b

IBM Power 770/780 and 795 Servers CEC Hot Add & Repair Maintenance Technical Overview

May, 2011

Authors:

Curtis Eide
Al Kitamorn
Aditya Kumar
Dave Larson
Erlander Lo
Chet Mehta
Naresh Nayar
Jayesh Patel
Adam Stallman

Table of Contents

1 Executive Overview

IBM continues to introduce new and advanced RAS (Reliability, Availability, Serviceability) functions in the IBM Power™ Systems to improve the overall system availability. With advanced functions in fault resiliency, recovery and redundancy design, the impact to the system from hardware failure has been significantly reduced. With these system attributes, Power Systems continue to be leveraged for server consolidation. For customers experiencing rapid growth in computing needs, upgrading hardware capacity incrementally with limited disruption becomes an important system capability.

Concurrent add and repair capabilities for the Power Systems have been introduced incrementally since 1997, starting with power supply, fan, I/O device, PCI adapter and I/O enclosure/drawer. In 2008 IBM introduced significant enhancements to the enterprise Power Systems 595 and 570 that highlighted the ability to add/upgrade system capacity and repair the CEC, i.e. central electronic complex, or the *heart* of a large computer system, which includes the processors, memory and I/O hubs (GX adapters), without powering down the system.

The IBM Power 770, 780 and 795 servers continue to improve on the CEC hot add and repair functions that were introduced with the Power Systems 595 and 570. Best practice experience from client engagements in 2008-2010 calls for a new level of minimum enablement criteria that includes proper planning during system order, configuration, installation, I/O optimization for RAS, etc.

This paper provides a technical overview and description of the Power 770, 780 and 795 CEC Hot Add & Repair Maintenance (CHARM) functions. It also includes best practices, minimum enablement criteria and planning guidelines to help the system administrator obtain the maximum system availability benefits from these functions.

With proper advanced planning and minimum criteria being met, the Power Systems 770, 780 and 795 CEC Hot Add & Repair Maintenance (CHARM) function allows the expansion of the system processors, memory, and I/O hub capacity, and their repair as well, with limited disruption to the system operation.

2 Introduction

This section provides a brief overview of the CEC Hot Add & Repair Maintenance (CHARM) functions and the Power 770/780/795 servers that these functions support. It also outlines the roles and responsibilities of the IBM System Service Representative (SSR) and the system administrator to jointly complete an end-to-end CHARM operation. Section 3 provides an overview of the best practices (minimum enablement criteria, partition configuration, I/O availability considerations and redundant I/O options) for the system administrator to obtain the maximum availability benefits from these functions. Section 4 provides planning guidelines and prerequisite for both the system administrator and SSR. Section 5 provides additional details of the CHARM functions. Section 6 provides a brief overview of the user interfaces for the CHARM functions. Section 7 provides additional reference information and terminology used in this document.

2.1 Function Overview

The CHARM functions provide the ability to add/upgrade system capacity and repair the CEC, i.e. central electronic complex, or the *heart* of a large computer, without powering down the system. The CEC hardware includes the processors, memory, I/O hubs (GX adapters), system clock, service processor and associated CEC support hardware. The CHARM functions consist of special hardware design, service processor, hypervisor, and HMC(hardware management console) or SDMC(system director management console) firmware. The HMC/SDMC provides the user interfaces for the system administrator and SSR to perform the tasks for the CHARM operation.

The CHARM functions utilize the existing Resource Monitoring and Control (RMC) connection between the HMC/SDMC and the logical partitions (LPARs) to determine the impact on the partitions' resources (processor, memory and I/O) in preparation for a hot node/GX repair or hot node upgrade operation. The HMC/SDMC communicates with the impacted partitions to perform the DLPAR (dynamic logical partition) remove and add operations during the CHARM operation. The HMC/SDMC communicates with the service processor and hypervisor to perform CHARM tasks such as a system “readiness check”, administrative fail-over (for redundant service processor and clock functions), workload evacuation, hardware deactivation, power off, power on, CEC hardware diagnostics, hardware activation, resource configuration, resource integration and error handling/reporting, during the CHARM operation.

During a hot node repair, I/O hub repair, or node upgrade operation, the node or I/O hub is powered off for safe removal and insertion. Consequently, all resources dependent on the node (CEC drawer for the Power 770/780/795) or I/O hub including processors, memory, internal I/O, and externally attached I/O are not functional during the CHARM operation. The CHARM functions identify the resources that will be impacted, and guide the system administrator through the process of gracefully deactivating and/or freeing up resources as necessary to complete the operation. The CHARM functions ensure that all resources that will be impacted by the operation have been explicitly deactivated by the user or workloads using those resources can be evacuated to available resources elsewhere in the system, before allowing the operation to proceed.

2.1.1 Hot Add or Upgrade

Hot node add: This function allows an SSR to add a node to a system to increase the processor, memory, and I/O capacity of the system. After the physical node installation, the system firmware activates and integrates the new node hardware into the system. The new processor and memory resources are available for the creation of new partitions or to be dynamically added to existing partitions. Additional GX adapters (I/O hubs) can be physically added to the node and activated as part of the hot node add operation. After the completion of the node add operation, additional I/O expansion units can be attached to the new GX adapter(s) in the new node in a separate concurrent I/O expansion unit add operations.

Hot node upgrade (memory): This function allows an SSR to increase the memory capacity in a system by adding additional memory DIMMs to a node, or upgrading (exchanging) existing memory with higher-capacity memory DIMMs. The system must have 2 or more nodes to utilize the hot node upgrade function. Since the node that is being upgraded is functional and possibly running workloads prior to starting the hot upgrade operation, the system administrator uses the “Prepare for Hot Repair or Upgrade” utility, system management tools, and operating system (OS) tools to prepare the node for evacuation. During the hot node upgrade operation, the system firmware performs the node evacuation by relocating the workloads from the target node to other node(s) in the system and logically isolating the resources in the target node. It then deactivates and electrically isolates the node to allow the removal of the node for the upgrade.

After the physical node upgrade, the system firmware activates and integrates the node hardware with additional memory into the system. After the completion of a hot node upgrade operation, the system administrator then restores the usage of processor, memory, and I/O resources, including redundant I/O configurations if present. The new memory resources are available for the creation of new partitions or to be dynamically added to existing partitions by the system administrator using DLPAR operations.

Concurrent GX adapter add: This function allows an SSR to add a GX adapter to increase the I/O capacity of the system. The default settings allow one GX adapter to be added concurrently to a Power 770/780 system and two GX adapters to be added concurrently to a Power 795 system, without planning for a GX memory reservation. To concurrently add additional GX adapters, additional planning is required (refer to section 5.2.1 for more details). After the GX adapter is installed, the system firmware activates and integrates the new I/O hub into to the system. After the completion of the concurrent GX adapter add operation, I/O expansion units can be attached to the new GX adapter through separate concurrent I/O expansion unit add operations.

2.1.2 Hot Repair

Hot node repair: This function allows an SSR to repair defective hardware in a node of a system while the system is powered on. The system must have 2 or more nodes to utilize the hot node repair function. Since the node that is being repaired may be fully or partially functional and running workloads prior to starting the hot repair operation, the system administrator uses the “Prepare for Hot Repair or Upgrade” utility, system management tools and OS tools to prepare the system. During the hot node repair operation, the system firmware performs the node evacuation by relocating workloads in the target node to other node(s) in the system and logically isolating the resources in the target node. It then deactivates and electrically isolates the node to allow the removal of the node for repair.

After the physical node repair, the system firmware activates and integrates the node hardware back into the system. After the completion of a hot node repair operation, the system administrator restores the usage of processor, memory, and I/O resources, including redundant I/O configurations if present.

Hot GX adapter repair: This function allows an SSR to repair a defective GX adapter in the system. The GX adapter may still be in use prior to starting the hot repair operation, so the system administrator uses the "Prepare for Hot Repair or Upgrade" utility and OS tools to prepare the system. During the hot GX repair operation, the system firmware logically isolates the resources associated with or dependent on the GX adapter , then deactivates and electrically isolates the GX adapter to allow physical removal for repair.

After the repair, the system firmware activates and integrates the hardware into the system. After the completion of the hot repair operation, the system administrator restores the usage of I/O resources, including redundant I/O configurations if present.

C**oncurrent system controller repair:** The concurrent System Controller (SC) repair function allows an SSR to concurrently replace a defective SC card. The target SC card can be fully or partially functional and in the primary or backup role, or it can be in a nonfunctional state. After the repair, the system firmware activates and integrates the new SC card into the system in the backup role.

2.1.3 Prepare for Hot Repair or Upgrade Utility

The Prepare for Hot Repair or Upgrade (PHRU) utility is an tool on the HMC/SDMC used by the system administrator to prepare the system for a hot node repair, hot node upgrade or hot GX adapter repair operation. Among other things, the utility identifies I/O resources that will be impacted by the operation, identifies I/O resources that may be in use by operating systems and must be deactivated or released by the operating system, identifies additional processor and memory capacity that must be made available, and identifies system configuration issues that preclude CHARM operations.

Based on the PHRU utility's output, the system administrator may reconfigure or vary off impacted I/O resources using operating system tools, remove reserved processing units from shared processor pools, reduce active partitions' entitled processor and/or memory capacity using DLPAR operations, or shut down low priority partitions.

This utility also runs automatically near the beginning of a hot repair or upgrade procedure to verify that the system is in the proper state for the procedure. The system firmware will not allow the CHARM operation to proceed unless the necessary resources have been deactivated and/or made available by the system administrator, and the configuration supports it.

The table below summarizes the usage of the PHRU utility based on specific CHARM operations and expected involvement of the system administrator and service representative.

CHARM Operation	*Minimum # of nodes to use operation*	*PHRU Usage*	*System Administrator*	*Service Representative*
Hot Node Add	1	No	Planning only	Yes
Hot Node Repair	2	Yes	Yes	Yes
Hot Node Upgrade (Memory)	2	Yes	Yes	Yes
Concurrent GX Adapter Add	1	No	Planning only	Yes
Hot GX Adapter Repair	1	Yes	Yes	Yes
Concurrent System Controller Repair	1	No	Planning only	Yes

Table 1: PHRU utility usage

2.2 System Hardware Overview

This section provides a brief overview of the Power 770/780 and Power 795 system hardware that is supported by CEC Hot Add & Repair Maintenance functions.

2.2.1 Power 770 and 780 Servers

The Power 770 and 780 servers consist of 1 to 4 nodes (also known as building blocks or drawers). Each node has 2 processor modules (each with 6 or 8 cores), 32-512 GB of memory, 2 integrated I/O hubs/bridges with integrated virtual Ethernet (IVE), and 2 slots for GX adapters. In addition the first two drawers each include a service processor card on which the processor clock resides. Each Power 770/780 building block also contains 6 PCI-e adapter slots, 6 disk drive slots, and 1 media bay.

For a multi-node server configuration, processor fabric bus cables and clock/service processor interface cables are required to connect the nodes in a stack.

For additional details, please visit:

http://www-03.ibm.com/systems/power/hardware/770/index.html

http://www-03.ibm.com/systems/power/hardware/780/index.html

2.2.2 Power 795 Servers

The Power 795 server consist of 1 to 8 POWER7 processor books. Each processor book has either 24 or 32 cores that are packaged on four POWER7 processor chips. Each processor chip has 2 MB of on-chip L2 cache and 32 MB of eDRAM L3 cache, and each core supports four hardware threads. Each processor book also contains 32 DDR3 memory DIMM slots, support for up to four GX adapter (I/O hub) cards with 12x interface for connection to external I/O drawers, and two node controllers (NC) that serve in primary and backup roles for redundancy.

For Additional details, please visit:

http://www-03.ibm.com/systems/power/hardware/795/index.html

2.2.3 Logical View: Multi-Node System

Figure 1 provides a logical view of a multi-node system like the Power 770, 780 or 795. The figure shows the system configuration with redundant VIOS (Virtual I/O Server) and redundant I/O adapters to improve I/O availability and reduce the impact of a hot node repair or upgrade operation. Note that unlike the 795, Power 770 and 780 servers have PCI-e and disk I/O in the CEC nodes as well as GX adapters for external I/O expansion drawers. Thus, while the figure is accurate for both 770/780 and 795 servers, for 770/780 it would be possible to have the PCI-e adapters in the CEC nodes instead of, or in addition to, external I/O drawers.

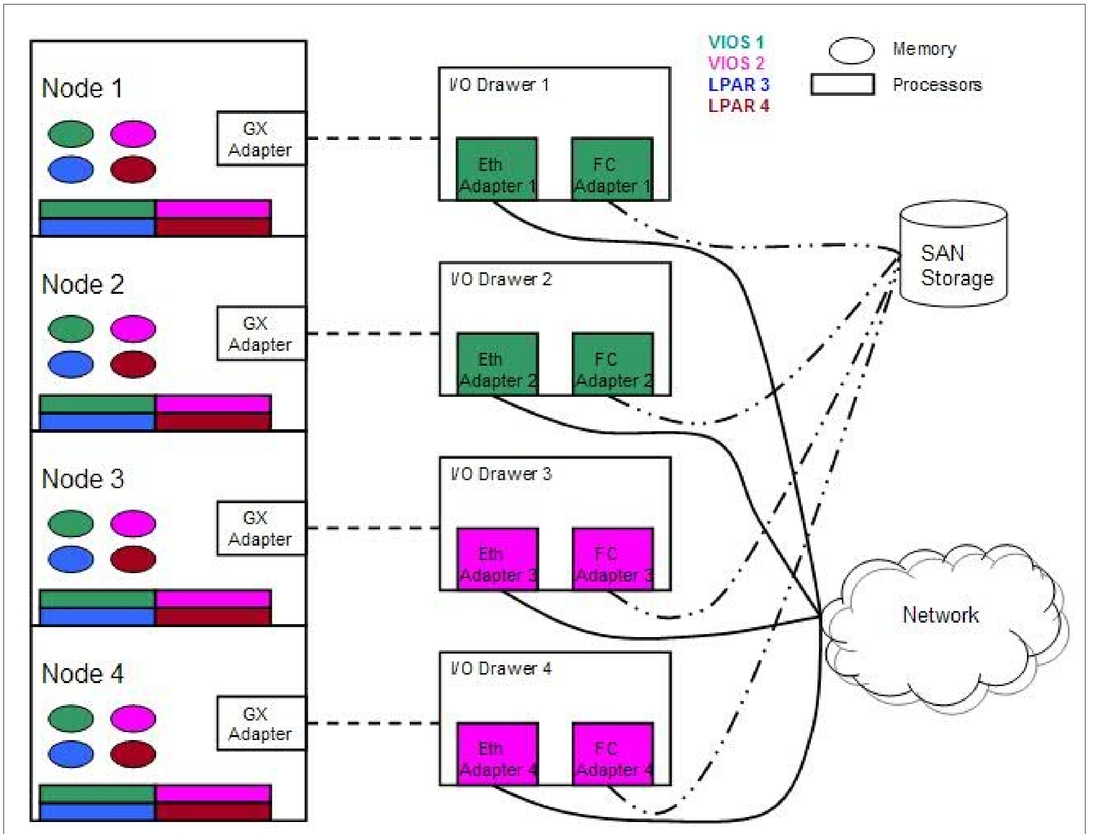

Figure 1: Multi-Node System

2.3 System Administrator Roles and Responsibilities

This section provides a brief overview of the system administrator's roles and responsibilities in the end-to-end CHARM operation:

1. The system administrator participates in the MES planning with the account team and the SSR.
2. The system administrator reviews the planning guidelines, checklist and the estimated time for the CHARM operation.
3. For a hot node upgrade, the system administrator confirms the location(s) of the physical node(s) with the SSR.
4. The system administrator obtains the FRU and node location code(s) from the SSR for the PHRU utility.
5. The system administrator coordinates the CHARM activity with the SSR to optimize the service window.
6. For a hot node upgrade, hot node repair, or I/O hub repair, the system administrator uses the PHRU utility and OS tools to prepare the system before the SSR performs the CHARM procedure.
7. After the SSR completes the CHARM procedure, the system administrator restores the processor, memory, I/O and partitions to the original or desired configuration.

2.4 Service Representative Roles and Responsibilities

This section provides a brief overview of the SSR's roles and responsibilities in the end-to-end CHARM operation:

1. The SSR participates in the MES planning with the account team and the system administrator. The SSR reviews the WCII (World-Wide Customized Installation Instructions) for a hot add or upgrade operation. The SSR also verifies that all of the required parts are available.
2. The SSR reviews the planning guidelines, checklist, and the estimated time for the CHARM operation.
3. For a hot node upgrade, the SSR works with the system administrator to identify the physical node(s) and the locations that are available for the upgrade.
4. For a hot node or I/O hub repair, the SSR communicates to the system administrator the location code(s) of the FRU to be repaired, and the location code of the node that houses the target FRUs.
5. For a hot node upgrade, hot node repair, or I/O hub repair, the SSR communicates to the system administrator the need to use the PHRU utility and OS tools to prepare the system before the CHARM procedure begins.
6. The SSR coordinates the CHARM activity with the system administrator to optimize the service window.
7. The SSR performs the CHARM procedure using the HMC's Repair and Verify tool, and the WCII instructions for an MES.
8. After the completion of the CHARM procedure, the SSR informs the system administrator that the processor, memory, I/O and partitions need to be restored to the original or desired configurations.

3 Best Practices

3.1 Minimum Enablement Criteria

CHARM operations are very complex, they entail numerous steps that are performed by service personnel and/or the system administrator while the system is powered on. Since the likelihood of failure increases with the complexity of the operation, the following minimum enablement criteria offer the best protection against any unforeseeable situations:

1. It is strongly recommend that all scheduled hot adds, upgrades or repairs are done during "non-peak" operational hours.
2. It is a prerequisite that critical I/O resources are configured with redundant paths for hot node repair, hot node upgrade, and hot GX adapter repair.
3. It is a prerequisite that ESA (Electronic Service Agent) is enabled to ensure timely resolution of all hardware failures. This will minimize the opportunity for a situation in which there are multiple hardware failures during a CHARM operation.
4. It is a prerequisite that critical business applications are moved to another server using Live Partition Mobility (LPM), if available, OR critical applications are quiesced for hot node add, hot node repair, hot node upgrade and hot GX adapter repair.
5. IBM recommends that the system administrator does not dynamically change the size of the 16M large page pool in AIX partitions with the vmo command while a CHARM operation is in progress. Please refer to http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=/com.ibm.aix.cmds/doc/aixcmds6/vmo.htm for more information on vmo command.

IBM remains committed to addressing our customers' requirements. We recognize that quiescing critical workloads or using LPM to migrate applications presents a significant restriction for our customers. As such, IBM plans to closely monitor performance of the CHARM function through ESA. Based on positive field performance, we expect to relax the prerequisite to quiesce critical applications. IBM anticipates the relaxation to be rolled out on a function-by-function basis, based on the availability of statistically significant results. Given the complexity of CHARM operations, other recommendations and prerequisites will continue to apply.

The table below summarizes the minimum enablement criteria for each CHARM function, as well as other concurrent maintenance functions.

IBM Power Systems

IBM

CCM / CHARM Minimum Enablement Criteria

Functions	Criteria			
	Off Peak[1]	Redundant I/O[2]	ESA enabled[3]	LPM or Quiesce[4]
Fan/Blower/Control Add or Repair	Recommend			
Power Supply/Bulk Power Add or Repair	Recommend			
Op Panel	Recommend			
DASD / Media Drive & Drawer Add	Recommend			
DASD / Media Drive & Drawer Repair	Recommend	Prerequisite		
PCI Adapter Add	Recommend			
PCI Adapter Repair	Recommend	Prerequisite		
I/O Drawer Add	Recommend			
I/O Drawer Repair, Remove	Recommend	Prerequisite		
System Controller Repair	Recommend			
GX Adapter Add	Recommend		Prerequisite	
GX Adapter Repair	Recommend	Prerequisite	Prerequisite	Prerequisite
Node Add	Recommend		Prerequisite	Prerequisite
Node Upgrade (memory)[5]	Recommend	Prerequisite	Prerequisite	Prerequisite
Hot Node Repair[5]	Recommend	Prerequisite	Prerequisite	Prerequisite

1. Highly recommend that schedule upgrades or repairs are done during "non-peak" operational hours.
2. Prerequisite that critical I/O resources are configured with redundant paths.
3. ESA (Electronic Service Agent) enablement highly recommended for Power6 systems and prerequisite for Power7 systems
4. Prerequisite that business applications are moved to another server using Live Partition Mobility, if available, OR critical applications quiesced
5. IBM recommends that you not dynamically change the size of the 16M large page pool in AIX partitions with the vmo command while a CCM/CHARM operation is in progress

7 Power your planet.

Table 2: Minimum Enablement Criteria

3.2 System Hardware Configurations

To obtain the maximum system and partition availability benefits from the CEC Hot Add and Repair Maintenance functions, follow these best practices and guidelines for system hardware configuration:

1. Request the free pre-sales "I/O Optimization for RAS" services offering to ensure that your system configuration is optimized to minimize disruptions when using CHARM.

2. The system should have spare processor and memory capacity to allow a node to be taken offline for hot repair or upgrade with minimum impact. Unlicensed Capacity on Demand processors and memory will be used by the system firmware (automatically) during node evacuation. Alternatively, LPM can be used to move partition(s) to other system(s), active partitions can be powered off, and memory and processor capacity can be dynamically removed from active partitions to free up sufficient processor and memory capacity to take the node offline.

3. All critical I/O resources should be configured using an operating system multi-path I/O redundancy configuration. Please refer to the specific operating system user guides for additional information.

4. Redundant I/O paths should be configured through different nodes and GX adapters. Properly configured redundant I/O paths ensure continued availability of I/O resources in the event of a critical node or GX adapter failure, and also during a hot repair or upgrade operation. For multi-node systems, the redundant paths should be attached to different nodes. For single-node systems, the redundant paths

should be connected to different GX adapters. If possible.

5. All logical partitions should have RMC network connections with the HMC(s)/SDMC(s).

6. The HMC/SDMC should be configured with redundant service networks with both service processors in the CEC.

3.3 Partition Configuration Considerations

Maintaining the availability of partitions during a hot node repair or upgrade requires advanced planning to ensure that the impacts to the resources are minimized. It is important to realize that the operation is not completely transparent to the logical partitions on the system. Since an action of this type temporarily reduces the memory, processing, and I/O capacity of the system, it will very likely impact individual partitions if proper planning is not done.

Memory and processors are the easiest types of resources to plan for since PowerVM™ can transparently change processor and memory assignments to accommodate the removal of a node. The only requirement that the system administrator must plan for is that there must be enough free memory and processor capacity available in the system to accommodate the entitled capacity of active partitions and shared resource pools with one node removed Or, it must be possible to make such capacity available through actions such as memory and processor DLPAR, LPM, or powering off partitions.

The following considerations should be taken into account when planning to free sufficient processor and memory capacity:

1. Capacity Upgrade on Demand (CUoD) resources that have not been activated will be used by the system firmware to meet the free memory and processor requirements during the CHARM operation (with no CUoD usage charge).
2. Processor and memory resources currently assigned to logical partitions that are powered off will be considered available resources to be used during the hot node repair or upgrade procedure .
3. Partitions can either be powered off, or have processor and/or memory resources removed from them using dynamic logical partitioning (DLPAR) operations, to free the resources needed to perform the hot node repair or upgrade. If DLPAR is used, the minimum processor and memory settings for each logical partition should be planned for appropriately so that an entire node's worth of processor and memory resources could be removed from the partitions that will be running during the CHARM operation.

An additional logical partition configuration consideration is related to the system setting to "Power off the system after all the logical partitions are powered off." This setting can be displayed and modified by selecting the "Properties" for your server on the HMC and viewing the "General" tab (figure 2). Preparation for a hot node repair or upgrade may impact resources on the system and may require logical partitions to be powered off. If the last logical partition is powered off as a part of preparing for the hot node repair or upgrade, the system will also power off if this option is selected. In order for the system to successfully complete a hot repair or upgrade operation, this setting should not be selected if logical partitions are being deactivated. If this option is selected and the last logical partition is powered off, the system will attempt to power off in the middle of the hot repair or upgrade operation.

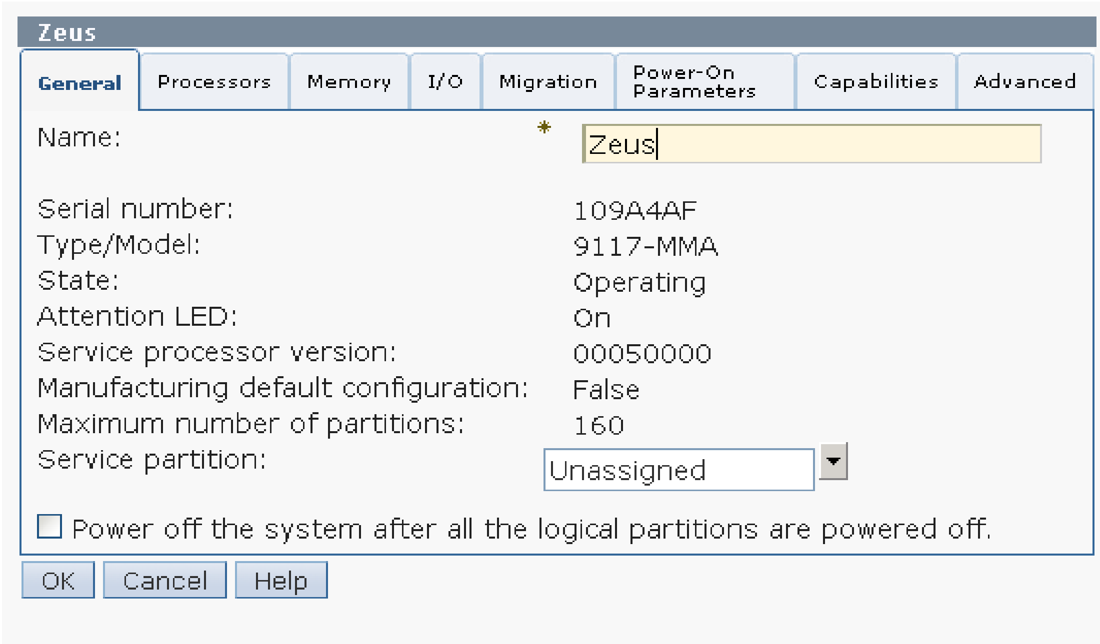

Figure 2: Power off the system after all the logical partitions are powered off

3.4 I/O Availability Considerations

I/O resource planning is more complex. During a hot node repair or upgrade procedure, the integrated I/O resources in the CEC drawer, and in I/O expansion units attached to the GX adapters in that node, will be unavailable. If a partition needs to continue running during the procedure and is using I/O resources either internal to the CEC drawer or externally attached to the target node, careful planning must be done to ensure that access to critical I/O resources will be maintained.

The CHARM functions utilize the existing “administrative network” between the HMC and the logical partitions (LPARs) to prepare the system for the procedure. Therefore, the availability of the “administrative network” is key to a successful CHARM operation.

To maintain access to the I/O internal to the CEC drawer or externally attached through the target node, multi-path I/O solutions must be utilized so that there are multiple I/O adapter (IOA) paths to the physical I/O resource such as network or SAN-attached storage. Care must be taken to ensure that the redundant IOAs for each resource are located in different I/O expansion units that are attached to different GX adapters located in different nodes.

These redundant paths to the I/O resources could be assigned to the same partition and use the multi-path physical I/O solutions that have been implemented by each OS. Another example would be having each of the redundant paths assigned to a different Virtual I/O Server (VIOS) partition, with the VIOS partitions being redundant VIOS servers for VIOS clients that need to remain available during the CHARM procedure..

In the example shown in Figure 3, four I/O drawers are attached to four GX adapters that have been placed in different nodes. There is an Ethernet adapter and Fiber Channel adapter in each I/O drawer to provide redundant access to the Ethernet and storage networks. As a result, if access to one drawer is lost, it would not disrupt access to the Ethernet and storage networks for partitions that have redundant paths to the I/O configured through different nodes. Figure 4 shows the connections that would be disrupted during the hot repair of node 1. A partition with paths to the I/O configured through node 1 and at least one other node would continue to have access to the Ethernet and storage networks.

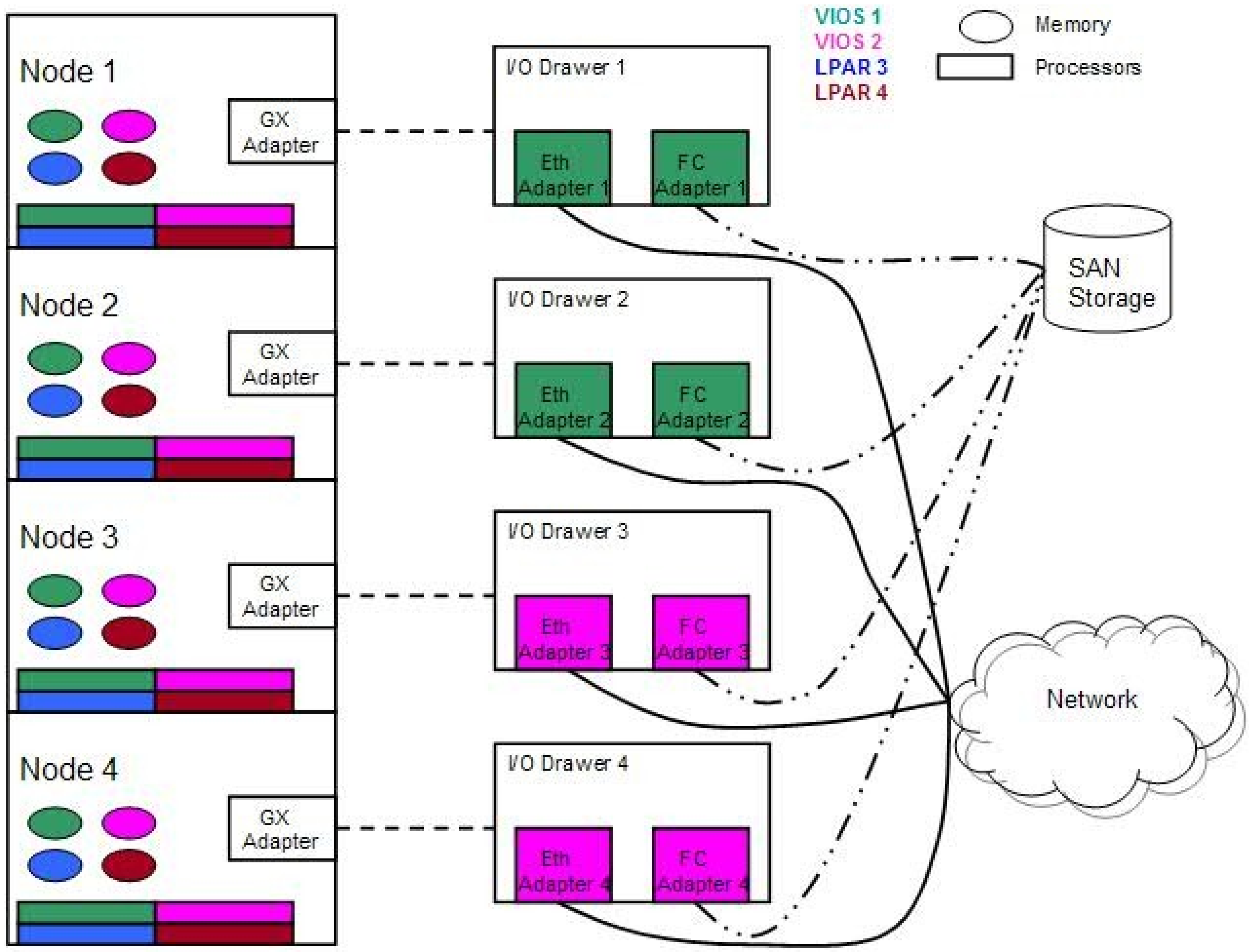

Figure 3: System configuration with redundant VIOS and I/O paths

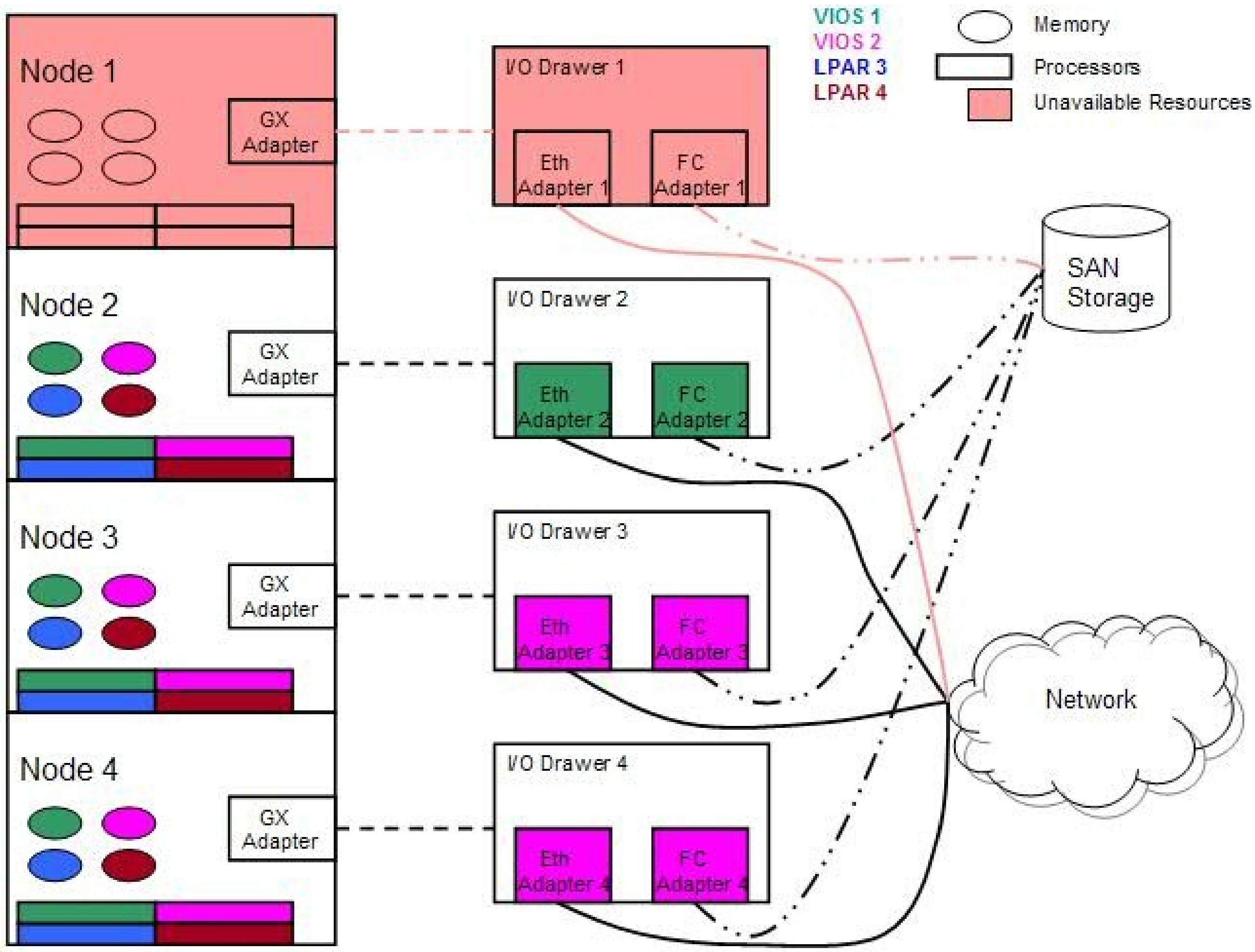

Figure 4: System configuration with redundant VIOS and I/O paths, during hot repair of node 1

3.5 Redundant I/O Options

There is a variety of methods and technologies described in this section that can be implemented in order to ensure continuous availability of I/O resources to partitions during a CHARM operation. The use of properly configured dual, or redundant, Virtual I/O Server (VIOS) partitions providing virtualized I/O to clien partitions running customer workloads is perhaps the most convenient method for achieving this. With such a configuration, the physical I/O resources can be assigned to the VIOS partitions in a manner that ensures that the loss of one node for CHARM affects at most one of the VIOS partitions, allowing the other VIOS partition to continue to provide virtualized I/O resources to the client partitions.

Factors such as desired availability and performance levels, I/O throughput and capacity requirements, and system manageability need to be considered when planning the VIOS configuration. A small system may not have enough resources to configure a second VIOS and meet the required performance or throughput requirements.

The sections below provide information about strategies for creating a redundant VIOS configuration that allows a single node to be taken down for maintenance without disrupting access to I/O resources by client partitions. To achieve this, physical I/O resources must be assigned to the VIOS partitions so that a single node can be shut down for maintenance without requiring both VIOS partitions to be shut down, as described previously. There are Redbooks with additional information about the virtualization technologies described in these sections. A recommended one is "PowerVM Virtualization on IBM System p: Introduction and Configuration."

3.5.1 Disk/Storage Redundancy

Disk redundancy in a redundant VIOS configuration can be achieved using:

- Multi-path I/O (MPIO)
- Logical Volume Manager (LVM) mirroring

Note that that these offerings are available for both AIX and IBM i; however, terminologies may vary between the two operating systems.

3.5.1.1 Multi-path I/O

Redundant I/O adapter (IOA) paths to the physical I/O resources can be assigned to the same partition and use the MPIO solutions that are implemented by the operating system. MPIO can be implemented at both the VIOS level and the client partition level as shown in Figure 5. Building on the example shown in Figure 3, each VIOS owns two Fiber Channel (FC) adapters placed in different I/O drawers, which are in turn attached to different nodes. In each VIOS, MPIO is used to create redundant paths to the physical storage network. As a result, if one of the physical FC adapters is unavailable, the VIOS can use the second FC adapter to get access to the storage network as shown in Figure 6.

With MPIO in the client partition, the client partition has a virtual SCSI client adapter connected to each VIOS. The corresponding virtual SCSI server adapters in the VIOS partitions are backed by the same physical resources (LUNs in the storage network). As a result, if a VIOS needs to be powered down or both the physical FC adapters in a VIOS partition are unavailable, the client partition can use MPIO to access the storage network through the second VIOS as shown in Figure 7.

In order to balance the load of multiple client partitions across two VIOS partitions, the priority on each virtual SCSI disk in the client partition can be set to select the primary path, and therefore, a specific VIOS.

MPIO for virtual SCSI devices only supports fail-over mode. Only one path in a partition is used at a given time even if both paths are enabled. If the primary path fails, the partition will fail-over to the secondary path.

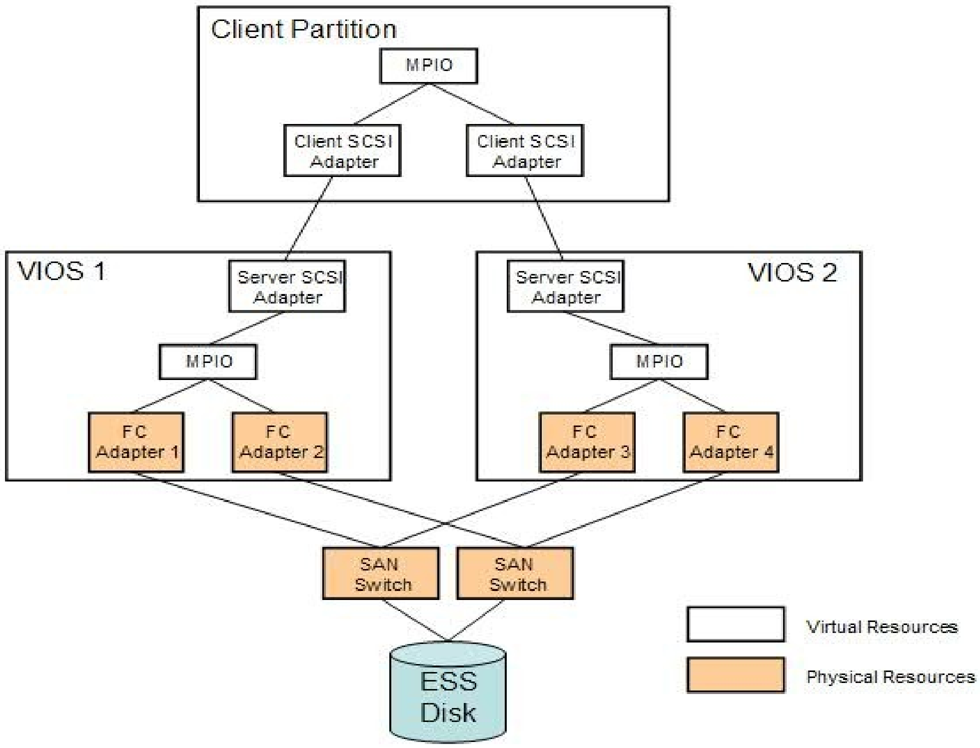

Figure 5:

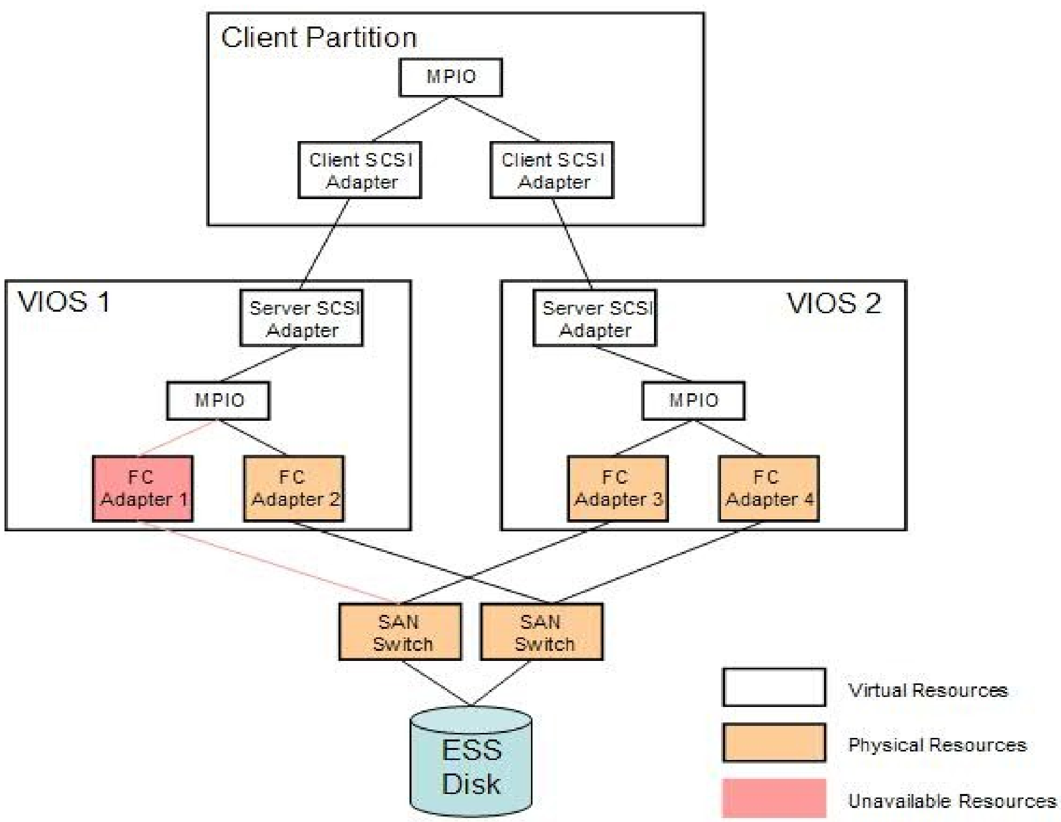

Figure 6:

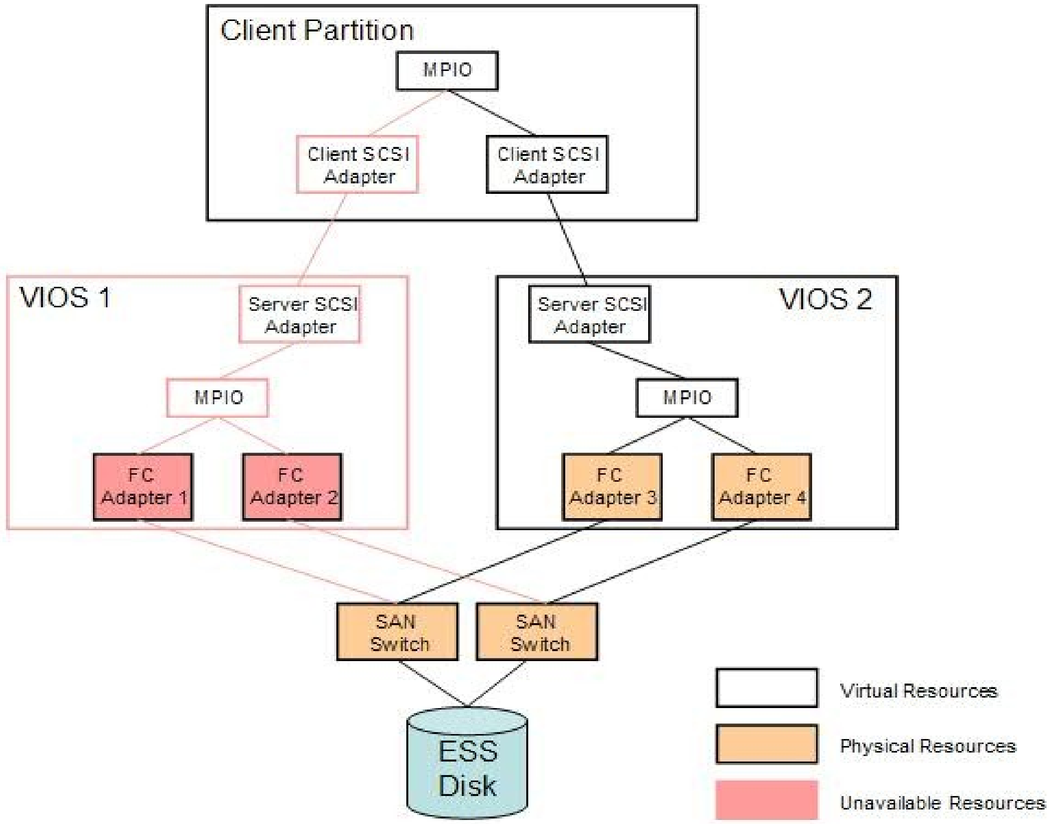

Figure 7:

3.5.1.2 Logical Volume Manager Mirroring

With the use of LVM mirroring in the client partition, each VIOS can present a virtual SCSI device with a backing device that is different than that for the virtual device presented by the other VIOS. The corresponding virtual disks in the client partition can then be in a normal mirrored volume group as shown in Figure 8. If one of the VIOS partitions becomes unavailable, the mirroring will allow the client partition to continue running with the virtual disk from the other VIOS as shown in Figure 9.

Note that while volume groups and volume group mirroring are AIX terms, the approach described in this section can also be applied in IBM i using Auxiliary Storage Pools (ASP) and disk mirroring.

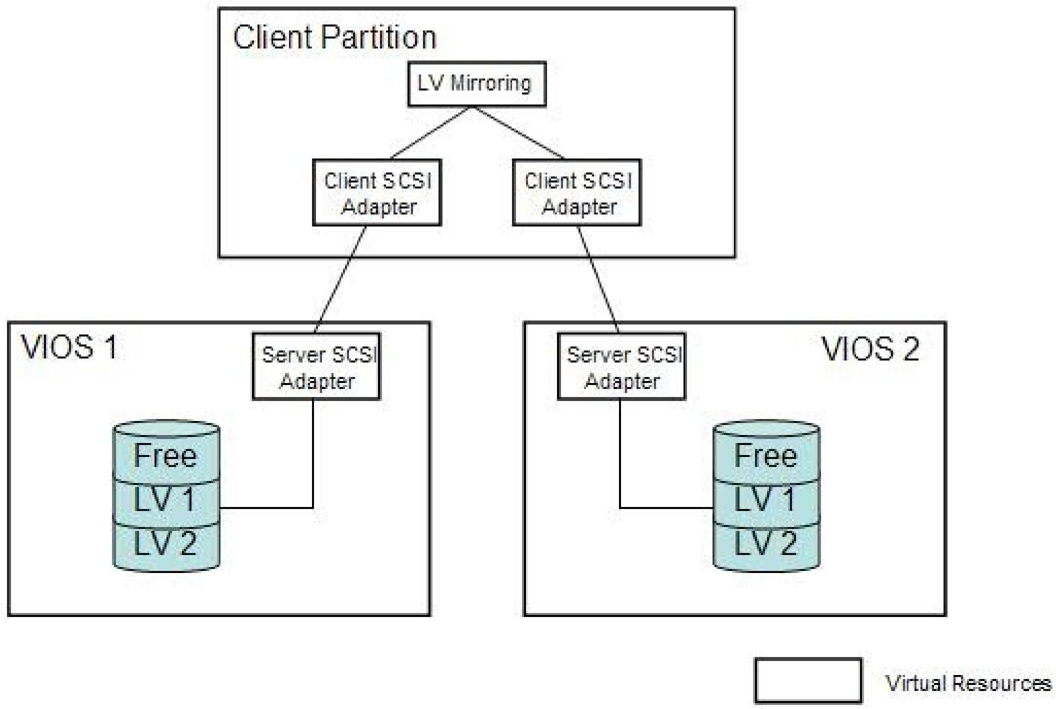

Figure 8:

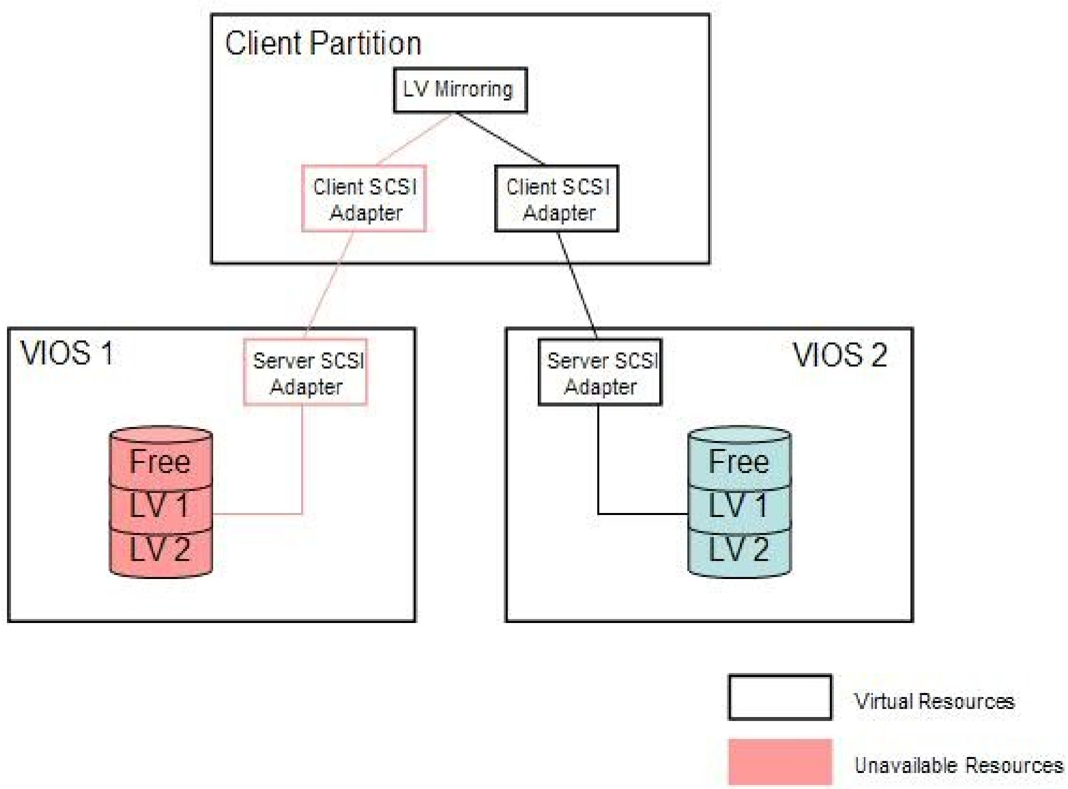

Figure 9:

3.5.2 Ethernet Redundancy

The most common approaches for continuous availability of Ethernet in a redundant VIOS configuration are:

- Shared Ethernet Adapter (SEA) fail-over
- Network Interface Backup (NIB)

Another approach that is intended primarily to increase network bandwidth but provides some degree of increased availability is link aggregation.

3.5.2.1 Shared Ethernet Adapter (SEA) fail-over

Shared Ethernet adapter fail-over offers Ethernet redundancy to the client at the VIOS level. In an SEA fail-over configuration two VIOS partitions have the bridging functionality of the shared Ethernet adapters, and a control channel is used to determine which of the VIOS partitions is supplying the Ethernet service to the client. The control channel serves as a channel for the exchange of keep-alive or heartbeat messages between the two VIOS partitions and therefore controls the fail-over of the bridging functionality. No network interfaces have to be attached to the control channel Ethernet adapter and it should be placed on a dedicated VLAN that is not used for anything else. Note that the keep-alive messages are only sent over the control channel and so an SEA fail-over would not occur for an external network failure. The SEA fail-over feature can be configured to periodically check the reach-ability of a given IP address. Each SEA has a priority value which determines which will be the primary (active) and which will be the backup (standby). The bandwidth of the standby physical Ethernet adapter is not used. If one VIOS is shut down for maintenance an automatic fail-over will occur. Similarly, if the SEA is configured to periodically check the accessibility of a given external IP address, loss of access to the external network will also cause an automatic fail-over. A manual fail-over can also be triggered.

The client partition has no special protocol or software configured and uses the virtual Ethernet adapter as though it was bridged by only one VIOS. This helps simplify the configuration of clients and virtual network administration. It allows the approach to be used for all operating systems that support virtual ethernet. NIM installations are also simplified as the Ethernet configuration does not need to be modified during installs.

If a system is set up as shown in Figure 3, SEA fail-over can be configured using two VIOS partitions configured for SEA fail-over, where the physical adapters that the SEAs bridge to are connected through different nodes. The logical partition configuration view of such a setup is shown in Figure 10. If the primary SEA becomes unavailable, the VIOS that owns the backup SEA will detect it via the control channel and initiate an SEA fail-over as shown in Figure 11. As a result, the other VIOS will now provide access to the external network using the SEA that it owns.

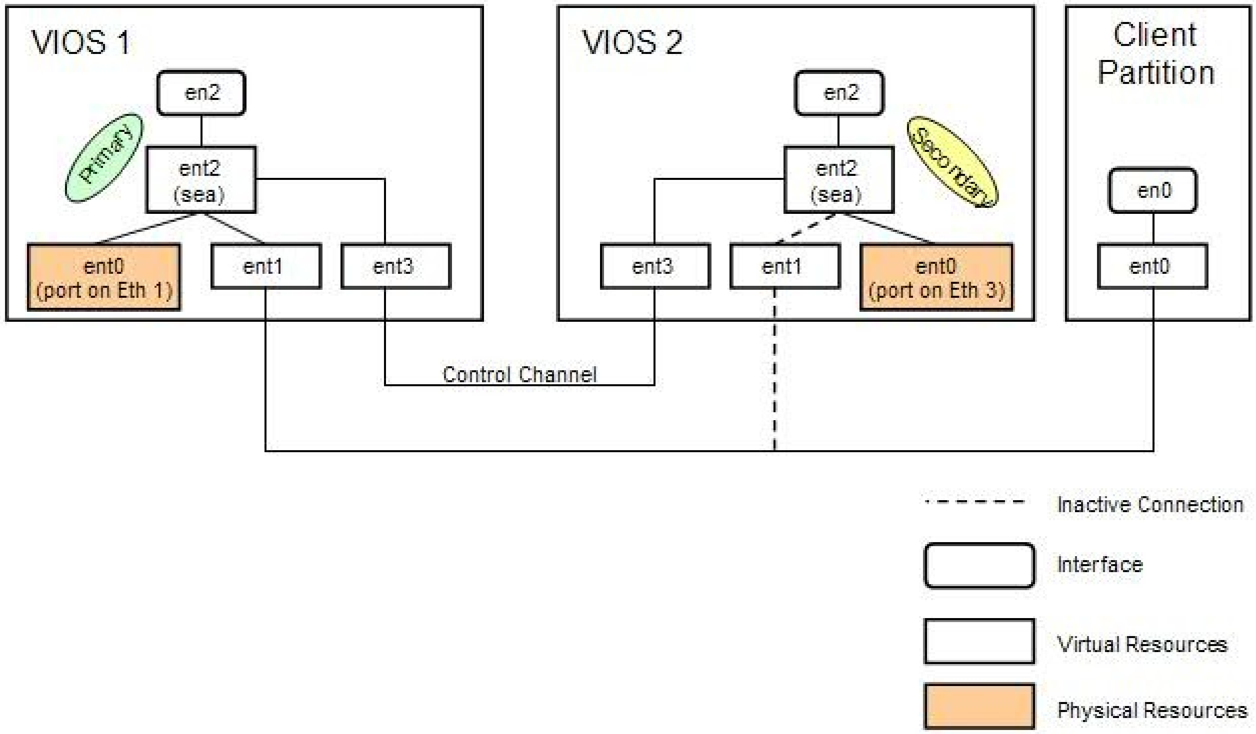

Figure 10:

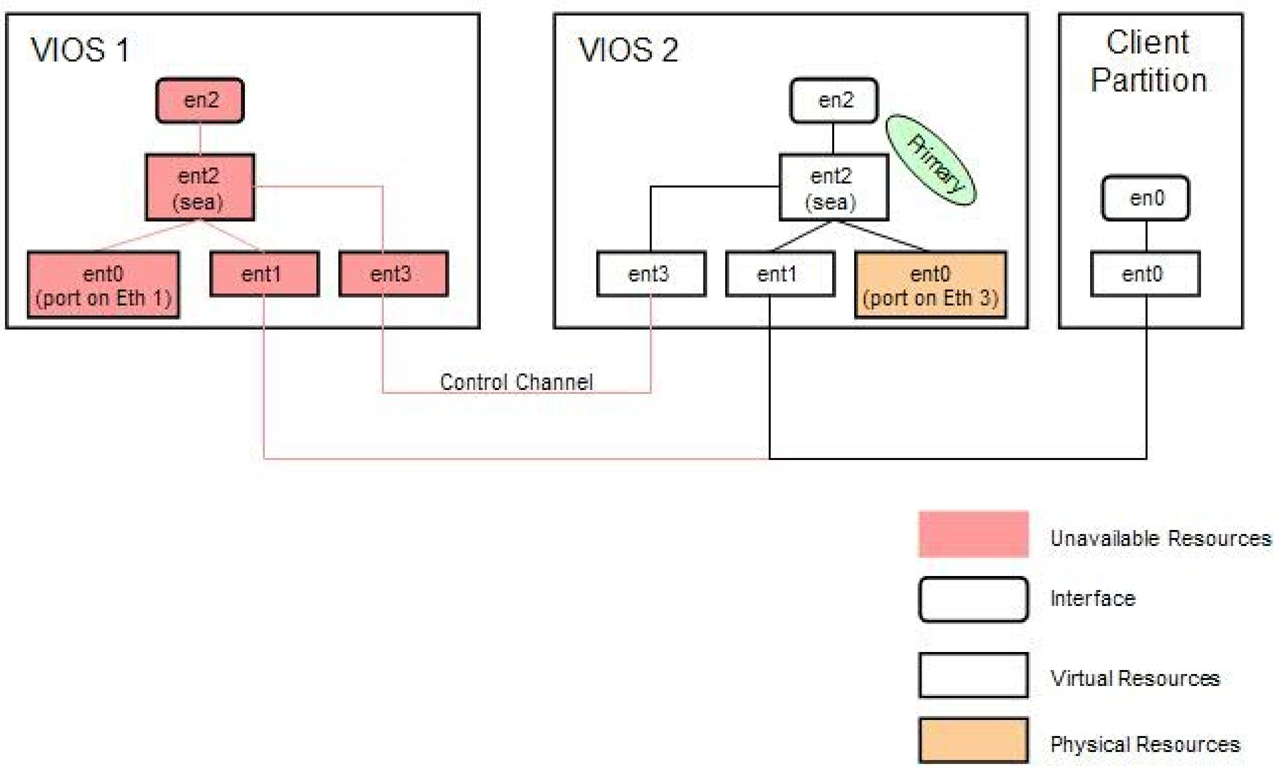

Figure 11:

3.5.2.2 Network Interface Backup (NIB)

Network interface backup can be used at the client partition level to achieve network redundancy when using two VIOS partitions. The client partition uses two virtual Ethernet adapters to create an Ethernet channel (EtherChannel) that consists of a primary adapter and a backup adapter. If the primary adapter becomes unavailable, the NIB switches to the backup adapter.

When configuring NIB in a client partition, each virtual Ethernet adapter has to be configured on a different VLAN. The two different internal VLANs are then bridged to the same external VLAN.

If a system is setup as show in Figure 3, NIB can be configured using two VIOS partitions each of which owns a physical Ethernet adapter where the two adapters are connected through different nodes. NIB is implemented in the client partition with the primary channel being through a port on the Ethernet adapter owned by one VIOS, and the secondary channel being through a port on the Ethernet adapter owned by the other VIOS, as shown in Figure 12. If the Ethernet adapter being used by the primary channel becomes unavailable, NIB will switch to using the secondary channel through the Ethernet adapter owned by the other VIOS, as shown in Figure 13.

NIB enables the bandwidth of the physical Ethernet adapters in both the VIOS partitions to be used concurrently by different client partitions. Half of the clients can use the SEA assigned to one VIOS as the primary adapter and the other half can use the SEA assigned to the other VIOS as the primary adapter. As a result, NIB is ideal when one wants to load balance client partitions to use both Virtual I/O Servers.

Note that NIB is not supported by Linux partitions.

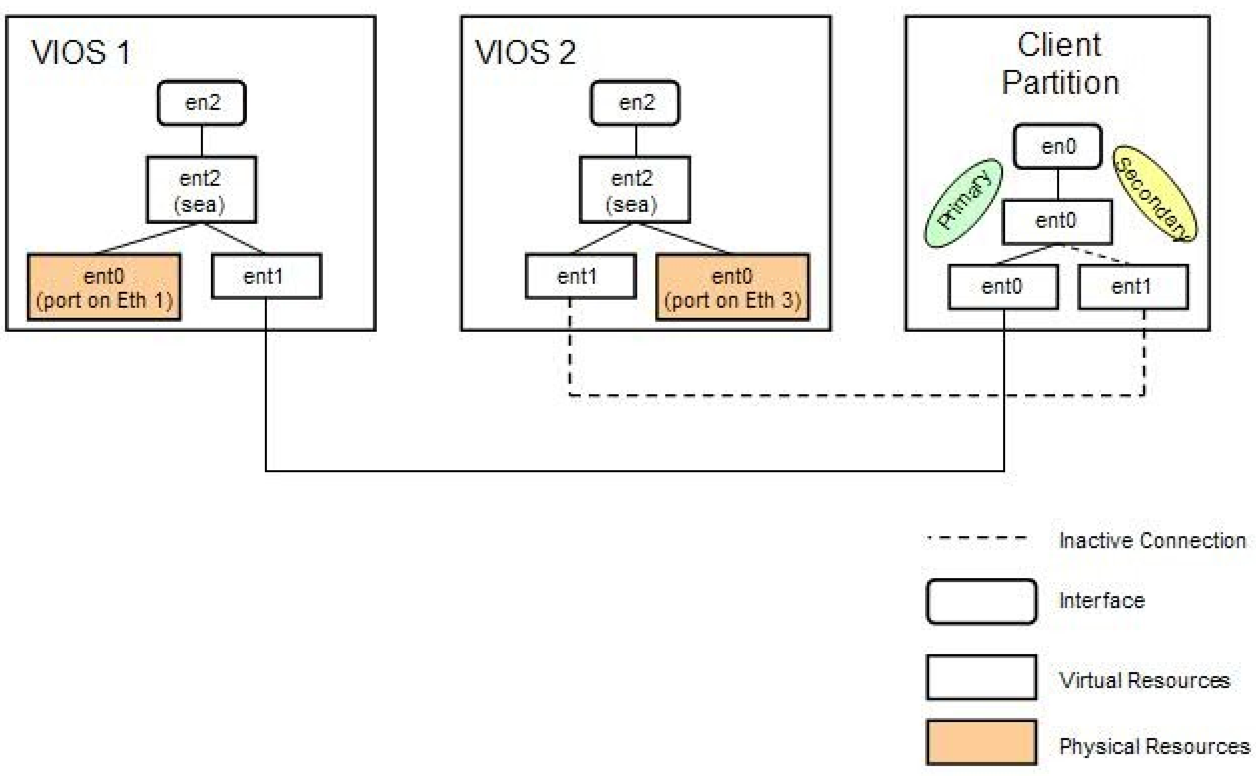

Figure 12:

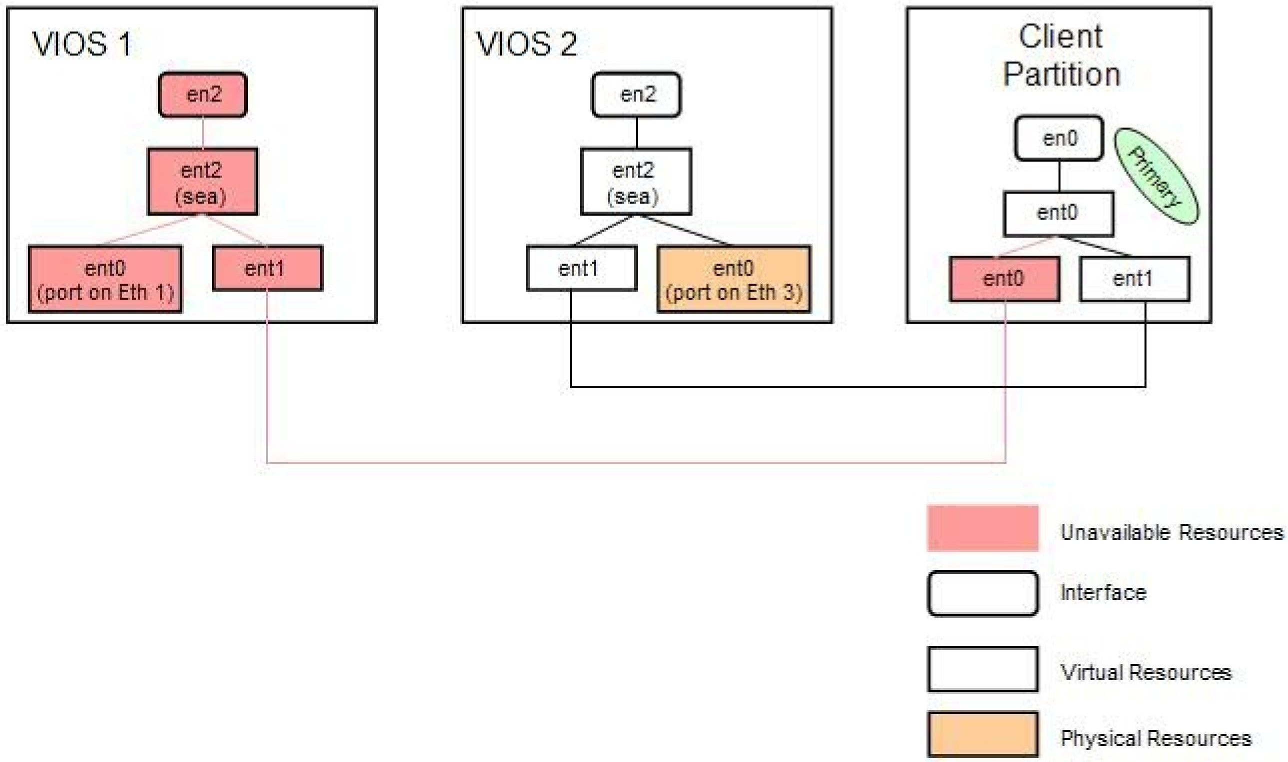

Figure 13:

3.5.2.3 Link Aggregation

Link aggregation is a network port aggregation technology that allows several Ethernet adapters to be aggregated together to form a single pseudo-Ethernet adapter. If an adapter fails or is unavailable, the packets are automatically sent on the next available adapter without disruption to existing user connections. A link or adapter failure will therefore lead to performance degradation, but not a disruption. The adapter is automatically returned to service on the link aggregation when it recovers.

An added advantage of link aggregation is the increase in bandwidth compared to using a single network adapter. As a result, any bottlenecks that can be caused by one network adapter being shared among many client partitions can be avoided.

All the aggregated links must connect to the same switch. A backup adapter connected to another Ethernet switch with the same VLAN can be used to provide a single additional backup link.

If a system is set up as shown in Figure 3, link aggregation can be configured in a VIOS using Ethernet adapters connected through different nodes. The VIOS owns multiple Ethernet adapters, two of which are aggregated to a single virtual Ethernet device, while another physical adapter provides the backup link as shown in Figure 14. If one of the Ethernet adapters being used for link aggregation is unavailable, all the packets will automatically be sent through the other adapter without causing any disruption to the client partitions, as shown in Figure 15. If for some reason both the Ethernet adapters used for link aggregation fail, the backup link being provided by the third Ethernet adapter will be used to provide the client Ethernet services, as shown in Figure 16. Note that link aggregation would provide no benefits if the VIOS needs to be taken down.

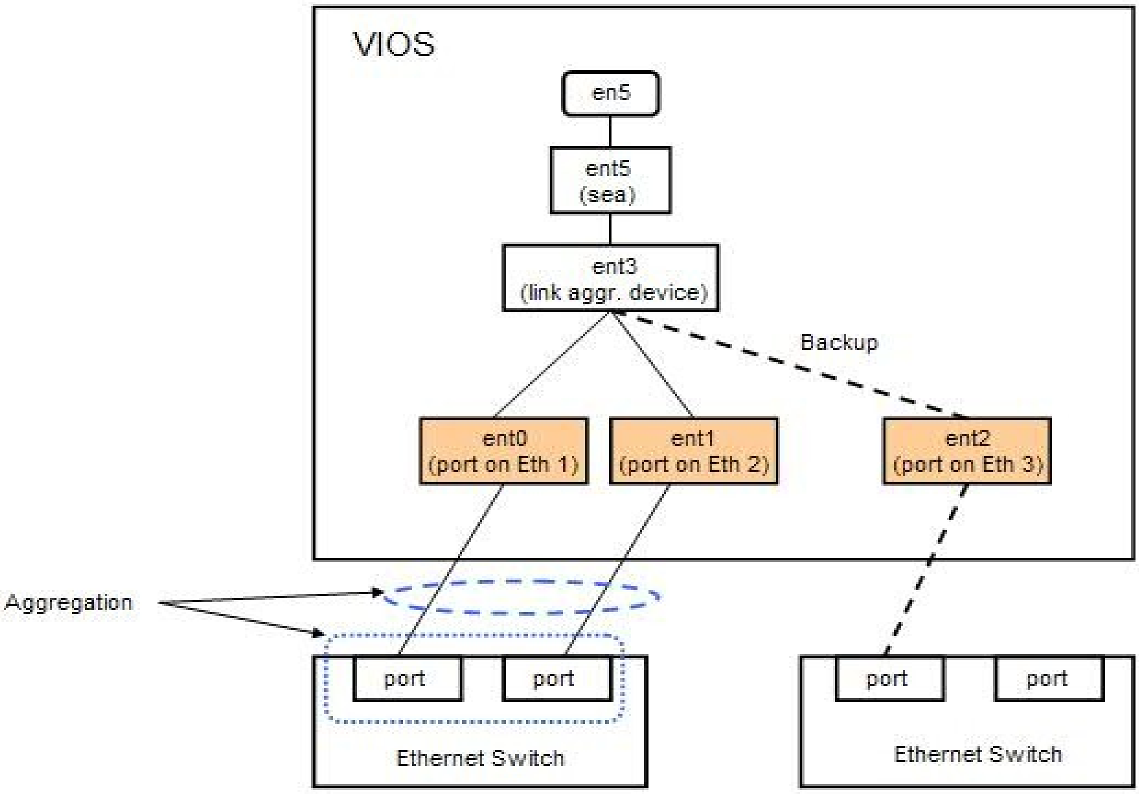

Figure 14:

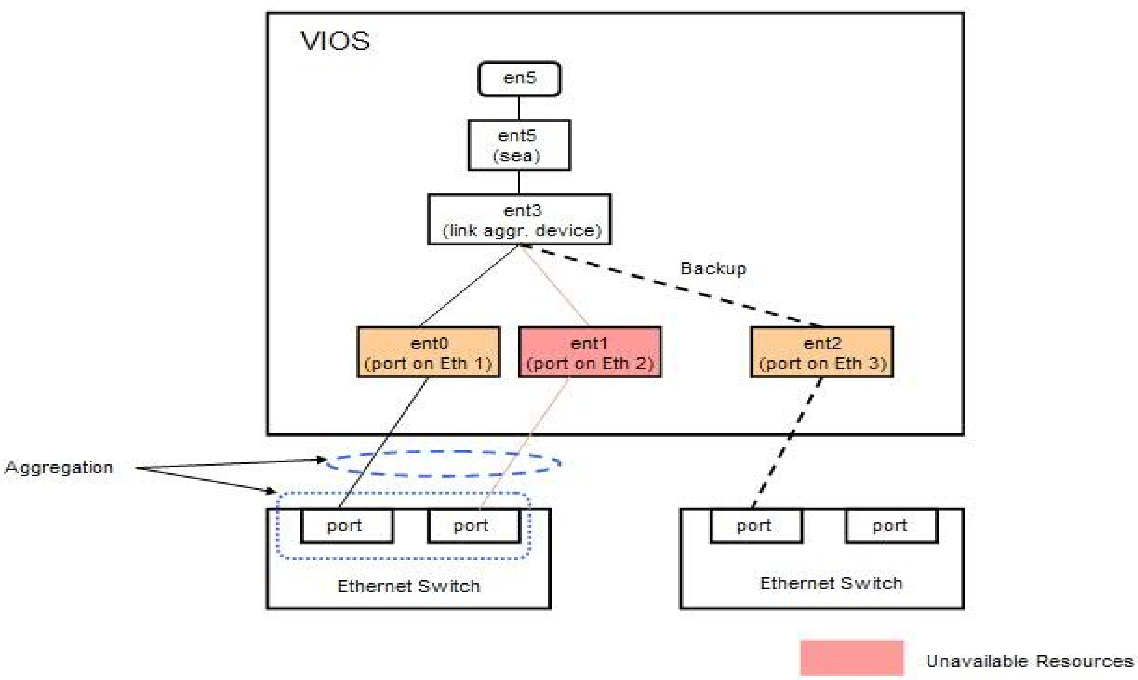

Figure 15:

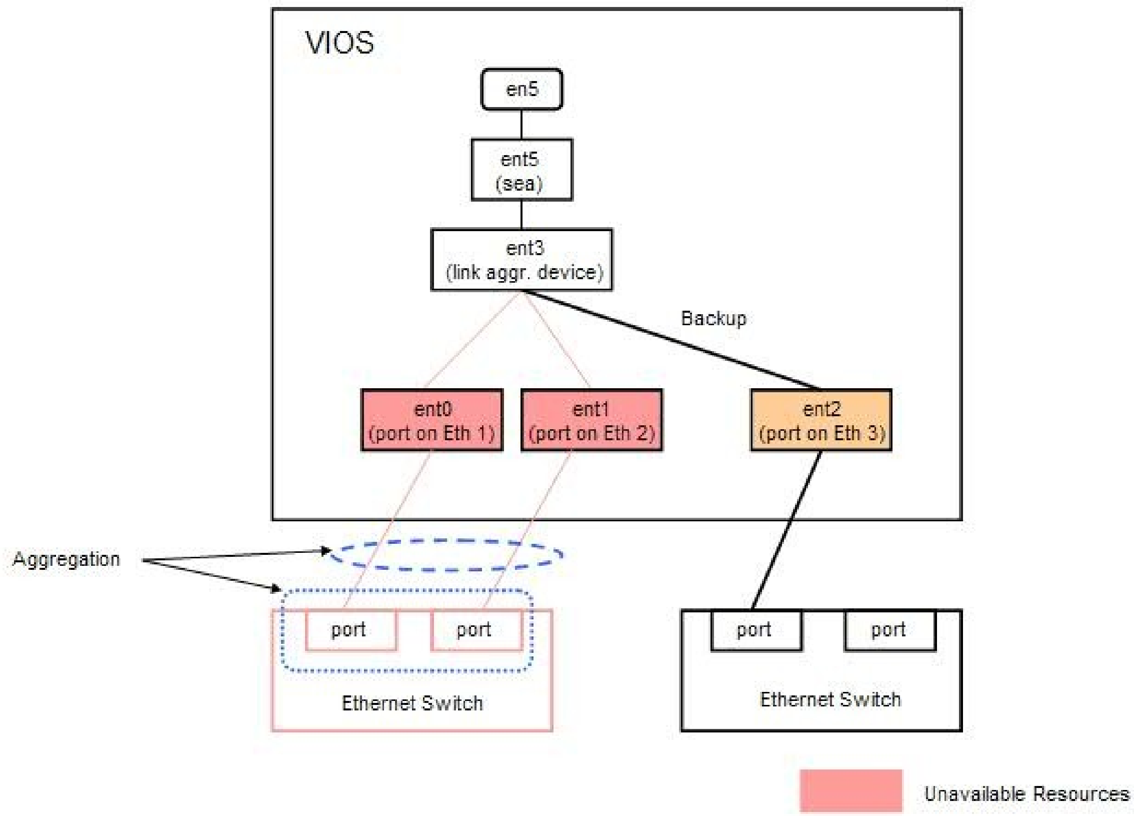

Figure 16:

3.5.3 Assessing VIOS Client Partition Impacts

If you have a properly configured redundant VIOS configuration as described in the previous sections, you should be able to do a hot node upgrade or repair, hot GX adapter repair, or concurrent repair on an I/O expansion unit or PCI adapter without causing disruption to client partitions in the form of loss of connectivity to I/O resources. If physical I/O resources are assigned to the VIOS partitions such that each redundant VIOS partition's resources are connected through a different node than those of the other redundant VIOS partitions, at most one VIOS partition may have to be powered off if a node is shut down for maintenance. Employing the multi-path strategies described previously in the client partitions ensures connectivity to I/O devices is not lost if one of the VIOS partitions must be shut down. By employing multi-path at the VIOS partition level in addition to client partition level, and configuring the paths to a given I/O device through different nodes, it is possible to shut down a node without having to power off any of the redundant VIOS partitions.

If you do not have a redundant VIOS configuration, or the configuration does not conform to the best practice strategies described in previous sections, a CHARM procedure that requires physical I/O resources assigned to the VIOS partitions to be taken offline may impact client partitions. This section describes how to assess those impacts.

3.5.3.1 Client Partitions Impacted by Virtual SCSI Configuration

3.5.3.1.1 Determining Impacts Using VIOS Command Line

Run the following command from the VIOS command line:

lsmap –all

An example of an entry from the output is shown in Figure 17.

```
$ lsmap -all
SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost0          U9406.675.107074C-V17-C13                    0x0000000f

VTD                   vtopt0
Status                Available
LUN                   0x8200000000000000
Backing device        /var/vio/VMLibrary/WindowsServer2003.iso
Physloc

VTD                   vtscsi0
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk13
Physloc               U5796.001.10536A3-P1-C5-T1-L2-L0
```

Figure 17:

The output includes the hexadecimal client partition ID as well as the backing devices associated with each server virtual SCSI adapter (SVSA), also referred to as a virtual SCSI(vSCSI) server adapter. For a given vSCSI adapter the client partition ID will be 0x00000000 if no virtual target device (VTD) resources have been configured for it, or the client partition it is configured to connect to is powered off.

3.5.3.1.2 Determining Impacts Using HMC

On the HMC the "Hardware Information -> Virtual I/O Adapters -> SCSI" utility (Figure 18) can be run for a given VIOS to determine which client partition profiles are configured to use the VIOS VSCSI services.

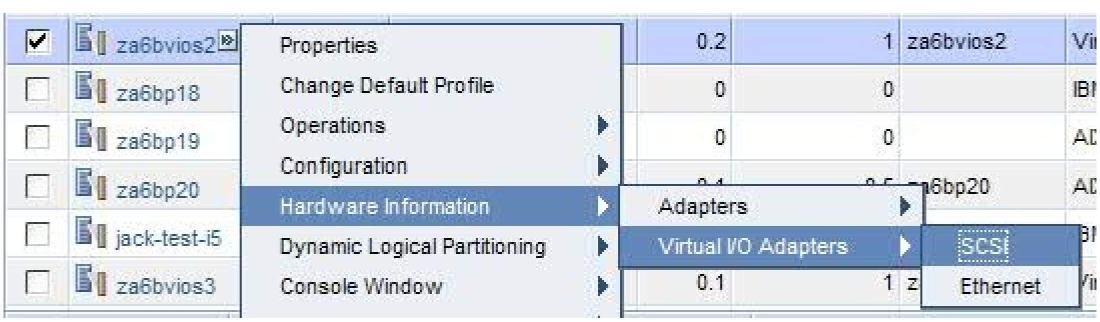

Figure 18:

Clicking on "SCSI" in Figure 18 will show the screen in Figure 19.

hmc16: SCSI - Mozilla Firefox

https://

Virtual SCSI Topology: za6bvios2@za6be-9406-675-SN107074C - za6bvios2

Current virtual SCSI adapter configuration

Virtual Adapter	Backing Device	Remote Partition	Remote Adapter	Remote Backing Device
vhost4		za6bp6(6)	6	
vhost5		za6bp13(13)	2	
vhost0	/var/vio/VMLibrary/WindowsServer2003.iso,hdisk13	za6bp15(15)	5	DD006, OPT01
vhost6		za6bp4(4)	6	
vhost1	/var/vio/VMLibrary/slic611190007.iso	za6bp8(8)	4	
vhost2	hdisk9	za6bp6(6)	2	
vhost7	/var/vio/VMLibrary/xpf61010B292402.iso	za6bp10(10)	4	OPT01
vhost12	,	za6bp3(3)	4	OPT01, OPT02
vhost8	/var/vio/VMLibrary/slic710410002.iso	za6bp9(9)	4	
vhost9	/var/vio/VMLibrary/xpf61010B292402.iso	za6bp7(7)	4	

Close Help

Done

Figure 19:

As shown in Figure 19, the HMC displays all virtual SCSI (vSCSI) server adapters assigned to the VIOS partition. For each vSCSI server adapter, the local backing devices, partition ID and name of the client partition currently connected to the vSCSI server adapter (referred to as the 'Remote Partition' on the display), adapter/slot ID of the vSCSI client adapter in the client partition (referred to as 'Remote Adapter' on the display), and virtual devices exported to the client partition (referred to as 'Remote Backing Device' on the display) are shown. Note that the HMC will list virtual SCSI server adapters along with the remote partitions (client partitions) they are connected to even if the client partitions are inactive or no VTD's have been configured. Therefore, a client partition listed here may not actually be impacted if the physical I/O used by the VIOS to provide virtual SCSI services becomes unavailable. The use of "lsmap –all" at the VIOS command line (described in Section 3.5.3.1.1) may be used in conjunction with the HMC utility described here to gain a more accurate view of the actual client partition impacts.

3.5.3.2 *Client Partitions Impacted by SEA Configuration*

3.5.3.2.1 Information Available at VIOS Command Line

Run the following command from the VIOS partition command line:

lsmap –all –net

An example of an entry from the output is shown in Figure 20.

```
$ lsmap -all -net
SVEA   Physloc
------ --------------------------------------------
ent2   U9406.675.107074C-V16-C11-T1

SEA                    ent4
Backing device         ent0
Status                 Available
Physloc                U5796.001.10536A3-P1-C1-T1
```

Figure 20:

3.5.3.2.2 Determining Impacts Using HMC

In order to determine the list of client partitions using Ethernet services provided by a given VIOS, follow the steps outlined below on the HMC. The procedure involves collecting information from both the VIOS partition and all the potential client partitions on the system.

Run the “Hardware Information->Virtual I/O Adapters->Ethernet” utility (Figure 21) for the VIOS partition in question.

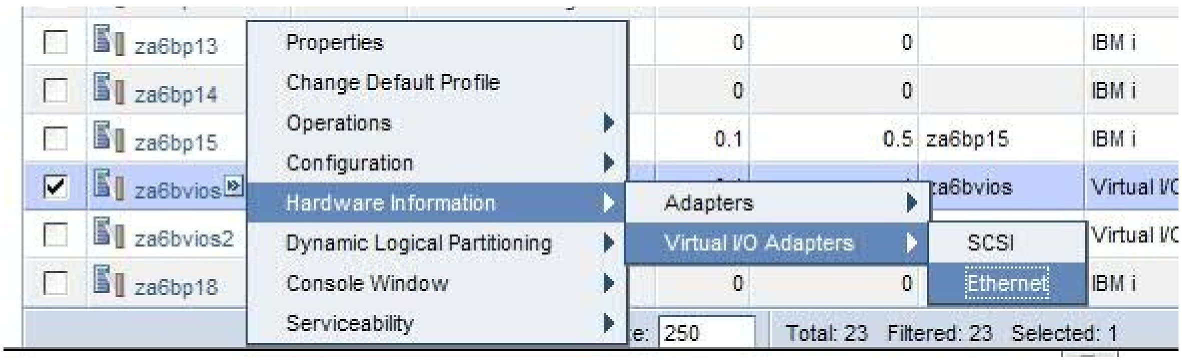

Figure 21:

The window in Figure 22 will be shown when “Ethernet” is clicked in Figure 21. It will list all the virtual Ethernet adapter resources, along with the virtual LAN (VLAN) IDs and physical shared adapters, associated with the virtual adapters. To map a physical adapter resource to its physical location code use the VIOS command described in Section 3.5.3.2.1.

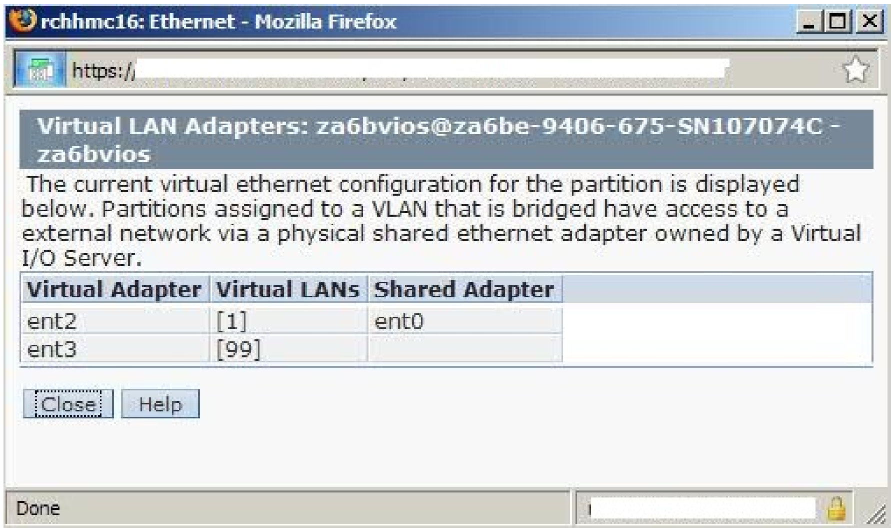

Figure 22:

In Figure 22, ent0 is the physical adapter port that connects the virtual network with VLAN ID 1 to the external network. The VIOS partition accesses VLAN ID 1 via the virtual adapter ent2. This is accomplished by configuring a Shared Ethernet Adapter (SEA) in the VIOS and specifying ent0 as the physical adapter and ent2 as a virtual adapter.

In order to determine which client partitions would lose access to the external network if the physical adapter providing port ent0 were to become unavailable to the VIOS partition, for each possible client partition go to the "Dynamic Logical Partitioning ->Virtual Adapters" utility, as illustrated in Figure 23. Note that the same information can also be obtained by using the "Properties->Virtual Adapters" utility for a given partition.

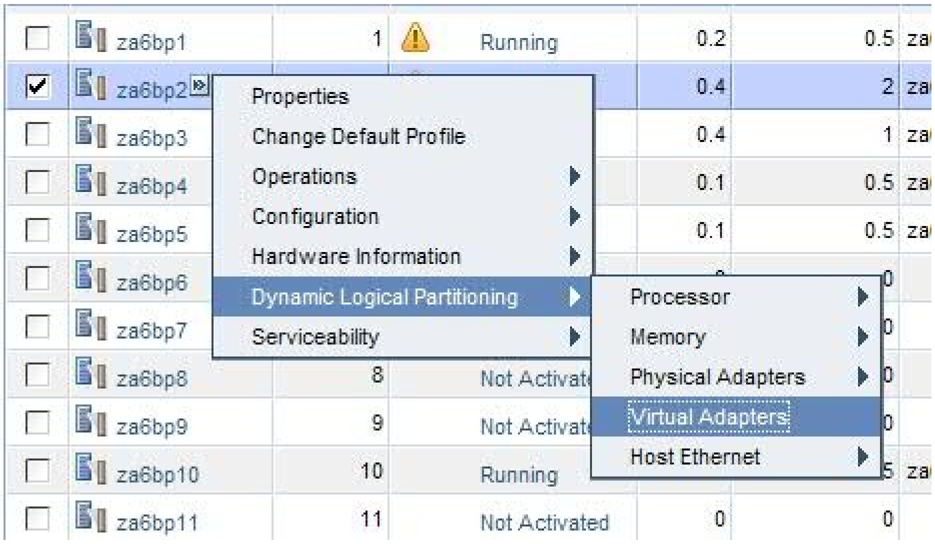

Figure 23:

Clicking on "Virtual Adapters" in Figure 23 will show the screen in Figure 24.

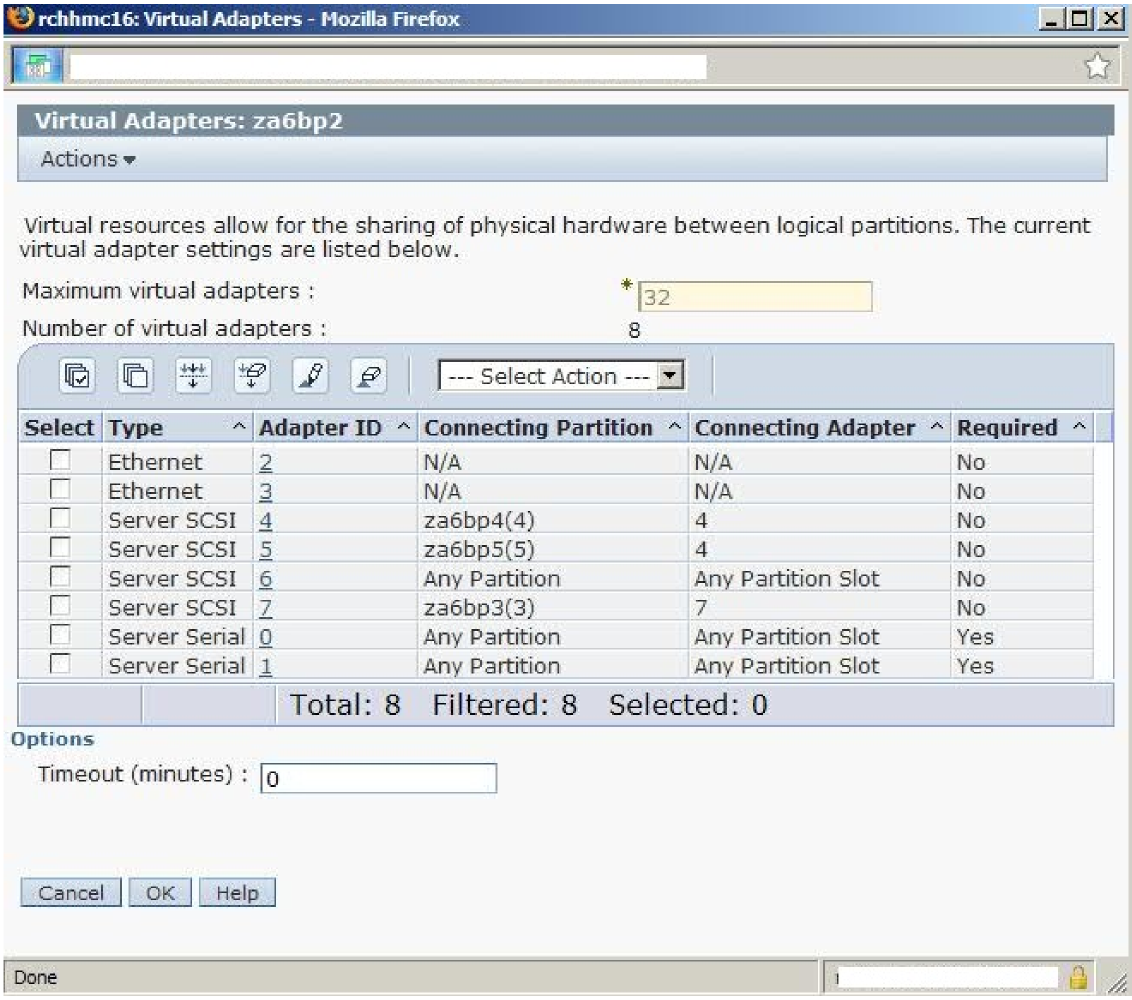

Figure 24:

Figure 24 lists all of the virtual adapters configured in the partition. For those of type ‘Ethernet’, click on the ID in the “Adapter ID” column and the screen in Figure 25 will open up and list the VLAN ID.

If the VLAN ID shown in Figure 25 matches one listed in Figure 22, the client partition would be impacted by the corresponding physical resource listed in Figure 22 becoming unavailable.

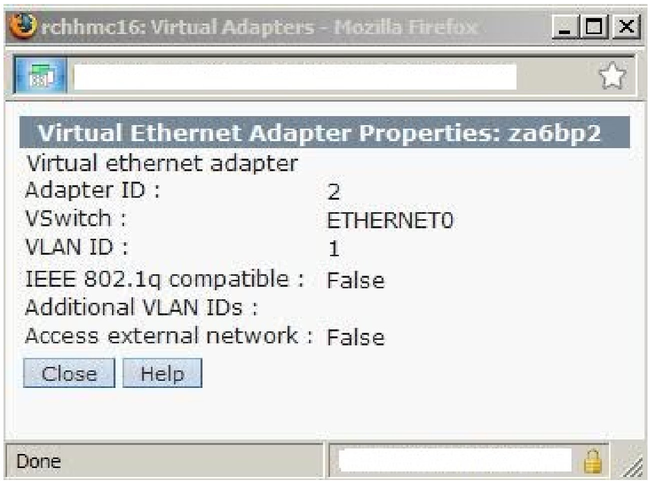

Figure 25:

According to Figure 25, this client partition is using VLAN ID 1. Figure 22 shows that the VIOS partition named ze6bvios is bridging that VLAN to the external network via the physical adapter port ent0. Consequently, this client partition will lose access to the external network if ze6bvios must be shut down or ent0 must be deconfigured.

To determine if a given client partition is impacted, each Ethernet adapter listed in Figure 24 must be examined as described above. The entire process must be repeated for each client partition listed in Figure 23 to determine whether any client partitions will be impacted. Since multiple VLANs can be bridged to the same physical adapter port in the VIOS partition, multiple client partitions may be impacted by a given VIOS owned physical adapter port becoming unavailable.

4 Planning Guidelines and Prerequisites

4.1 Supported Firmware Level

4.1.1 770/780 systems

Please refer to IBM Power Systems Hardware Information Center, CEC Hot Add & Repair Maintenance, Power 770 and Power 780 system firmware and HMC/SDMC levels section

http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/p7ed3/p7ed3cm_matrix_mmb.htm

for the latest update on the minimum and recommended firmware levels for 770 and 780 systems..

4.1.2 795 systems

Please refer to IBM Power Systems Hardware Information Center, CEC Hot Add & Repair Maintenance, Power 795 system firmware and HMC/SDMC levels section

http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/topic/p7ed3/p7ed3cm_matrix_mmb_9119.htm for the latest update on the minimum and recommended firmware levels for 795 systems.

4.2 IBM i Planning Considerations

The following PTF must be installed before a hot node repair, hot node upgrade (memory), or GX adapter hot repair is performed on a system running IBM i partitions:

MF52971

If PTF is not activated, the IBM i partitions must be powered off before the CHARM operation begins.

4.3 Planning Guidelines for CHARM Operations

In addition to the “Best Practices” described in section 3, this section summarizes the design attributes of all the CHARM functions and other operational requirements. The SSR and system administrator must take these factors into account when planning for a CHARM operation.

CHARM Operation	Design Attributes and Planning Guidelines
All	Only one CHARM operation can be performed at a time from one HMC.
All	A second CHARM operation cannot be started until the first one has been completed successfully. If there is an interruption during the first CHARM operation, it must be corrected, and the operation restarted and completed, before starting a second CHARM operation.
All	Multiple CHARM operations must be completed by doing a series of single CHARM operations.
All	Enable the service processor redundancy capability, if it has been disabled, before a

CHARM Operation	Design Attributes and Planning Guidelines
	CHARM operation, except on a Power 770/780 with single node.
All	Review the estimated times for each CHARM operation (section 7.1) to optimize the overall service window.
All	The execution of CHARM procedures and the physical hardware removal and replacement must be performed by a IBM service representative (SSR).
Node Add, Node Upgrade GX add	All serviceable hardware events must be repaired and closed before starting a hot add or upgrade operation. This eliminates the possibility that an existing hardware failure will cause the hot add operation to fail. Prior to a node add operation, the existing nodes in the system must be in a functional state.
Node Repair, GX Repair	Other hardware serviceable events not related to the node or GX adapter that is the target of the CHARM procedure must be repaired before starting the procedure. This eliminates the possibility of other hardware failures interfering with the hot repair operation.
Node Upgrade, Node Repair, GX Repair	• The "Prepare for Hot Repair or Upgrade" (PHRU) utility must be run by the system administrator to get the system in the proper state prior to the start of the hot upgrade or repair procedure by the SSR.
Node Upgrade Node Repair	The following features and capabilities are not supported in conjunction with hot node repair or upgrade: • If the system is clustered to another system using RIO-SAN technology, CHARM is not supported. This technology is used only by i users that are using clustering, switchable towers, and virtual opti-connect technologies. • If the system is clustered to another system using InfiniBand clustering technology, CHARM is not supported. This capability is typically used by High Performance Computing type-clients that have several systems clustered together using an InfiniBand switch. • I/O Processors (IOPs) used by IBM i partitions do not support CHARM. Any IBM i partitions that have IOPs assigned to them must either have the assigned IOPs powered off, or the partition must be powered off to allow the CHARM operation. • 16 GB memory pages, also known as huge pages, do not support memory relocation. Any partitions that have 16 GB pages assigned must be powered off to allow the CHARM operation. Furthermore, due to memory layout changes that take place during the memory relocation phase of a node deactivation, the number of huge pages after a hot node repair may be reduced.
Node Add, GX Add	System firmware enforces the node and GX adapter plugging order. Only the next GX adapter slot or node position based on the plugging order is available.
Node Add, GX Add	When multiple hot adds are planned that include node and GX adapter adds, the GX adapter(s) can be installed in the node before the node is installed in the system. If this order is observed, the memory in the new node will be used for the ~128 MB of memory that is required for the GX adapter. (The node must have approximately 128 MB of memory per GX adapter to support the adapter's translation control entry (TCE) table). The GX adapter in the new node will be activated during the node add

CHARM Operation	Design Attributes and Planning Guidelines
	if the ~128 MB memory requirement is met. If the GX adapter is installed after the new node is installed and activated, the plugging rules for GX adapters will be enforced. In this case, the GX adapter must be installed in another node with another hot add operation.
Node Add, GX Add	For multiple upgrades that include node or GX adapter adds, as well as I/O drawer adds, the node or GX adapter add must be completed first. The I/O drawer can then be added later as a separate concurrent I/O drawer add.
Node Repair, GX Repair	Repair with same FRU type: The node repair procedure does not allow for changes to the hardware beyond the repair of the target FRU. The same FRU type must be used to replace the failing FRU and no additional hardware can be added or removed during the procedure. For example, if a 4GB DIMM fails, it must be replaced with a 4GB DIMM – not a 2GB or 8GB DIMM. If a RIO GX adapter is being repaired, it must be replaced with a RIO GX adapter, not an InfiniBand GX adapter.
Node Repair, GX Repair	Hardware upgrades and downgrades are not supported during a repair operation.

Table 3: Planning Guidelines for CHARM Operations

4.4 Additional Planning Checklist

For the latest version of the planning checklist, please refer to IBM Power Systems Hardware Information Center, CEC Hot Add & Repair Maintenance section

http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/p7ed3/ared3kickoff.htm

5 CEC Hot Add & Repair Maintenance Functions

This section provides additional details of the CHARM functions.

The HMC/SDMC provides the user interfaces for the system administrator and SSR to perform their respective tasks during the CHARM operation.

For the hot node upgrade, hot node repair, or GX adapter repair, the system administrator uses the "Prepare for Hot Repair or Upgrade" (PHRU) utility on the HMC/SDMC to determine the impact to the resources, then uses the DLPAR function and other operating system or system management tools to free up processor, memory, and/or I/O resources if necessary.

The SSR uses the Repair & Verify (R&V) function on the HMC/SDMC to perform all the CHARM procedures. R&V provides the step-by-step guided instructions to the SSR to complete the CHARM procedure. For an MES add or upgrade, the WCII instructions provide the physical installation instructions as well as other information. The guided procedures provided by the HMC/SDMC or WCII must be precisely followed to minimize potential unrecoverable error during a CHARM operation.

The HMC/SDMC uses the RMC network connection with the logical partitions (LPARs) during the execution of the PHRU utility and the CHARM procedure. The HMC/SDMC communicates with partitions to determine impacts to the partitions and to deactivate impacted resources.

The HMC/SDMC uses the service network to send queries and requests to the service processor and hypervisor for the overall functional flow. The service processor and hypervisor perform specific CHARM tasks such as a system "readiness check", administrative fail-over (for redundant service processor and clock functions), node evacuation, hardware deactivation, power off, power on, CEC hardware diagnostics, hardware activation, resource configuration, resource integration and error handling/reporting during the CHARM operation.

5.1 Hot Node Add

This function allows an SSR to add a node to a system to increase the processor, memory, and I/O capacity of the system. After the physical node installation, the system firmware activates and integrates the new node hardware to the system. The new processor and memory resources are available for the creation of new partitions or to be dynamically added to existing partitions. Additional GX adapters (I/O hubs) can be physically added to the node and activated as part of the hot node add operation. After the completion of the node add operation, additional I/O expansion units can be attached to the new GX adapter(s) in the new node in a separate concurrent I/O expansion unit add operation.

5.1.1 Node Add Functional and Procedural Flow

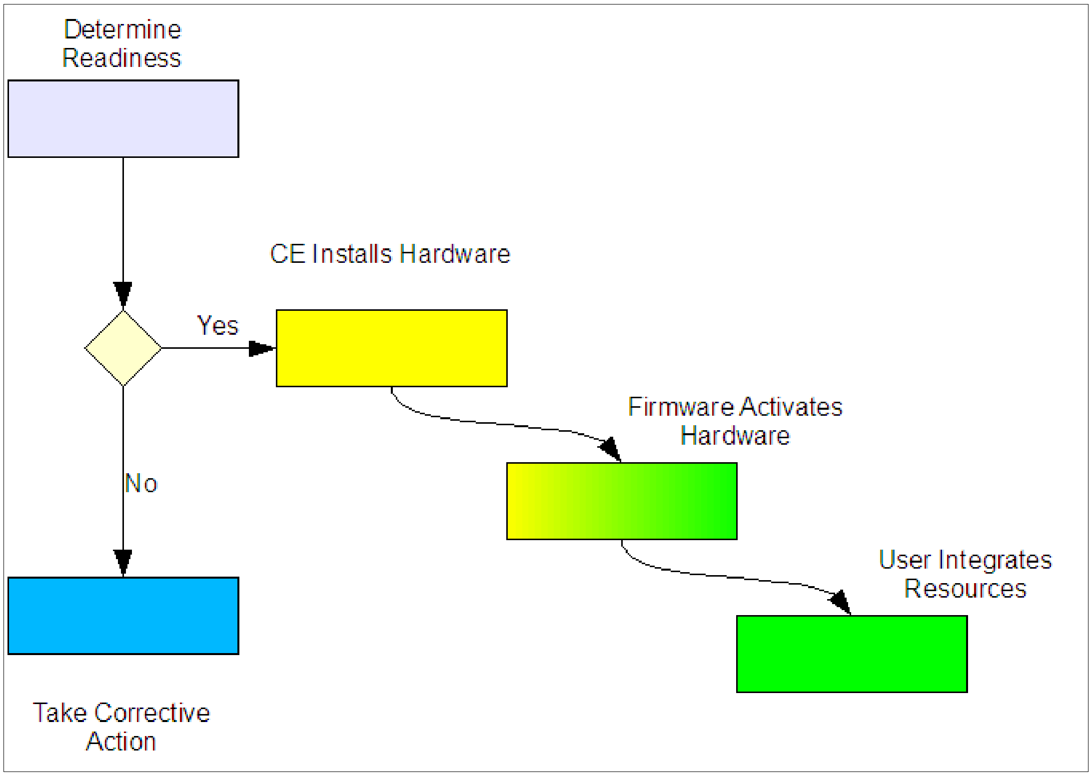

Figure 26: Hot Node Add Overview

The overall flow consists of the following major tasks performed by the system administrator, SSR, and system firmware.

1. MES planning and preparation – The system administrator and SSR coordinate the MES activity, and review the prerequisites and planning guidelines for a hot node add.
2. Determine Readiness – This step verifies that the system is ready for the hot node add operation. The firmware checks the system state to ensure that the system is ready.
3. SSR Installs Hardware – The step by step instructions are presented to the SSR to install the new node.
4. Firmware Activates Hardware – Power is supplied to the new node, new hardware is initialized, CEC diagnostics are run to verify the new hardware, then it is electrically and logically added to the system. The resources in the node are integrated into the system.
5. System Administrator Allocates Resources – New processor and memory can be assigned to partitions.

Note: When a Power 770/780 with a single node is upgraded to 2 or more nodes using the node add function, the redundant service processor in the second node will be enabled and activated automatically.

5.1.2 Node Plugging Rules

The firmware automatically determines where the next node can be added into the system. The guided procedure on the HMC/SDMC will direct the SSR to place the node in the location that follows the system design plugging order.

5.1.3 Allocation of Resources

With proper planning, the newly-added processor and memory resources can be assigned to logical partitions without a partition reboot or entering capacity upgrade on demand (CUoD) activation codes. If the system administrator wants to dynamically add the new processor and memory resources to activated partitions after a hot add operation, the profiles that were used to activate the partitions must have the maximums for processors and memory set to the appropriate values. When the hot add operation is complete, the processor and memory resources can be added to the partitions by the system administrator using dynamic logical partitioning (DLPAR) operations.

5.2 Concurrent GX Adapter Add

This function allows an SSR to add a GX adapter to increase the I/O capacity of the system. To concurrently add GX adapters, additional planning is required (refer to section 5.2.1 for more details).

After the GX adapter is installed, the system firmware activates and integrates the new I/O hub into to the system. After the completion of the concurrent GX adapter add operation, an additional I/O expansion unit can be attached to the new GX adapter as a separate concurrent I/O expansion unit add operation.

5.2.1 GX Adapter Memory Reservations

Prior to the concurrent GX adapter add operation, system memory must be reserved for the I/O address translation table (TCE) for each adapter. The TCE memory is contiguous and must be reserved during system IPL. Each GX adapter requires approximately 128MB of TCE memory.

The system firmware automatically makes GX adapter memory reservations to support a concurrent GX adapter add. The default TCE memory reservations are as follows:

Platform	*Default Memory Reservation*
Power 770/780	1 GX adapter (~128MB) maximum, if an empty slot is available
Power 795	2 GX adapter (~256 MB) maximum, if empty slots are available

Table 4: Default GX Adapter Reservation

The default value can be changed at the system administrator's discretion by changing the "Total number of slots to reserve" for GX adapters from zero to the total number of empty GX adapter slots in the system using the service processor's advanced system management interface (ASMI) menus. Section 6.1 provides additional detail procedure on accessing the ASMI menu from the HMC/SDMC and a example of the user interface.

Once the system administrator saves the settings, they become active on the next system IPL. The firmware attempts to allocate the memory based on the desired number of slots to reserve, but may not be able to due to the system memory configuration. The more configured memory that is available, the higher the probability that all of the requests will be fulfilled. Once the system reaches the runtime state, the menu can be viewed again to show exactly how many reservation requests were satisfied.

Because reservation requests are fulfilled during a system IPL, reservation requests that are made while the system is powered on do not take effect until the next system IPL. A reservation request made while the system is powered off is fulfilled during the next system IPL. Once a reservation is made, it will remain the same through system power off, power on/IPL.

Once a GX adapter is successfully added through a concurrent GX adapter add operation, the memory that was reserved is utilized, and the firmware decrements the total number of reserved slots. For example, if a system administrator requests three slots, when a concurrent GX adapter add operation is performed, the total number of slots reserved decreases to two.

5.2.2 Plugging Rules

In order to optimize performance, the firmware automatically determines where a GX adapter can be added to the system. A system administrator only needs to specify how many GX adapters are desired, and the firmware enforces the plugging rules.

5.2.3 Functional and Procedural Flow

The GX flow is similar to the hot node add flow, with primary difference being the hardware. Please refer to the node add section for details.

5.2.4 Allocation of Resources

After a concurrent GX adapter add operation, the new adapter will be available to the system.   An I/O expansion unit can be added by following the appropriate procedure, and the resources in the new I/O expansion unit will then be available for the system administrator to assign to the partitions.

5.3 Hot Node Upgrade (Memory)

This function allows an SSR to increase the memory capacity in a system by adding additional memory DIMMs to a node, or upgrading (exchanging) existing memory with higher-capacity memory DIMMs. The system must have 2 or more nodes to utilize the hot node upgrade function. Since the node that is being upgraded is functional and possibly running workloads prior to starting the hot upgrade operation, the system administrator uses the "Prepare for Hot Repair or Upgrade" utility, system management tools, and operating system (OS) tools to prepare the system. During the hot node upgrade operation, the system firmware performs the node evacuation by relocating workloads from the target node to other node(s) in the system and logically isolating the resources in the target node. It then deactivates and electrically isolates the node to allow the removal of the node for the upgrade.

After the physical node upgrade, the system firmware activates and integrates the node hardware with additional memory into the system. After the completion of a hot node upgrade operation, the system administrator then restores the usage of processor, memory, and I/O resources, including redundant I/O configurations if present. The new memory resources are available for the creation of new partitions or to be dynamically added to existing partitions by the system administrator using DLPAR operations.

5.3.1 Prepare for hot node upgrade

Since the node that is being upgraded may be running workloads prior to starting the hot upgrade operation, the HMC/SDMC "Prepare for Hot Repair or Upgrade" (PHRU) utility is used by the system administrator to identify the impact of and get the system in the proper state for the hot node upgrade operation. The PHRU utility may instruct the user to de-configure impacted I/O resources using operating system tools or free up additional processor and/or memory capacity using DLPAR operations or other means in preparation for the operation.

5.3.2 Functional and Procedural Flow

The following diagram shows the basic hot node upgrade (memory) flow:

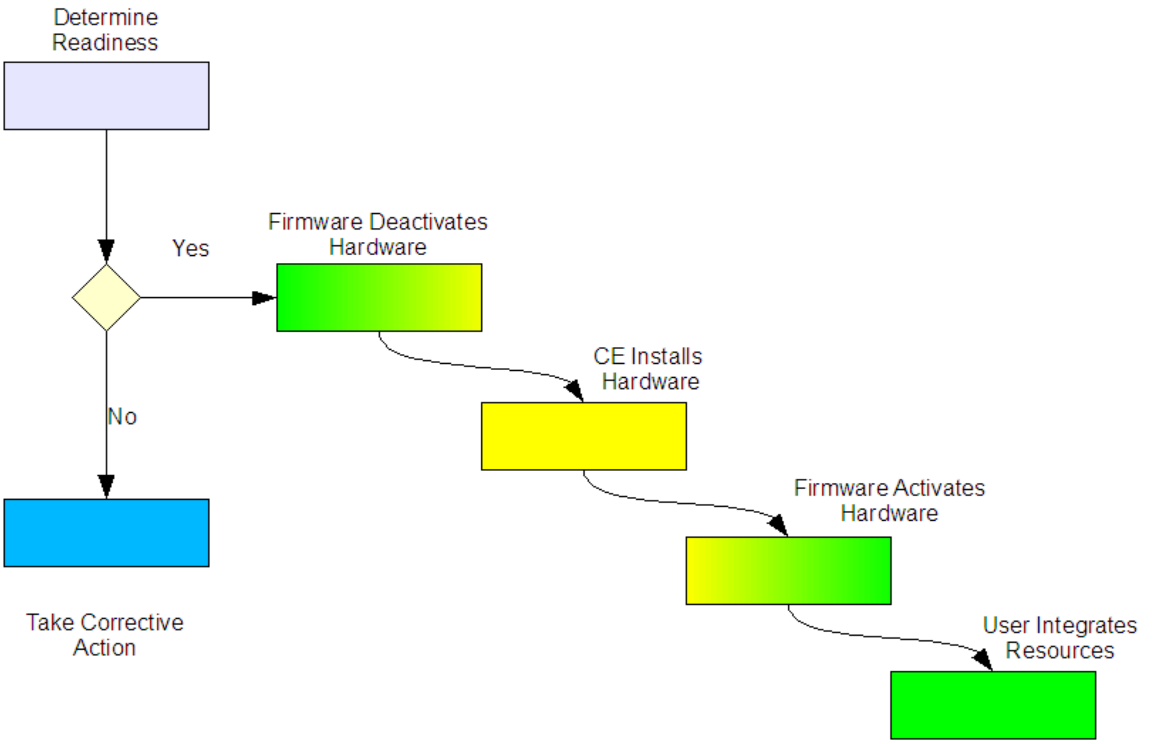

Figure 27: Hot Node Upgrade Overview

The overall flow consists of the following major tasks performed by the system administrator, SSR, and system firmware.

1. Planning and Preparation - The system administrator and SSR coordinate the MES activity, and review the prerequisites and planning guidelines for a hot node upgrade (memory). The system administrator uses the PHRU utility to identify the impacted resources, then if necessary uses other OS and system management tools to free up system resources in preparation for the operation.
2. Determine Readiness – This step verifies that the system is ready for the hot node upgrade operation. The firmware checks the system state to ensure that the system is ready and will remain functional when the target node is removed. Then the system firmware verifies that the required resources have been freed up by the system administrator during the planning and preparation phase. If the required resources have not been freed up, the system firmware does not allow the operation to continue, and the system administrator must take necessary steps to prepare the system for the operation.
3. Firmware Deactivates Hardware –During this step of the procedure, the system firmware logically isolates the I/O resources in or dependent on the target node, and relocates processor and memory workloads from the target node to other node(s) in the system. It then deactivates, electrically isolates, and powers off the node to allow the removal of the node for the upgrade.

4. SSR Removes/Installs Hardware – The SSR follows the step by step instructions to remove the node, install new memory, and reinstall the node.
5. Firmware Activates Hardware – Power is applied to the node, new hardware is initialized, CEC diagnostics are run to verify the new hardware, then the firmware electrically and logically adds the node hardware back to the system. The resources in the node are reintegrated into the system.
6. System Administrator Allocates Resources – The steps taken by the system administrator during the planning and preparation phase to free up processor, memory and I/O resources can now be reversed manually at the system administrator's discretion. The new memory resources are available for new partitions or to be dynamically added to existing partitions by the system administrator using DLPAR operations.

5.4 Hot Node Repair

This function allows an SSR to repair defective hardware in a node of a system. The system must have 2 or more nodes to utilize the hot node repair function. Since the node that is being repaired may be running workloads prior to starting the hot repair operation, the system administrator uses the "Prepare for Hot Repair or Upgrade" utility to identify impacts to the system, and then if necessary uses system management tools and OS tools toget the system in the proper state for the operation. During the hot node repair operation, the system firmware logically isolates the I/O resources in or dependent on the target node, and relocates processor and memory workloads from the target node to other node(s) in the system. It then deactivates and electrically isolates the node to allow the removal of the node for repair.

After the physical node repair, the system firmware activates and integrates the node hardware back into the system. After the completion of a hot node repair operation, the system administrator restores the usage of processor, memory, and I/O resources, including redundant I/O configurations if present.

5.4.1 Node Hardware State

This function can be used to repair a node with hardware in the following states:

1. All hardware in the node is functional – Some hardware may have had a predictive error (i.e. a recoverable error over threshold) but continues to function.
2. The majority of the hardware in the node is functional – Some hardware may have had an unrecoverable error, and was removed from the system configuration.
3. The node is not functional – Some node hardware had a specific type of hardware error that required all the node hardware to be removed and electrically isolated from the system configuration during system IPL.

5.4.2 Prepare for hot node repair

This step is the same as the hot node upgrade. Please refer to section 5.3.1 for details.

5.4.3 Functional and Procedural Flow

The hot node repair flow is similar to the hot node upgrade, with the differences in planning for repair and replacing defective hardware in place of an upgrade operation. Please refer to section 5.3.2 for details.

5.5 Hot GX Adapter Repair

This function allows an SSR to repair a defective GX adapter in the system. The GX adapter may still be in use prior to starting the hot repair operation, so the system administrator uses the "Prepare for Hot Repair or Upgrade" utility to identify impacts to the system, and then if necessary uses system management tools and OS tools to get the system in the proper state for the operation. During the hot GX repair operation, the system firmware logically isolates I/O resources dependent on the target GX adapter, and then deactivates and electrically isolates the GX adapter to allow physical removal for repair.

After the repair, the system firmware activates and integrates the hardware into the system. After the completion of the hot repair operation, the system administrator restores the usage of I/O resources, including redundant I/O configurations if present.

5.5.1 GX Adapter Hardware State

This function can be used to repair a GX adapter with hardware in the following states:

1. The adapter is functional – Some part of hardware may have had a predictive error (i.e. a recoverable error over threshold) but continues to function.
2. The adapter had an unrecoverable error during system runtime and it is now in a "failed" state. The I/O resources connected through this I/O hub are unavailable.
3. The adapter had an unrecoverable error and it was removed and electrically isolated from the system configuration during system IPL. The I/O resources connected through this I/O hub are unavailable during system IPL and partition boot.

5.5.2 Prepare for hot GX repair

The HMC/SDMC "Prepare for Hot Repair or Upgrade" (PHRU) utility is designed for the system administrator to identify the impact to the I/O resources of a hot GX repair operation. Once the impacted system resources are identified, the system administrator uses the OS tools to remove the I/O resources from usage, or shuts down partition(s) to free up the I/O resources in preparation for hot repair.

5.5.3 Functional and Procedural Flow

The GX repair flow is similar to the hot node upgrade, with the differences in planning for I/O hub repair and replacing defective hardware in place of a node upgrade operation. Please refer to section 5.3.2 for details.

5.6 System Controller Repair

This function allows an SSR to repair a defective system controller. After the repair, the system firmware activates and integrates the hardware into the system.

5.6.1 System Controller Hardware State

This function can be used to repair a system controller with hardware in the following states:

1. The system controller is functional – Some part of hardware may have had a predictive error (i.e. a recoverable error over threshold) but continues to function.
2. The system controller had an unrecoverable error during system runtime and it is now in a "failed" state.
3. The system controller had an unrecoverable error and it was removed and electrically isolated from the system configuration during system IPL.

5.6.2 Functional and Procedural Flow

The system controller repair flow is similar to the hot node upgrade, with the differences in planning for repair and replacing defective hardware in place of an upgrade operation. The system controller operation will not perform any node evacuation steps that include the use of the PHRU utility to identify the resources and operating systems that are impacted. Also, the system controller repair operation does not use system management tools that free up the node resources in preparation for evacuation. Please refer to section 5.3.2 for details.

6 User Interfaces

The HMC/SDMC and the server's Advanced Systems Management Interface (ASMI) provide the interfaces for the activities related to CEC Hot Add & Repair Maintenance actions.

6.1 Reserving TCE Memory for Concurrent GX Adapter Add

The system administrator can alter the TCE memory reservation for a GX adapter from zero to the total number of empty slots in the system through the server's ASM interface.

To verify and reserve slots for installing GX adapters using HMC, complete the following steps:

1. In the HMC navigation pane, select Systems Management > Servers.
2. In the Contents of: Servers pane, select the check box in the Select column for the target managed server.
3. Select Tasks > Operations > Advanced Systems Management (ASM).
4. Click OK in the Launch ASM Interface window. Click Accept in any Certificate signer not found window that might appear.
5. Log in to the ASM panel.
6. Expand the Concurrent Maintenance list.
7. Select RIO/HSL Adapter Slot Reservation.

To verify and reserve slots for installing GX adapters using SDMC, complete the following steps:

1. On the SDMC welcome page, under the Resources tab, select the server
2. Select Actions > Operations > Launch Advanced System Management (ASM)
3. Log on to ASMI using the administrator user ID and password
4. Expand the Concurrent Maintenance list.
5. Select RIO/HSL Adapter Slot Reservation.

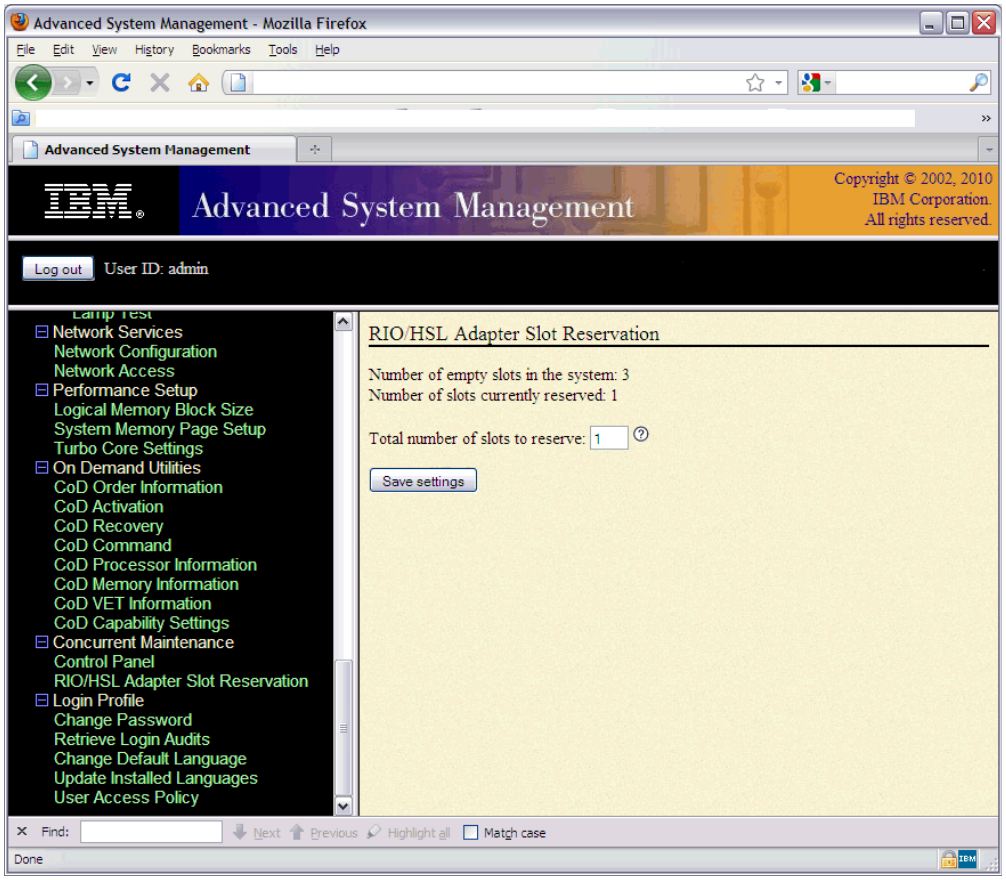

Figure 28: Altering GX Adapter Memory Reservation with ASM

Please also refer to "Reserving RIO/HSL Adapter Slots" in the "Managing the Advanced System Management Interface" document in the IBM Power System Hardware Information Center at:

http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/topic/p7hby/p7hby_77x_78x.pdf

for the latest version of this procedure.

6.2 Prepare for Hot Repair or Upgrade Utility

The Prepare for Hot Repair or Upgrade (PHRU) utility is a tool for the system administrator to identify the impacts to the system resources in preparation for a hot node repair, hot node upgrade, or hot GX adapter repair operation. The utility identifies I/O resources that must be deactivated, and processor and memory capacity that must be freed up before the system firmware will allow the operation to occur. It also identifies any other aspects of the configuration that would prevent the start of the operation. .

Once the impacted processor, memory or I/O resources are identified, they must be freed up or deactivated by the system administrator before the start of the hot repair or upgrade procedure.

This utility also runs automatically at the start of a hot repair or upgrade procedure to confirm that the system is in the proper state for a successful operation.

For hot node add or GX add operation, there is no impact to existing system resources; this step is not required.

This utility is available with HMC V7R7.2.0 + MH01235. It replaces the "Display Service Effect" utility in the prior versions of the HMC code.

The PHRU Utility is also available on the SDMC. To access the PHRU utility on the SDMC, go to the SDMC welcome page and under the resources tab, select the server > "Actions" pull down menu > Service and Support > Hardware > Prepare for Hot Repair/Upgrade

Figure 29, 30, and 31 show the HMC user interfaces for the PHRU utility.

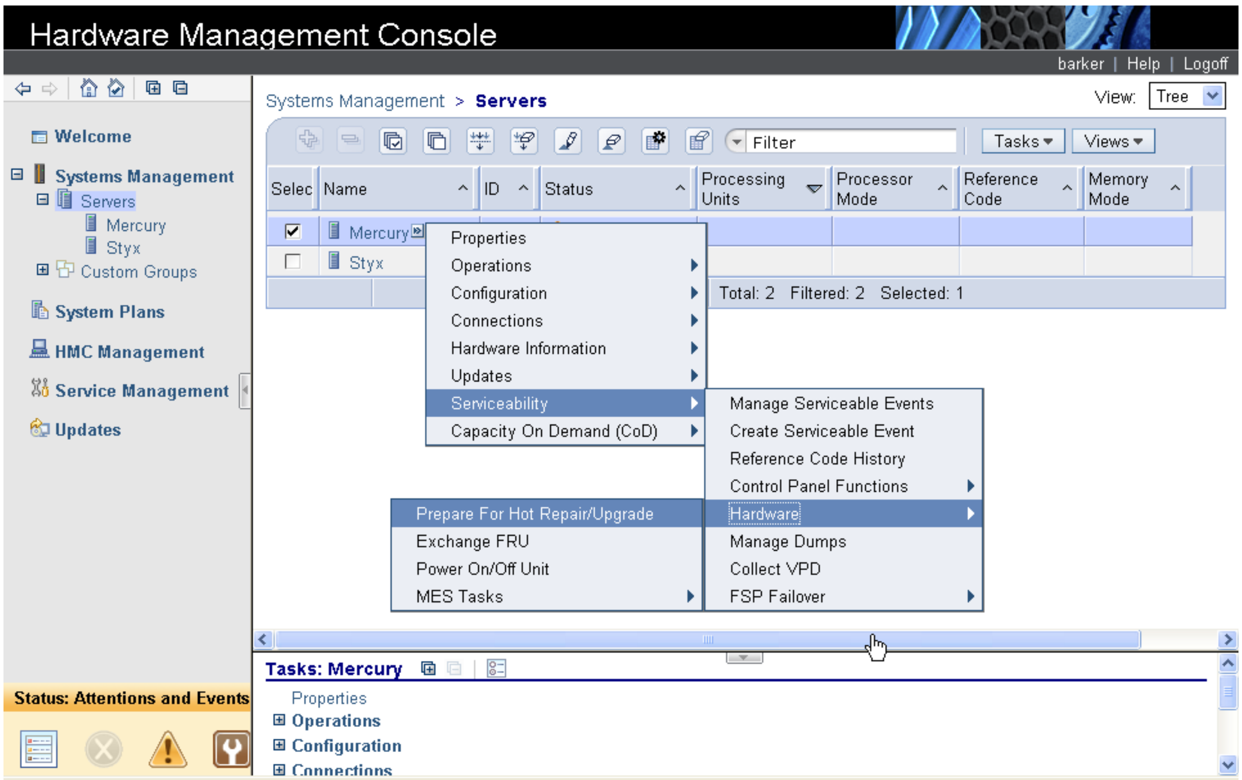

Figure 29: HMC menu navigation for the PHRU utility

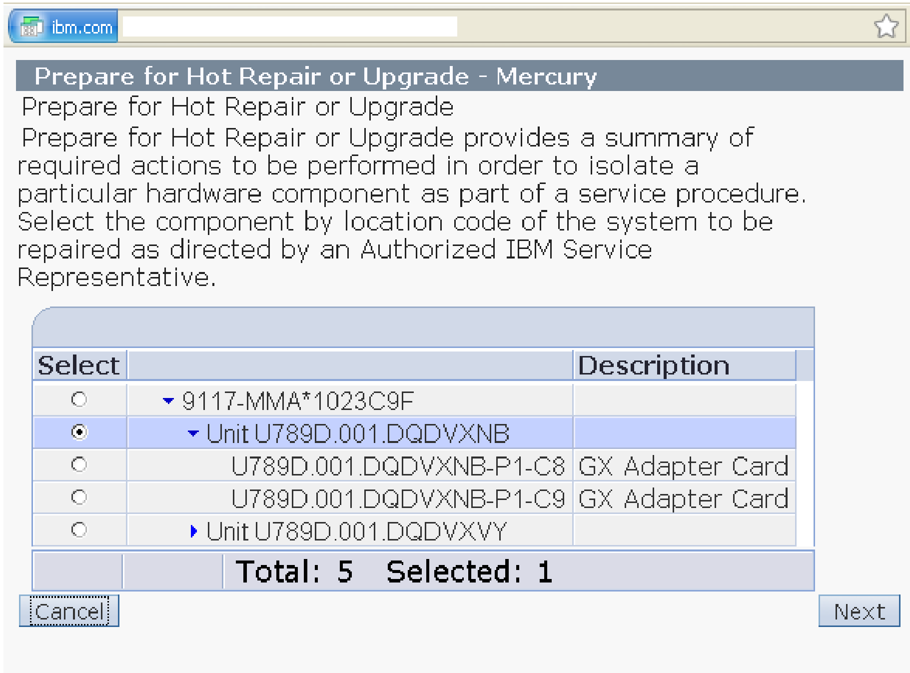

Figure 30: Prepare for Hot Repair or Upgrade utility (1)

Figure 31: Prepare for Hot Repair or Upgrade utility (2)

The procedure for using and responding to the information provided by the Prepare for Hot Repair or Upgrade is described in the IBM Power Systems Hardware Information Center, CEC Hot Add & Repair Maintenance section at the following location:
http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/p7ed3/ared3nodeevac.htm

6.3 Initiating a Hot Add & Repair Maintenance Procedure

CEC Hot Add & Repair maintenance operations are HMC/SDMC-guided procedures performed by an SSR. Repair procedures can be initiated directly from a serviceable event which calls out the FRU(s) to be repaired. They can also be started via the “Exchange FRU” option in the HMC Serviceability > Hardware task or SDMC Service and Support > Hardware task. Procedures to add FRUs can be initiated by using the “Add FRU” option under the HMC Serviceability> Hardware task or SDMC Service and Support > Hardware task. For an MES add or upgrade, the WCII instructions provide the physical installation instructions.

6.3.1 Initiating a Service Action from a Serviceable Event

When a serviceable event is used to initiate a repair procedure, the FRU(s) called out in the serviceable event will be repaired. If there are multiple FRUs called out in the serviceable event, the FRUs may be either grouped for repair or listed in order of priority. If the FRUs are grouped together for repair, each of the FRUs will be replaced in the guided procedure. Grouped sets of FRUs are most common for DIMMs and processors. If there are multiple FRUs in the serviceable event, but they are not grouped, the FRUs will be prioritized. The repair procedure will provide the steps to repair the first FRU in the priority list. If replacement of that FRU does not fix the problem, the repair procedure will direct the replacement of the next highest priority FRU in the serviceable event. Figure 32 shows a repair procedure being initiated from a serviceable event using the HMC user interface. The SDMC provides a similar interface under System Status and Health > Problems > Active problem event > Actions > Repair.

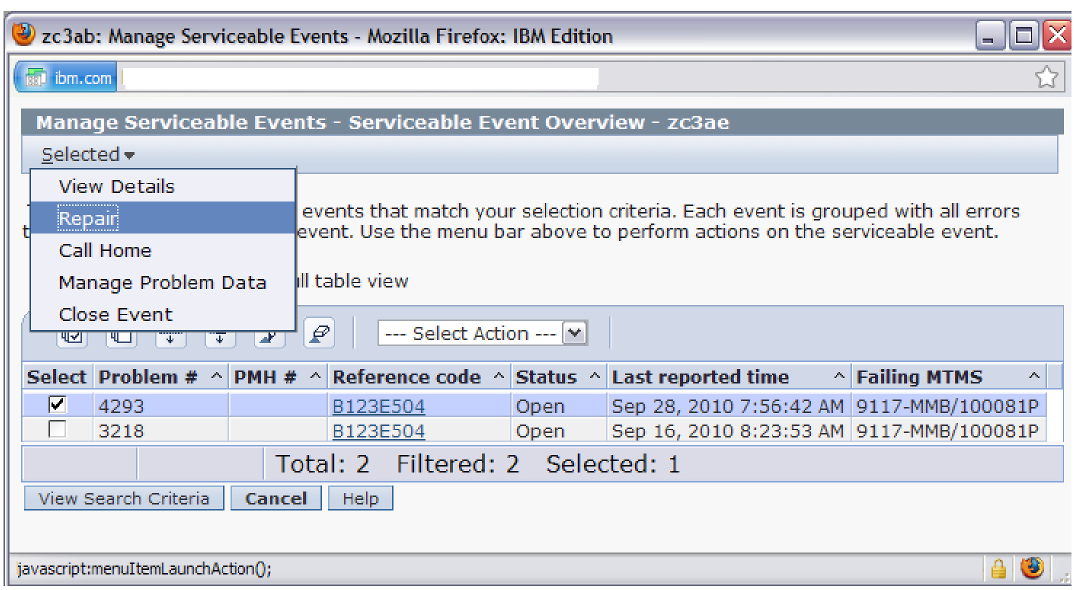

Figure 32: HMC interface for Initiating a Service Action from a Serviceable Event

6.3.2 Initiating a Service Action from the Exchange/Add FRU panels

The Exchange FRU and Add FRU options allow for the selection of a FRU type followed by the selection of a specific location code. For Add FRU, the location code will be automatically chosen based on the plugging order of the FRU type selected (if plugging order must be followed). Figure 33 shows the selection of a FRU type on the Exchange FRU screen. Figure 34 shows the selection of the location code of a FRU to exchange.

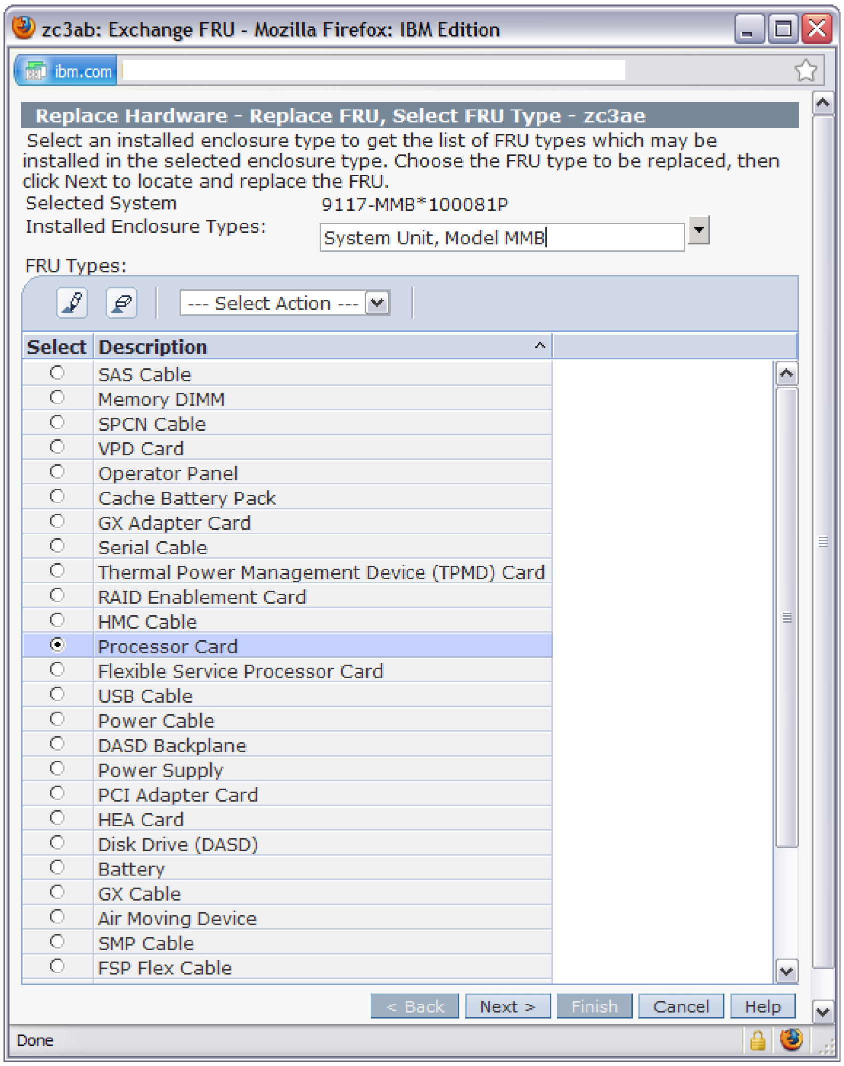

Figure 33: HMC Interface for Selecting a FRU Type Using Exchange FRU

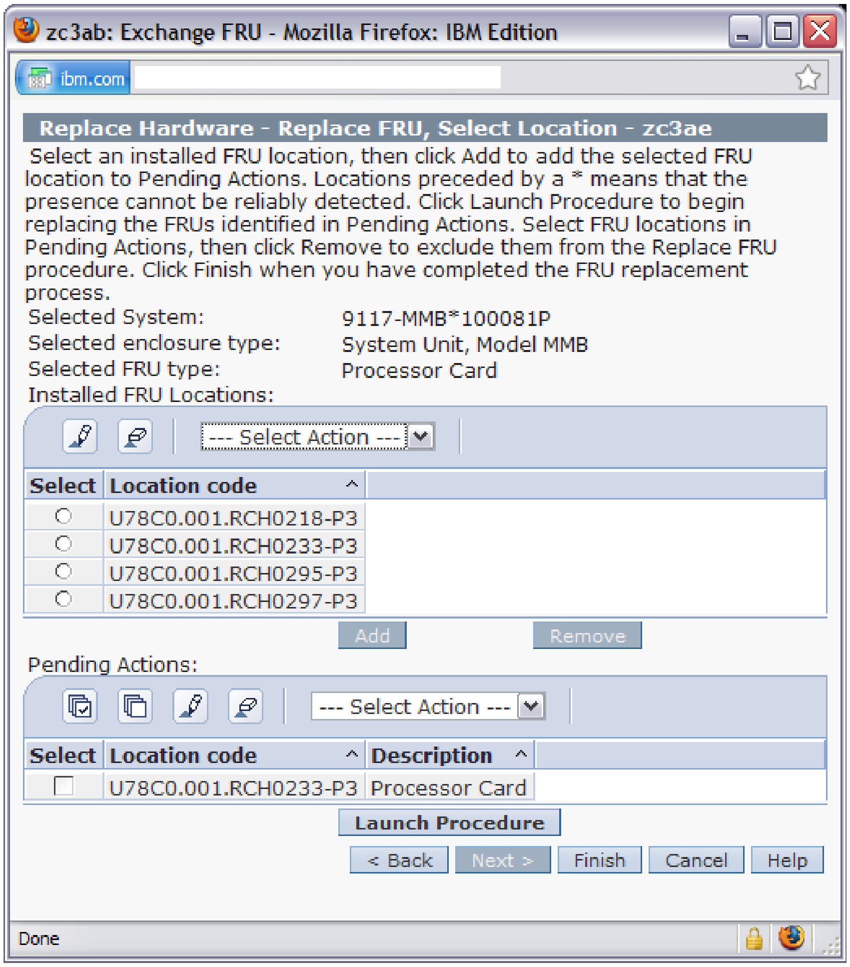

Figure 34: HMC Interface for Selecting a System Processor Card Location Code to Exchange

7 Appendix

7.1 Estimate Time for Hot Add, Upgrade and Repair Operation

The following matrices provide estimated times in minutes for each activity (by role) in a CHARM operation on Power 770/780 (see Table 5a) and 795 (see Table 5b) Servers. The times shown are in minutes, and are approximate.

In a typical hot repair or upgrade scenario, the system administrator and SSR should coordinate their activities in advance to optimize the end-to-end operation time. When the system administrator's tasks are complete, the SSR can proceed with the hot repair or upgrade operation. When the SSR has completed the operation, the system will be turned back over to the system administrator to be returned to its original state (at his or her discretion).

As stated in the previous section, only one CHARM operation can be performed at a time from one HMC/SDMC. The estimated times below are for a single operation. For a large MES upgrade with multiple nodes or GX adapters, careful planning by the system administrator and SSR must be done to optimize the overall upgrade window. If multiple SSRs are available to perform upgrade activities in parallel, a non-concurrent upgrade with system powered down may result in a shorter time window.

	System Admin. Time (minutes)		**SSR Time** (minutes)			**Total Time** (minutes)	
Operation	**Prepare for Node/GX Evacuation**	**Resource Allocation/ Restore**	**Memory Relocation (32-512GB)**	**Firmware Deactivate/ Activate**	**Physically Remove/ Install**	**System Admin.**	**SSR**
Node Add	N/A	~30	N/A	~30-45	~60	~30	~90-105
Node Upgrade	~30-60	~30	~11-77	~25-40	~15	~60-90	~60-120
Node Repair	~30-60	~30	~11-102	~25-40	~15-20	~60-90	~60-160
GX Add	N/A	~15	N/A	~10	~5	~15	~15
GX Repair	~10-30	~15	N/A	~15-20	~8	~25-45	~23-28

Table 5a: Estimated time for CHARM operation on a Power 770/780 Server

	System Admin. Time (minutes)		SSR Time (minutes)			Total Time (minutes)	
Operation	**Prepare for Node/GX Evacuation**	**Resource Allocation/ Restore**	**Memory Relocation (64-1024GB)**	**Firmware Deactivate/ Activate**	**Physically Remove/ Install**	**System Admin.**	**SSR**
Node Add	N/A	~45	N/A	~30-45	~90	~45	~120-135
Node Upgrade	~30-60	~30	~25-155	~25-40	~45-90	~60-90	~95-285
Node Repair	~30-60	~30	~25-205	~25-40	~45-90	~60-90	~95-335
GX Add	N/A	~15	N/A	~10	~10	~15	~20
GX Repair	~10-30	~15	N/A	~15-20	~10	~25-45	~25-30

Table 5b: Estimated time for CHARM operation on a Power 795 Server

7.2 FRU Type and Location Code Reference

7.2.1 Hot Add/Upgrade

The following tables list the FRUs which are supported for hot node add:

Power 770/780	
FRU Type	**Location Code**
Node/Building Block / Drawer	Uxxyy.001.abcdefg

Table 6a: FRUs supported for Hot Node Add

Power 795	
FRU Type	**Location Code**
Node/Processor Book	Uxxyy.001.abcdefg-Pm (m = 2 to 8)

Table 6b: FRUs supported for Hot Node Add

The following tables list the FRUs that can be added or upgraded through a hot node upgrade:

Power 770/780	
FRU Type	**Location Code**
Memory DIMM	Uxxyy.001.abcdefg-P3-Cn

Table 7a: FRUs Supported for Hot Node Upgrade

Power 795	
FRU Type	**Location Code**
Memory DIMM	Uxxyy.001.abcdefg-Pm-C2 through Uxxyy.001.abcdefg-Pm-C38 (m = 2 to 9)

Table 7b: FRUs Supported for Hot Node Upgrade

The following tables list the FRUs which are supported for GX adapter add.

Power 770/780	
FRU Type	**Location Code**
GX Adapter Card	Uxxyy.001.abcdefg-P1-C8, Uxxyy.001.abcdefg-P1-C9

Table 8a: FRUs Supported for GX Adapter Add

Power 795	
FRU Type	**Location Code**
GX Adapter Card	Uxxyy.001.abcdefg-Pm-C39 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C40 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C41 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C44 (m = 2 to 9)

Table 8b: FRUs Supported for GX Adapter Add

7.2.2 Hot Repair

The following tables list the FRUs and corresponding location codes which can be replaced through a hot repair procedure:

Power 770/780	
FRU Type	**Location Code**
SCM (Processor Single Chip Module)	Uxxyy.001.abcdefg-P3-C22, Uxxyy.001.abcdefg-P3-C25
Memory DIMM	Uxxyy.001.abcdefg-P3-C1 through Uxxyy.001.abcdefg-P3-C4, Uxxyy.001.abcdefg-P3-C6 through Uxxyy.001.abcdefg-P3-C13, Uxxyy.001.abcdefg-P3-C17 through Uxxyy.001.abcdefg-P3-C20
Processor card regulator (VRM)	Uxxyy.001.abcdefg-P3-C21, C23, C24, C26, C5, C14, C16
TPMD card	Uxxyy.001.abcdefg-P3-C15
Processor card	Uxxyy.001.abcdefg-P3
Mid-plane	Uxxyy.001.abcdefg-P1
I/O card	Uxxyy.001.abcdefg-P2
FSP/Clock card (drawer 1&2)	Uxxyy.001.abcdefg-P1-C1
Clock Passthru card(drawer 3&4)	Uxxyy.001.abcdefg-P1-C1
Battery	Uxxyy.001.abcdefg-P1-C1-E1
HEA Card	Uxxyy.001.abcdefg-P2-C8
DASD Backplane	Uxxyy.001.abcdefg-P2-C9
RAID card Battery	Uxxyy.001.abcdefg-P2-C9-C1
GX Adapter Card (block FRU repair with others)	Uxxyy.001.abcdefg-P1-C2, C3

Table 9a: FRUs Supported for Hot Node Repair

Power 795	
FRU Type	**Location Code**
Memory DIMM	Uxxyy.001.abcdefg-Pm-C2 through Uxxyy.001.abcdefg-Pm-C38 (m = 2 to 9)
SCM (Processor Single Chip Module)	Uxxyy.001.abcdefg-Pm-C22 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C23 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C31 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C32 (m = 2 to 9)
Thermal Power Management Device	Uxxyy.001.abcdefg-Pm-C30 (m = 2 to 9)
Processor Book	Uxxyy.001.abcdefg-Pm (m = 2 to 9)
Processor Book FSP- Node Controller	Uxxyy.001.abcdefg-Pm-C43 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C42 (m = 2 to 9)
Processor Book – LED controller	Uxxyy.001.abcdefg-Pm-C1 (m = 2 to 9)
GX Adapter Card	Uxxyy.001.abcdefg-Pm-C39 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C40 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C41 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C44 (m = 2 to 9)

Table 9b: FRUs Supported for Hot Node Repair

The following tables list the FRUs which are supported for GX adapter repair:

Power 770/780	
FRU Type	**Location Code**
GX Adapter Card	Uxxyy.001.abcdefg-P1-C2, C3

Table 10a: FRUs Supported for Hot GX Adapter Repair

Power 795	
FRU Type	**Location Code**
GX Adapter Card	Uxxyy.001.abcdefg-Pm-C39 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C40 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C41 (m = 2 to 9) Uxxyy.001.abcdefg-Pm-C44 (m = 2 to 9)

Table 10b: FRUs Supported for Hot GX Adapter Repair

The following tables list the FRUs which are supported for system controller repair:

Power 795	
FRU Type	**Location Code**
System Controller	Uxxyy.001.abcdefg-P1-C2 Uxxyy.001.abcdefg-P1-C5

Table 10c: FRUs Supported for System Controller Repair

7.3 Terminology

7.3.1 System Resource, Function and Operation

CEC: Central Electronic Complex, the heart of a large computer system, i.e. processor, memory, I/O hub, and associated controlling and supporting hardware (e.g. system clock, service processor, etc). The CEC does not include I/O drawers, I/O adapters, I/O devices, power or cooling hardware.

Node: a physical group of processors, memory, and I/O hubs in the system.

GX Adapter: an I/O hub which connects I/O adapters to the processor and memory in the system.

DIMM: Dual Inline Memory Module, a hardware part for computer memory.

FRU: Field Replaceable Unit, a computer hardware part that can be replaced in a client environment (e.g. data center) by a trained service representative.

Service Processor: an embedded controller which controls the system power on/off, hardware initialization, and other internal system hardware operations, e.g. CEC Hot Add & Repair Maintenance, CEC runtime diagnostics, service network, etc.

HMC: Hardware Management Console, an appliance with platform management software which manages one or multiple large computers.

SDMC: System Director Management Console, an appliance with platform management software which manages one or more large computers.

RV: Repair and Verify, a serviceability application that runs on the HMC.

TCE: Translation Control Entry, an I/O address translation table in system memory.

DLPAR: Dynamic Logical Partitioning provides the capability to change the assignment of processor, memory, and I/O hardware resources dynamically, without restarting the operating system partition.

CUoD: Capacity Upgrade on Demand.

VIOS: Virtual I/O Server is software that is located in a logical partition. This software facilitates the sharing of physical I/O resources between client logical partitions within the server. Client logical partitions can share SCSI devices, fiber channel adapters, Ethernet adapters, and expand the amount of memory available using the VIOS.

SEA: Shared Ethernet Adapter is a VIOS component that bridges a physical Ethernet adapter and one or more virtual Ethernet adapters. Using SEA, logical partitions on the virtual network can share access to the physical network and communicate with stand-alone servers and logical partitions on other systems.

VLAN: Virtual Local Area Network allows a physical network to be divided administratively into separate logical networks.

LVM: Logical Volume Management provides a method of allocating space on mass-storage devices.

MPIO: Multi-path I/O provides a redundant path to a given logical unit number (LUN) device.

SCSI: Small Computer System Interface is a set of standards for physically connecting computers and peripheral devices and transferring data between them. These standards define commands, protocols, electrical and optical interfaces.

SAN: Storage Area Network is an architecture to attach remote computer data storage devices (such as disk arrays) to servers so the devices appear to be locally attached to the operating system.

VTD: Virtual Target Device

ESS: Enterprise Storage Server is IBM's disk Storage server.

IPL: System power on, initial program load to start the computer system.

reIPL: System restart after failure or restart initiated by the system administrator.

RMC: Resource Monitoring and Control

7.3.2 Service Operation or Procedure

Hot Add: add new CEC hardware to the system while the system is powered on and critical applications are quiesced.

Hot Upgrade: add new CEC hardware to the system or exchanges hardware to increase system capacity while the system is powered on and critical applications are quiesced.

Hot Repair: repair the CEC hardware in the system while the system is powered on and critical applications are quiesced.

CHARM: CEC Hot Add & Repair Maintenance is a common term used to refer to the hot add, upgrade or repair functions of CEC hardware.

Concurrent Add: adds new hardware to the system while the system and all applications are running.

Non-Concurrent Repair: repairs the hardware in the system while the system is powered off.

Disruptive Service or Operation: adds, upgrades or repairs the hardware in the system while the system is powered off.

SSR: IBM System Service Representative

PE: Product Engineer, a highly trained product support person that has direct access to product development team for design information.

The table below summarizes the terminologies used in the Power6 CEC Concurrent Maintenance (CCM) and Power7 CHARM function.

IBM Power Systems

New POWER7 Terminology

POWER6 CEC CM Terminology	New POWER7 Terminology
"Concurrent" when referring to CEC hardware	"Hot" when referring to CEC hardware
CCM: CEC Concurrent Maintenance	CHARM: CEC Hot Add & Repair Maintenance
Concurrent Node Add	Hot Node Add
Concurrent Node Upgrade (memory)	Hot Node Upgrade (memory)
Concurrent Hot Node Repair Concurrent Cold Node Repair	Hot Node Repair
Concurrent GX Adapter Add	Concurrent GX Adapter Add
Concurrent Cold GX Adapter Repair	Hot GX Adapter Repair
Concurrent System Controller Repair	Concurrent System Controller Repair

6 Power your planet.

Table 11: Power 6 CCM and Power7 CHARM Terminologies

IBM Corporation
Systems and Technology Group
Route 100
Somers, New York 10589

Produced in the United States of America
April 2009

The IBM home page on the Internet can be found at: http://www.ibm.com.

The IBM Power Systems home page on the Internet can be found at: http://www.ibm.com/systems/power/

POW03058-USEN-01