh "lslpp -Lc | egrep \"HPS|sni\" | cut -d : -f 2,3 | sed 's:/ /' "|dshbak|more
```

Refer to Section 3 - Detailed LPP Level Check for correct levels.

If you are upgrading from Service Pack 6 or lower:

**do not reboot logical partitions (LPARs). See note 1.**

Go to Step 3b

If you are upgrading from Service Pack 7 or higher:

Verify 64 bit kernel is currently in use. Verify the TLP settings. **See Note 2.**

Reboot LPARs. You should reboot the LPARs as soon as possible to properly integrate the changes and to avoid disruption of current functionality.

Go to procedure # 4 **Install GFW.**

Notes:

It is strongly recommended that users read and become familiar with the items covered in this sections notes: ([HPS/SNI Notes](#) )

1. If you are upgrading from Service Pack 6 or below, do not reboot logical partitions (LPARs) until after procedure #4 "Install GFW" is complete. Rebooting prematurely will generate "phantom" SNI devices! See Problem #1 in the "Known Problems" section of this document for more information.

2. If you are upgrading from Service Pack 7 or higher you should already have set up LPARs to boot the 64 bit kernel and enabled Technical Large Page (TLP) option.  
as described in step 3b.

Verify 64 bit kernel is currently in use (on an LPAR): bootinfo -K  
64

Verify the TLP settings:

```
vmo -a|grep lg  
lgpg_size = 16777216  
lgpg_regions=YYY where YYY is the amount of technical large pages to export  
(Ex: lgpg_regions = 256)
```



soft\_min\_lgpgs\_vmpool = 0

3b IBM pSeries HPS now requires that you set up LPARs with 64 bit kernel and enable Technical Large Page (TLP) option.

After successful installation of HPS Filesets from levels at Service Pack 6 or lower, 64 bit kernel and technical large page support option must be enabled.

**To set up your LPARs with 64 bit kernel:**

- 1) Check which kernel is currently in use: `bootinfo -K`  
a response of "32" is a 32bit Kernel
- 2) `ln -fs /usr/lib/boot/unix_64 /unix`
- 3) `ln -fs /usr/lib/boot/unix_64 /usr/lib/boot/unix`
- 4) Determine which rootvg hdisk contains the boot logical volume (usually hd5).

This hdisk will be your "ipldevice".

a) `lspv |grep rootvg`

hdisk0	009b982332a1f9b8	rootvg	active
hdisk1	009b982332a2321a	rootvg	active

b) `lspv -l hdisk0 |grep hd5`

hd5	1	1	01..00..00..00	N/A
-----	---	---	----------------	-----

(hdisk0 is your ipldevice)

5) Issue: `bosboot -ad /dev/<ipldevice>` (eg. `bosboot -ad /dev/hdisk0`)

6) Reboot: `shutdown -Fr`

7) Verify 64 bit kernel is running after reboot:

`bootinfo -K`

64

@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@

**To setup Large Page Option: ( For configuration details, see Large Page Support in AIX 5L Version 5.2 Performance Management Guide.)**

**Setup Large Page Option using the vmo command for each node or node group:**

`vmo -r -o v_pinshm=1 -o lgpg_size=16777216 -o lgpg_regions=YYY`

where YYY is the amount of Technical Large Pages to export, 256 max

Ex: for a system where `lgpg_size = 16777216`, to export `lgpg_regions = 256`:

`dsh -v[Nn] <nodelist> "echo y|vmo -r -o v_pinshm=1 -o lgpg_size=16777216 -o lgpg_regions=256`

( Use the echo command, because vmo will ask for verification to run bosboot)

The number of TLP depends on customer configuration and relates to the number of windows required for each adapter(sni).

**For Example:** To setup a MAX config:

`dsh -vn <nodelist> "echo y|vmo -r -o v_pinshm=1 -o lgpg_size=16777216 -o lgpg_regions=256"`

Would generate this response:

Setting v\_pinshm to 1 in nextboot file

Setting lgpg\_size to 16777216 in nextboot file

Setting lgpg\_regions to 256 in nextboot file

Warning: some changes will take effect only after a bosboot and a reboot

Run bosboot now?

bosboot: Boot image is 19624 512 byte blocks.

Warning: changes will take effect only at next reboot

**NOTE:** The vmtune sample program is being phased out and is not supported in future releases. It is replaced with the vmo command (for all the pure VMM parameters) and the ioo command (for all the I/O related parameters) which can be used to set most of the parameters that were previously set by vmtune. The -v flag has been added to vmstat to replace the -A flag which display counter values instead of tuning parameters. For AIX 5.2, a compatibility script calling vmo and ioo is provided to help the transition.

To Check that Large Page Option is set:

```
vmo -a|grep lg
```

```
lgpg_size = 16777216
```

```
lgpg_regions = YYY <where YYY is the amount of Technical Large Pages to export>
```

```
soft_min_lgpgs_vmpool = 0
```

---

## **[HPS/SNI Notes](#)**

### **1. Notes on using TLP (Large Page) Settings in an HPC environment:**

It is strongly recommended that users familiarize themselves with TLP basics and configuration options available to them, at this location. Federation switch adapter requires TLP usage and these TLP requirements are documented (see "Here is a formula to calculate the required TLP" in the HPS/SNI LPPs section below ) in a latter section of this document.

You should also consult the section on Large page feature on AIX in

[The AIX 5L Version 5.2 Performance Management Guide](#)

**NOTE:** Users need to be aware of the usage of the LoadLeveler pre-emption features with TLP (Large Pages) - specifically the fact that jobs that are using TLP that are pre-empted will essentially "lock up" the real memory the TLP's use, which is pinned by AIX. Unwise use of TLPs with LoadLeveler pre-emption can result in exhausting real memory available for jobs. If one wants LoadLeveler to schedule jobs based on the availability of large page, (especially if the job is going to run in mandatory Large Page mode) he/she may consider making use of the LoadLeveler consumable resource feature. The consumable resource function, which has already been around for several years, is documented in the LoadLeveler manual.

Here is a formula to calculate the required TLP

The following section is a cut and paste ksh script that will run on any Federation node.

The resulting output is a plug in number for the above mentioned Large Page Option -

vmo -r -o v\_pinshm=1 -o lgpg\_size=16777216 -o lgpg\_regions=YYY

where YYY is the number of Technical Large Pages to export, 256 max

```
## ## -----
## ## -----
#
# Formula to calculate required TLP (technical large page)used by HPS / SNI code.
#
#
# The calculation for setting Large Page presumes that:
# Each Page will be 0x1000000 bytes (16777216)
# There are 8 non-user windows per adapter
# Adapters do not share Large Pages
#
# To Calculate the number of LPs we need 4 quantities:
# number of TLP required = A + B + C + D
#
# o A.) first quantity is the device driver global page:
# driver_lp = 1;
#
# o B.) second quantity is the recv FIFOs for each adapter:
# one LP for each windows recv FIFO.
# recv_lp = (num_win) # total number of windows per node
#
# o Quantities C and D are the send_lp value - the send FIFOs for each adapter:
# C.) The IP window has 1 LP for it's send FIFO - 1 for each adapter.
# send_ip_lp = $num_adpt # 1 Large Page per adapter
#
# D.) The rest of the windows have send FIFO buffers allocated to a TLP -
# At least one Large Page per adapter -
# Plus an additional Large Page if there are more than 56 windows.
# send_win_lp = 1 + ( windows / 56 ) # Per adapter
# Quantities C and D are the send_lp value:
# send_lp = send_ip_lp + send_win_lp # Per adapter
# send_lp = 1 + 1 + ( windows / 56 ) # Per adapter
# send_lp = 2 + ( windows / 56 ) # Per adapter
#
# Therefore the TLP calculation is:
# driver_lp + recv_lp + send_lp
```

```

# 1 # per node + (num_win #total per node ) + (2 + ( num_win / 56 ) #Per adater )
#
#
## ## num_windows, send pool size, and receive pool size: are attributes whose
## ## values may be obtained with lsattr -El sniX
## ## (where X is the device minor number; 0, 1, 2, etc.)
## ##
## ## -----
## ## -----

# get num_adpt and num_windows with lsattr -El sniX
for sni in `lsdev|grep sni[0-8]|awk '{print $1}'`; do
    _num_win=`printf "%d" `lsattr -El sni0 -a num_windows -F "value"``
    (( _num_win = $_num_win + 8 )) # Per adater
    (( send_lp = $send_lp + 1 + ( _num_win / 56 ) )) # D = total send_lp per adapter
    (( num_win = $num_win + $_num_win )) # total number of windows
    (( num_adpt = $num_adpt + 1 )) # total number of adapters
done

# o Quantity A = device driver global page: driver_lp = 1;
(( A = 1 ))

# o Quantity B = rcv FIFOs per adapter rcv_lp = (num_win);
(( B = num_win ))

# o Quantity C = IP window : send_ip_lp = $num_adpt # 1 Large Page per adapter
(( C = num_adpt ))

# o Quantity D = send windows - 1 per adapter + more if windows > 56 :
# send_win_lp = 1 + ( windows / 56 ) # Per adapter

(( D = send_lp ))

# Total =
(( TLP = A + B + C + D ))
echo "Total Number of Adapters=$num_adpt; Total number of windows=$num_win!!"
echo "A($A) + B($B) + C($C) + D($D): $TLP"
tot_lp=$TLP
printf "Total required LPs == 0x%x [%d]\n" $tot_lp $tot_lp

```

## 2. Notes on tuning Virtual Memory Settings in an HPC environment:

Customers should be advised that the AIX VMM parameters (set by the vmo command) minfree

and maxfree will most likely have to be adjusted (increased) in an HPC environment based on your cluster size, the amount of system memory, and number of processors per CEC. These settings when tuned properly will ensure enough memory remains available for core cluster infrastructure applications (RSCT, GPFS, LL). The recommended initial value for these tunables are minfree = 10000 and maxfree = 12000. Users are strongly urged to consult the following AIX documentation on virtual memory and vmstat tools and tune their system accordingly.

[http://publib16.boulder.ibm.com/doc\\_link/en\\_US/a\\_doc\\_lib/aixbman/prftungd/memperf.htm](http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/prftungd/memperf.htm)

[http://publib16.boulder.ibm.com/doc\\_link/en\\_US/a\\_doc\\_lib/aixbman/prftungd/memperf1.htm#i50853](http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/prftungd/memperf1.htm#i50853)

The AIX 5L Version 5.2 Performance Management Guide (SC23-4876-00, May 2004) should also be consulted.

NOTE: Tuning these settings have been shown to help users avoid conditions where core cluster applications shut down and restart due to extensive blockage caused by "out of memory" issues. Keep in mind that all cluster applications should be designed and cluster tuned accordingly as to avoid oversubscribing to the real memory available.

## **GFW Installation Guidelines**

- 4 Install GFW.  
(using the recommended AIX command line (update\_flash) method with a locally available image)  
Level Check: ROM Level (alterable).....3H050215 (P690)  
ROM Level (alterable).....3J050215 (P655)

### **Important Preliminary Notes:**

Notes on updating GFW code (system firmware) from the AIX command line

Document Reference: pSeries High Performance Switch Planning, Installation, and Service.

For "Code load requirements for existing p690 and p655 server frames" please see the section in Chapter 6 titled "Step 3: p690 GFW code load" or "Step 3: p655 GFW code load" respectively in the HPS guide.

To understand how to connect the rs422/rs232 cables see the HPS Guide:

Chapter 6; Step 6, "Install the Hardware Management Console (HMC)" thru Step 16. "Verify Installation is Complete"

For "Code load requirements for existing server frames" see Chapter 6

For p655 "Code load requirements for existing p655 server frames"

For p690 "Code load requirements for existing p690 server frames"

For each CEC on which you want to install the GFW code -

One partition running AIX must have "Service Authority" set. Linux does not support microcode download at this time.

The "Service Authority" is set on one LPAR per CEC in the LPAR's profile "other" tab. This designates the LPAR as authorized to provide update images to the CSP. All partitions except the one with "Service Authority" must be shut down. The partition with "Service Authority" must own the device from which the microcode update image will be read.

It is also recommended that the partition with "Service Authority" have a hard disk. If the required devices are not in the partition with "Service Authority", the customer or system administrator must reassign the appropriate resources to it. This requires rebooting the partition with "Service Authority".

If the firmware on a full system partition is being updated, no special steps are required to perform the firmware update using the service aid.

Ensure the GFW image file is not corrupted/truncated before you begin the update\_flash process. Check that /var and /tmp directories are not above 50% full on the partition with the service authority.

The update process can range from 20 minutes to 2 hours, depending on system configuration. The system reboots itself during the update process. Since SNM is disabled during this process, the SNI adapter interfaces will NOT be configured, or will be incorrectly configured when the LPAR(s) reactivate.

It is recommended that you use the ``update_flash -f`` command as opposed to the shutdown -Fu method.

AIX APAR IY49146 is required for update\_flash to work correctly. Level Check by running this command on the partitions:

```
instfix -ik IY49146
```

The ``update_flash`` command will reboot the CEC(s) and will activate the LPAR(s). You may find some more detailed instructions provided on the website with the latest image: <http://techsupport.services.ibm.com/server/mdownload2/download.html>

To install GFW update using Diskette method:  
For p690 systems follow the instructions in the HPS guide on "Step 3: p690 GFW (system firmware code load)" in chapter 6.

To install GFW update using NIM method:  
For a p655 CEC via NIM, follow the HPS guide Chapter 6, "Code load requirements for existing p655 server frames", Step 3. GFW (system firmware) code load.

**Installation steps:** (using the recommended AIX command line (update\_flash) method)

- |    |  |
|----|--|
| 4a | SNM Software should be still disabled. Verify from the SNM GUI Panel.  |
| 4b | For each CEC on which you want to install GFW code, shutdown all partitions except the one with service authority. |

4c	<p>Install the appropriate GFW driver on each CEC to be upgraded:  GFW 3H050215.img (p690) or  3J050215.img (p655):</p> <p>On the AIX partition with Service Authority:  Copy the GFW firmware update code (3x050215.img file) to /tmp  Where x = H for p690 or J for p655.  Enter the following command:  /usr/lpp/diagnostics/bin/update_flash -qf /tmp/3x050215.img  The system will apply the new firmware, reboot, and return to the AIX prompt.  If you use dsh to invoke update_flash then use the -q flag so it does not put out a prompt.  Eg. dsh /usr/lpp/diagnostics/bin/update_flash -qf /tmp/3x050215.img  more conveniently: (the '?' represents a single character 'wild card' and will select either /  tmp/3H050215.img or /tmp/3J050215.img. Dont have both on the same LPAR.)  dsh [-N &lt;nodegrp&gt;] "echo \"'/usr/lpp/diagnostics/bin/update_flash -qf /tmp/3?050215.  img\" at now"  job root.1102366157.a at Mon Dec 6 15:49:17 2004  job root.1102366120.a at Mon Dec 6 15:48:40 2004  job root.1102366242.a at Mon Dec 6 15:50:42 2004</p>
4d	<p>After the LPAR(s) are 'Running', Power OFF the CEC(s) from the GUI or using CSM rpower -  not from EPO red switch.</p>
4e	<p>Once the CEC(s) are powered off - follow these steps to enable the SNM/FNM Software from the  GUI:</p> <ol style="list-style-type: none"> <li>1. Select Switch Network Management from the Switch Management folder in the Navigation  area.</li> <li>2a Select Logical Topology on a 'per plane' basis.  Select the number of Planes and select the Logical Topology on a per plane basis (Number of  Endpoints on a plane) .  Ex: 2 Frames and 2 Switches with NO Switch-to-Switch Links is 2 planes; 1NSB_0ISB_16EP  (16 Endpoints)  Ex: 2 Frames and 2 Switches with ANY Switch-to-Switch Links is 1 plane; 2NSB_0ISB_32EP  (32 Endpoints)</li> <li>2b Enable SNM Software for Normal Operation  OR  2b Enable SNM Software for Switch Network Verification.</li> <li>3. After selecting one of the Enable SNM Software tasks, select the Management Properties task.  On the Management tab, the column labeled "SNM Version" should show:  "IBMhsc.SNM.1.3.2.0-1".</li> </ol> <p>NOTES:</p>



The "Enable SNM Software for Switch Network Verification" task is used during new system setup/installation or after reconfiguration to initialize and check out the system by NOT removing bad links or routes. The main purpose is to discover cable mis-wires and defective hardware. No Routing is modified or loaded. It is NOT intended to be used in a working environment.

Refer to the HPS Guide for more details on defining the Switch Network Topology:

Review - Chapter 4. System management components; Step 1: Enable SNM

Review - Appendix E. Switch Network Manager (SNM) - The SNM Graphical User Interface

The GUI Does NOT update its view automatically. You MUST refresh the display via the GUI menu "Reload" button

∨ ( the 5th icon from the left)

OR

from the "Menu">"View">"Reload" function

OR

simply press the 'F5' key

4f Power up CEC(s) from the HMC GUI and activate logical partition(s).

4g Enable technical large page support - Required for levels greater than Service Pack 6.  
Note: Refer to "HPS/SNI" sub-section in the "Installation Guidelines" section for details on technical large page setup.

#### 4h **Determining the level of firmware on the processor subsystem**

Firmware level is indicated as: 3xyymmdd.img; where

x = a firmware designation such as J or H. J = p655 (Regatta IH series), H=p690 (Regatta H series)

yy = year, mm = month, and dd = day of the release.

Check the GFW microcode level from a VTERM to the main SP Menu. This should show the correct level 3x040901 on the top line. You can also check the GFW level from the AIX command line on the active LPAR(s):

```
#lscfg -vp | grep alter | grep "\.3"
```

You should see:

ROM Level (alterable).....3H050215 - or -

ROM Level (alterable).....3J050215

4i **Determining the level of HPS adapter microcode**  
 The HPS adapter microcode (ucode) is shipped as part of the GFW update image.  
 To Level Check the ucode:  
 From AIX partition, issue: /usr/sni/aix52/debugtools/sni\_get\_ucode\_version -l sniz  
 where z = sni interface number on your system anywhere from 0 thru 7 (Eg. sni0) which can be seen in  
 "netstat -in" output. The timestamp should show build date: 09/1/04  
 NOTE: You will need to reinstall the GFW update if you:  
 - neglected to disable SNM during the GFW update  
 - or if you added/replaced an HPS adapter. Otherwise, the HPS ucode may not have been applied correctly.  
 Level Check the ucode as stated in the previous bullet to make sure you do not have to reinstall GFW.

### [Power Subsystem Microcode Installation Guidelines](#)

5 Install Power Subsystem Microcode on each frame by following these steps.  
 Level Check: 26a7

5a Power Subsystem Microcode: ptcodes-1.74.26a7-1.i386.rpm  
<https://techsupport.services.ibm.com/server/mdownload/other.html>  
 Install via the HMC GUI through the Software Maintenance -> Frame panels.  
 Receive Corrective Service  
 Install Corrective Service  
 If you are upgrading from Service Pack 9 or above, the installation of Power Subsystem Microcode: ptcodes-1.74.26a7-1.i386.rpm is complete.  
 Proceed to step 5d to level check installed version.

5b If you are upgrading from Service Pack 8 or lower, the new Power Subsystem Microcode requires the switches to be recycled for the changes to become effective.  
 Power cycle switch(s):  
 From the HMC GUI select "Switch Network Management > Switch Topology View"  
 For each switch plane  
 - select "Selected > Power <Off"  
 - refresh GUI to verify power status  
 - select "Selected <Power <On"  
 - refresh GUI to verify power status  
 Repeat procedure on all switch planes.

5c Recycle SNM daemon using the HMC GUI:  
 select Switch Network Management > Disable SNM Software  
 refresh display: Menu > File > Refresh  
 >Enable SNM Software.

Notes:

1. Recommended procedure for a complete power-cycle of the cluster ("EPOW") - a.k.a. Cold Boot Procedure

In the event that power-cycling the HPS switch boards (5b) is not enough

- as a last resort after trying 5b and failing - the following procedure is suggested:

For a complete discussion of the subject :

See pSeries High Performance Switch Planning, Installation, and Service; Chapter 9. Service procedures; Managed system power on and power off (LPAR reboot); p. 170)

The 10 minute wait is for switch frames to stabilize.

Suggested Procedure:

- shutdown all the lpars
- rpower -a cec\_off
- stop fnmd
- EPOW off all the CECs and the ISB frame
- start fnmd
- EPOW on the ISB frame and the NSB frames
- wait 10 minutes, EPOW everyone else on
- check for flashing lights on the (optical) risers in the ISBs
- rpower -a cec\_on
- tail /var/hsc/log/\*Init.log until that stops
- run hps\_check.pl
- activate all the lpars

5d Level Check - After completing the pcode installation, go to the HMC GUI and verify successful installation:

- Select: Software Maintenance --> Frame --> Install Corrective Service
  - Verify that the "Installed Version" matches the version you just installed - 26a7
- (Important: This window may not automatically refresh when installation completes. Manually refresh the window as necessary. )

5e

**Verify HPS is Functioning.**

At this point the basic HPS installation is complete. You should now be able to ping over the switch.

A general check is a good ping all script. The HPS Documentation covers this topic in the host based verification tool.

- Post Installation Task 2. ( pSeries High Performance Switch Planning, Installation, and Service; "Run the host-based verification tools", p116)

If SNM Software is not running then :

The "Switch Management" | Switch Network Management | views : Switch Topology View, and End Point View will not be populated and there is a dialog message indicating that and you will not be able to ping over the switch.

hps\_check.pl is not available in a closed box without the root password.

It should show the links as Timed and MPA Available :

Lpar Name	Lpar#	Sni# =>	Adapter#	Csp#	Cronus# =>	Frame	Cage	Chip	Port :
Timed?	MPA	TOD							
c661f1rp02	1	0	0 2	5	1 1 5 2	YES	YES	SLV	
c661f1rp02	1	1	1 3	4	4 4 5 2	YES	YES	BAK	

**[AIX Installation Guidelines](#)**

6

Install AIX base updates on each node by following these steps.  
The recommended AIX service level for AIX 5L version 5.2 is:  
5200-05 Recommended Maintenance package.

Document Reference: AIX 5L Version 5.2 Installation Guide and Reference (SC23-4389-03)

Level Check: bos.mp64                      5.2.0.54

6a

Download and install the recommended AIX 5L version 5.2 Maintenance Package 5200-05

To download:

Go to <http://www-1.ibm.com/servers/eserver/support/pseries/aixfixes.html>

Select Maintenance Packages and choose AIX52

Specify your "Current level"

Specify "Desired level": 5200-05 and click "GO"

Follow instructions on this page for downloading the gzip file.

6b

Level check AIX by running this command on the logical partition(s): " lslpp -ha bos.mp64 "  
and verify that the /usr part of the fileset is at or above 5.2.0.54

**[AIX LPP Installation Guidelines](#)**

7	<p>Install AIX LPP updates on each node by following these steps.  <a href="#">Level Check: see "Detailed LPP Level Check"</a></p> <p>Notes: CSM LPP's need to download and install openCIMOM-0.8-1 RPM update.  Document Reference: AIX 5L Version 5.2 Installation Guide and Reference (SC23-4389-03)  IBM RSCT: Administration Guide</p> <p>As noted in the Installation Guide, in order to use CSM LPP's, you need to download and install openCIMOM-0.8-1 RPM update.  You can download openCIMOM from the following web site:  <a href="http://www-1.ibm.com/servers/aix/products/aixos/linux/download.html">http://www-1.ibm.com/servers/aix/products/aixos/linux/download.html</a></p> <p>Select "Package" OpenCIMOM "Version" 0.8 (5.2) (For AIX 5.2)</p> <p>For complete details see: Chapter 4. Installing the management server; Step 6. Download Open Source Software of CSM Guide.  Document Reference:  IBM Cluster Systems Management for AIX 5L Planning and Installation Guide  AIX 5L Version 5.2 Installation Guide and Reference  IBM Reliable Scalable Cluster Technology Administration Guide</p> <p>DPCL is no longer a part of the IBM PE for AIX licensed program.  Instead, DPCL is now available as an open source offering that supports PE.  For more information and to download the DPCL open source project go to:  <a href="http://oss.software.ibm.com/developerworks/opensource/dpcl">http://oss.software.ibm.com/developerworks/opensource/dpcl</a></p> <p>Document Reference: IBM Parallel Environment for AIX 5L Installation Version 4 Release 1.1 (1.1 GA22-7943-01)  Chapter 1. Introducing PE 4.1.1</p>
7a	<p>Download and install the applicable IBM Virtual Shared Disk(VSD) , LAPI, HPS, PPE, LoadLeveler, GPFS, Parallel ESSL, ESSL, CSM and RSCT PTF updates.  Apply the listed APARs which are needed for HPS Service Pack 12. They are available from the website:  <a href="http://www-1.ibm.com/servers/eserver/support/pseries/aixfixes.html">http://www-1.ibm.com/servers/eserver/support/pseries/aixfixes.html</a></p> <p><b>If</b> you are upgrading to SP12 from a service pack that is earlier than SP9, then installing CSM1.4 is required.  The recommended AIX service level AIX 5L version 5.2 is ML 5200-05.  The Maintenance package contains CSM 1.4 which requires RPM update openCIMOM 0.8(5.2).</p>
7b	<p>To verify that the service levels for your LPP's are current for Service Pack 12, on each logical partition issue:  <pre>lspp -Lc   egrep "vsd LAPI HPS sni ppe LoadL mmfs rsct csm essl pessl"   cut -d : -f 2,3   sed 's:/ /'</pre></p>

7c **RSCT Migration Issues:**

[See Problem 2: "hagsglsm is not reporting the local switch membership group"](#)

For more information, refer to RSCT Administration Guide; Chapter 3. Creating and Administering an RSCT Peer Domain; Migration

>Section 6: Known problems/workarounds/Restrictions

Problems:

1. "Phantom" SNI devices may appear after upgrade from pre-Service Pack 7

**Users Affected:** Users upgrading systems from SP6 or earlier with existing SNI devices

**Problem Description:**

Changes in the HPS switch microcode and driver demand that it is absolutely necessary to install the HPS fileset updates and the GFW firmware updates without rebooting logical partitions before the system firmware is successfully installed. Rebooting the LPAR(s) prematurely will cause "phantom" SNI devices to be created on the partition(s).

The failure signature is:

- 1) There are twice the number of snX and sniX devices as expected in the "lsdev -C|grep sn" output. (where X = sni or sn interface number on your system anywhere from 0 thru 7 -- e.g. sni0 or sn0)
- 2) All the sn interfaces are in the Defined state.
- 3) Upper half of the sniX devices are in "Available" state (higher numbered devices) and lower half of the sniX devices are in "Defined" state (lower numbered devices).
- 4) All the sniX devices could also be in the "Defined" state.

**Common causes of "phantom" sniX devices:**

If you re-boot LPARs after upgrading HPS filesets, but prior to a successful GFW upgrade.

Failure during GFW upgrade after updating HPS filesets. These events/scenarios will cause all LPARs on a CEC to reboot without upgrading the GFW via AIX command line method (i.e. update\_flash command)

- If you have a corrupted/truncated GFW image file (3H050215.img or 3J050215.img)
- If either /var or /tmp is too full.
- If an LPAR other than the Set Service Authority LPAR is in "Running" state during update\_flash
- If an LPAR other than the Set Service Authority LPAR is used to run the update\_flash command

Note: There may be other factors that cause LPARs to reboot after the HPS fileset is upgraded and before the GFW is successfully updated.

## Recovery Procedure:

- 1) Complete the GFW update.
  - Verify that the firmware updates on all CEC(s) was successful.
- 2) Recover the sniX and snX devices. (Note: To recover requires at least 1 reboot. Two reboots are required if the sniX devices are busy)

2a) Make note of any customization to the SNI devices (e.g. num\_windows, driver\_debug, etc ...) using the command "lsattr -El sniX" so that they can be re-applied after the recovery procedure. The procedure will reset all values back to the defaults.

Remove all the SNI devices:

```
for each logical sniX; # e.g. for X in 0 1 2 3;
do
    rmdev -d -l sni$X
done
```

If the rmdev fails for any devices (e.g. device is busy), then unconfigure the device driver as follows. Otherwise, go to step 2b.

- rename the configuration method for the device: mv /usr/sni/aix52/cfgsni /usr/sni/aix52/cfgsni.orig
- reboot each LPAR that failed
- run the rmdev loop again
- restore the original configuration methods name: mv /usr/sni/aix52/cfgsni.orig /usr/sni/aix52/cfgsni

2b) Remove ONLY the top half of the snX devices:

The lower half (original half ) have the ipaddr and netmask attributes in the odm. You don't want to delete these, nor do you have to.

```
for UPPER HALF of logical snX # e.g. for X in 2 3;
do
    rmdev -d -l sn$X
done
```

- 3) Reboot the LPAR(s) --> shutdown -Fr
- 4) Restore any customization to the SNI devices (e.g. num\_windows, driver\_debug, etc ...)



2. "hagsglsm is not reporting the local switch membership group

**Component:** rsct

**Systems Affected:** All rsct Users at Service Pack 9

**Description:**

In order to complete the migration of a peer domain and update the active RSCT version to a new level, you must enter the runact command:

```
runact -c IBM.PeerDomain CompleteMigration Options=0"
```

This command should be run after every RSCT release upgrade.

For a more complete discussion see: "Avoiding Domain Partitioning When Migrating From RSCT 2.2.1.x or 2.3.0.x" in

[Reliable Scalable Cluster Technology \(RSCT\) Library](#): Administration Guide SA22-7889-04; Chapter 3. Creating and Administering an RSCT Peer Domain; Migration

**Component:** rsct

**Systems Affected:** All rsct Users at Service Pack 9

**Description:**

In order to complete the migration of a peer domain and update the active RSCT version to a new level, you must enter the runact command:

```
runact -c IBM.PeerDomain CompleteMigration Options=0"
```

This command should be run after every RSCT release upgrade.

For a more complete discussion see: "Avoiding Domain Partitioning When Migrating From RSCT 2.2.1.x or 2.3.0.x" in

IBM Reliable Scalable Cluster Technology Administration Guide SA22-7889-04; Chapter 3. Creating and Administering an RSCT Peer Domain; Migration

**Restrictions:**

1. **No switch should be powered off while the SNM software is running.**

**Component:** SNM - Switch Network Management

**Systems Affected:** High Performance Switch (HPS) users applying from pre-Service Pack 7

**Implications:**

- A CEC frame with a switch in it CANNOT be EPOWed.
- If a CEC has to be power cycled, power down the CEC and not the frame.
- If a CEC frame with a switch needs to be EPOWed, power down the frame, kill the SNM daemon after 5 minutes, power up the frame and restart the SNM daemon.
- If one or more switches need to be recycled, power down the switches, power them back up and then recycle the SNM daemon after 5 minutes.

## 2. Rules for swapping cables for fault isolation

**Component:** SNM - Switch Network Management

**Systems Affected:** All HPS Users applying from pre-Service Pack 7

### **Description:**

Only swap SNI attached cables at the switch ports to which they are attached.

Do not swap switch to switch cables with other switch to switch cables nor with SNI attached cables.

Do not swap cables on the SNI ports.

Do not swap cables between switches.

If adapters are accidentally "miswired" during the process of swapping cables, recable the adapters to their original positions.

## 3. Improved performance is more sensitive to bad links

**Component:** HPS/LAPI

**Problem Description:** The protocol (MPI or LAPI) will timeout if the job runs on bad links and the link routes are not fixed. If the link failure turns into adapter failure then the job gets terminated.

**Solution:** To resolve this issue monitor Service Focal Point for bad links and fix them.

## 4. HPS Cluster recommended LPAR reboot procedure

**Component:** HPS/SNI

**Systems Affected:** All HPS Users applying from pre-Service Pack 7

### **Description:**

To ensure the HPS switch links are properly shutdown and re-enabled, it is recommended that you use the commands below, to recycle and reboot all LPAR/AIX images in your cluster:

`"shutdown -F" <-- to shutdown LPAR`

`"shutdown -Fr" or "dsh -av shutdown -Fr" <-- to reboot LPAR`

or multiple LPARs simultaneously

Use of the "reboot" command or "rpower" commands will not shutdown the HPS switch links in an orderly fashion - when more than one frame at a time is cycled concurrently the SNM daemon may hang and Service Focal Point could end up with artificial errors. If/when these commands must be run concurrently on multiple LPARS, it is recommended that you use them one frame at a time in your cluster.

When the HMC gui is used to cycle an LPAR, it is recommended that you use the "shutdown" option to recycle the LPAR, not the "reset" option. The "shutdown" option will ensure that the HPS

switch links are shutdown and re-enabled cleanly - whereas the "reset" option(s) will essentially use the rpower command (and not do an orderly shutdown).

NOTE: Use of the rpower or HMC GUI reset options should be reserved as a "last resort" for LPARS that are not responding to a shutdown command.

## 5. Striping mode Restrictions

### Component: HPS/SNI

#### Description:

- Striping currently limited to two links per task. Multiple tasks can use all links as long as there are more tasks than half the links available.
- Performance degradation for single task per link of 3% for unidirectional and 5% for exchange bandwidth at large messages. However, multiple tasks per link get full link bandwidth.
- In non-striping mode you only get half the links on a p655 system using a 2 plane configuration when setting MP\_EUDEVICE=sn\_single. You must use MP\_EUDEVICE= csss or sn\_all to get all links.

>Section 7: HPS Service Pack 12 fix list (by component)

### LAPI Abstract: ( APAR IY65386 )

Unknown LAPI\_term rc -1 mpi+lapi pgm and prot

Kill job on SIGDANGER

INTERNAL LAPI USES FLOATING POINT ADDS WITH

LAPI USES EXTRA MEMORY AND CREATES ZOMBIE TH

### Loadleveler Abstract: ( APAR IY65389 )

writing LL account data does not resume after

CUSTOM\_METRIC\_COMMAND MAY INHERIT THE OPEN S

INCONSISTENCY IN LLCTL RESUME

turn on sticky bit for WLM directory created

WLM AND LOADL CALC CONSUMABLE MEMORY DIFFERE

LLCLASS -L OUTPUT, CHECKPOINT\_DIRECTORY TO B

USR CRED EXPIRE=USR EPILOG NOT RUN DUE TO SE

SOFT WALLCLOCK PREMATURELY HIT/100 DAYS HITS

NEED EASIER WAY TO DEBUG LOADL\_SCHEDD SIGABR

LoadLeveler needs to recognize additional adapters

WLM CLASSES INTERMITTENTLY NOT DELET

**Parallel Environment Abstract: ( APAR IY65388 )**

PE BENCHMARKER SESMGR FAILS ASSERT IN HANDLE

fine grain dgs\_rm failure

MPI 8k task job support

DEFECT SHOWN IN MPI\_TYPE\_STRUCT() AND/OR MPI

MPI\_GET NOT USING DISPLACEMENT CORRECTLY

INTERNAL LAPI USES FLOATING POINT ADDS WITH

**IBM Virtual Shared Disk Abstract: ( APAR IY65385 )**

NONE

**GPFS 2.1 Abstract: ( APAR IY65390 )**

ability to set environment var for da

assert:oldDiskAddrP==NULLORoldDiskAddrFound m

structure error:LogFSError StripeGroup LLOpen

envvar: allow quoted strings

Fix envvar quoting

SIGNAL 11 IN FLUSHDEALLOCQUEUE USING ALMOST 1

DEADLOCK ON QUOTA FILE IN INTOPENFILE

MMAUTOLOAD FAILS IN ALL BUT LC AND SINGLE NO

MMFENCENODEIN USING DELETED NODES

## **GPFS 2.2 Abstract: ( APAR IY63035)**

### **GPFS 2.2.1 supported with HPS Service Pack 7 or higher**

With the release of HPS Service Pack 7, customers will now be able to use GPFS 2.2 in their HPS environment.

The recommend level of service required for GPFS 2.2.1 on an HPS system is IY63035

failover from local devices to server needs message

GPFS Daemon assert during unmount - line 789 of file quallo

error messages vary greatly between aix and linux gpfs

mmgetstate takes 5.5 minutes to complete after node failure

c462f1rp01 hangs ... rsh/rlogin/telnet fails ping OK

double [common]'s in mmlsconfig output

assert:err != E\_OK OR (whichFlags metadata.C 12164

SLOW RANDOM READS USING MMAP OF A LARGE FILE

mmpmon connection

mmpmon > rhist nr allows 0 to 0 size range and latency range

signal11:StripeGroupDesc::operator==(StripeGroupDesc const&)

TSBACKUP DOES NOT HANDLE FILE PATHNAMES LONGER THAN 256

cs2b: linux nfs filesystem: ls: root: Input/output error

signal11:LLCachedRecAddr::isMigrating during mmfsck -o

Assert: `(q == 0)

RECLOCK RETRY FAILS FOR NON-ROOT USERS

NFS on cs2b tar causes many Stale NFS file handle

assert:!"I/O completion: Unexpected disk availability after qui

IDE DRIVES (/DEV/HD\*) MISSED DURING DISCOVERY

Handle daemon errors on mmadd/del/rpldisk

mmfsctl syncFSconfig silently does nothing if partial data at t

remote mount NSD recovery tests ends in failure

Remove NOSHIP defines except Opteron

mmpmon connections

Cant unmount fs because of mmgetobjd

mmcrfs failing on 2105 MPIO disks

mmdsm, mmimportfs

missing error message 376

Allow overriding the trace file and trace buffer sizes

mmfsck ends prematurely with 0:5:18 and RepStatusBad

mmfsctl syncFSConfig cannot reach contact nodes in remote clust

NSD redriver thread

Complete 471707 backport for nfsd\_iget

long delays in GPFS cmds if primary server down (pmr 83091)

assert:I/O completion: Unexpected disk availability sched.C 220

IF MMCHONFIG RELEASE=LATEST IS NEVER PERFORMED, THE /VAR/MMFS/

Failure during migrate log

"mismatched pvids" error when vsd server is down

assert !"Attempt to delete an object whose header has been corr  
GPFS\_FGETTATTRS CAUSES KERNEL PANIC WHEN CALLED ON A SPECIAL F  
RANDOM DIRECTORY ERRORS ON OPTERON  
Invalid entries in /var/mmfs/etc/mmfs.cfg  
The assert subroutine failed: in alloc.C, line 1934  
mmpmon ports  
printDentry needs non-blocking trace  
Fix parsing long lines in mmfs.cfg file  
assert:0, file fencing.C, line 871  
smart pointer debugging, GPFS\_HANDLE\_NFS\_FH  
mmtrace perf, non-blocking traces  
MmpmonApplelevel prototype Java code  
Node sendClusterSDRFiles is not available to run the command

### **SNM: Abstract (APAR IY62953)**

LSSWTOPOL COMMAND FAILED WITH -F FLAG.  
jt query dest should only show BPAs and switc  
Decrease the debug level in testRoute  
NEW TOOL TO DISPLAY HW REPORTED ERRORS FROM  
Change needed to SNM\_RPM so correct SNM RPM v  
Remove timedate stamp included in Primary blo

### **HPS/SNI: Abstract ( APAR IY63031 )**

Network ID support for mixed Regetta/Squadron  
re-work getsmadata for squadrons - new prd si  
Verification test tool improvements  
investigate tcb support.  
updates for new window fatal condition (rCxt  
Enable odm entry and chgsni support for  
Ship additional debug tools.  
Implement the re-cable flows  
TCEs should be traced at error level  
update UT\_window's window fatal tests  
vmgetinfo failing in UT\_rdma\_ioctl32 and UT\_r  
trace the lmt for window fatals in close wind  
fix exposure in system();  
final UT\_window checkin  
loopback driver and tools  
Alias Support for Federation  
return canp\_fifo\_rdma as css\_dev type on squa  
crashed in curtime+000168

dd needs to kill all bad window procs,  
No adapters 'Available' after rebooting LPAR  
HAL snap can be dumping user data

### **GFW: Abstract**

Refer to the following links for complete list of abstracts for GFW fixes in HPS Service Pack 7 or higher:

<http://techsupport.services.ibm.com/server/mdownload/7040681F.html> - for Regatta H

<http://techsupport.services.ibm.com/server/mdownload/7039651F.html> - for Regatta IH

### **HMC:**

#### **SP12 Abstracts:**

**Need to Package LedCmd library for ledfs.**

**NEW TOOL TO DISPLAY HW REPORTED ERRORS FROM A CENTRAL POINT**

**FNM Collect VPD when it does not need to**

**Link Sync logged in SFP after adapter hit prd threshold**

**new ledcmd library and ledfs command**

**MAN PAGES AND HNC DOC HAVE TO BE UPDATED FOR LSSWTOPOL AND LSS**

**RPM change needed for SP12**