InfoScale™ 9.0 Virtualization Guide - AIX

arctera™

Last updated: 2025-09-09

Documentation

Make sure that you have the current version of the documentation. Each document displays the date of the last update on page 2. The latest documentation is available on the Veritas website:
https://sort.veritas.com/documents

Documentation feedback

Your feedback is important to us. Suggest improvements or report errors or omissions to the documentation. Include the document title, document version, chapter title, and section title of the text on which you are reporting. Send feedback to:
infoscaledocs@veritas.com

You can also see documentation information or ask a question on the Veritas community site:
http://www.veritas.com/community/

Veritas Services and Operations Readiness Tools (SORT)

Veritas Services and Operations Readiness Tools (SORT) is a website that provides information and tools to automate and simplify certain time-consuming administrative tasks. Depending on the product, SORT helps you prepare for installations and upgrades, identify risks in your datacenters, and improve operational efficiency. To see what services and tools SORT provides for your product, see the data sheet:
https://sort.veritas.com/data/support/SORT_Data_Sheet.pdf

Technical Support

Technical Support maintains support centers globally. Technical Support's primary role is to respond to specific queries about product features and functionality. The Technical Support group also creates content for our online Knowledge Base. The Technical Support group works collaboratively with the other functional areas within the company to answer your questions in a timely fashion.

Our support offerings include the following:

- A range of support options that give you the flexibility to select the right amount of service for any size organization
- Telephone and/or Web-based support that provides rapid response and up-to-the-minute information
- Upgrade assurance that delivers software upgrades
- Global support purchased on a regional business hours or 24 hours a day, 7 days a week basis
- Premium service offerings that include Account Management Services

For information about our support offerings, you can visit our website at the following URL:

www.arctera.io/support

All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policy.

Legal Notice

Arctera US LLC | www.arctera.io

Contents

Section 1

Overview

Chapter 1

Storage Foundation and High Availability Solutions in AIX PowerVM virtual environments

This chapter includes the following topics:

Overview of the InfoScale Virtualization Guide

This document provides information about InfoScale products support for AIX virtualization technology. It contains:

- High-level conceptual information for InfoScale products and how they function in PowerVM virtual environments.
- Use case chapters with examples of how InfoScale products can improve performance outcomes for common logical partition (LPAR) use cases.

- High level implementation information for setting up InfoScale products in LPAR environments.

About the AIX PowerVM virtualization technology

AIX logical partition or LPAR is a virtual machine in the IBM PowerVM virtualization technology.

Virtualizated technologies such as IBM PowerVM enable:

- Better utilization of IT resources and sharing of I/O resources to lower costs
- Greater flexibility to re-allocate resources to applications as needed
- Simplification of the infrastructure management by making workloads independent of hardware resources

IBM LPAR virtualization components and terminology are identified in Table 1-1:

Table 1-1

IBM LPAR virtualization technology	Description
IBM logical partition (LPAR)	A virtual server with a subset of the physical server's processors, memory, and I/O adapter with its own operating system instance and applications.
Dynamic Logical Partition (DLPAR)	A virtual server with the ability to add or remove processors, network, or storage adapters while the server remains online.
Micro-partition	A virtual server with shared processor pools with support for up to10 micro-partitions per processor core. Depending upon the Power server, you can run up to 254 independent micro-partitions within a single physical Power server. Processor resources can be assigned at a granularity of 1/100th of a core. Also known as shared processor partition.
IBM LPARs with dedicated I/O	The baseline configuration is a traditional AIX deployment with dedicated HBAs and NICs. The deployment may include partitions with virtual CPUs or partitions that support dynamic re-configuration.
IBM LPARs with Virtual I/O Servers	With Virtual I/O Servers LPARs can share physical resources. The VIOS provides virtual SCSI, virtual fibre channel, and virtual networking for sharing. Sharing of resources between LPARs enables more efficient utilization of physical resources and facilitates consolidation.

Table 1-1 *(continued)*

IBM LPAR virtualization technology	Description
POWER Hypervisor	The POWER Hypervisor is responsible for dispatching the logical partition workload across the shared physical processors. The POWER Hypervisor also enforces partition security, and provides inter-partition communication that enables the Virtual I/O Server's virtual SCSI and virtual Ethernet function.
Virtual I/O Server (VIOS)	The Virtual I/O Server facilitates the sharing of physical I/O resources among LPARs within the server. The Virtual I/O Server provides virtual SCSI target, virtual fibre channel, Shared Ethernet Adapter, PowerVM Active Memory Sharing and PowerVM Client Partition Mobility capability to client logical partitions within the physical server.
VIO client	The VIO client is a client LPAR that consumes resources shared by the VIO Server.
Virtual SCSI	Virtual disks vSCSI provided by the VIO server to reduce the need for dedicated physical disk resources for client partitions. vSCSI can be full LUNs or logical volumes.
Virtual Ethernet	In-memory network connections between partitions by POWER Hypervisor that reduce or eliminate the need for separate physical Ethernet Adapters in each LPAR.
Shared Ethernet Adapter	The Shared Ethernet Adapter (SEA) enables network traffic outside the physical server by routing it through a software-based layer 2 switch running in the VIO Server.
N_Port ID Virtualization	Virtual HBAs which enable multiple LPARs/micro-partitions to access SAN devices through shared HBAs providing direct Fibre Channel connections from client partitions to storage. Fibre Channel Host Bus Adapters (HBAs) are owned by VIO Server Partition.
Workload Partitions (WPARs)	Workload Partitions enable administrators to virtualize the AIX operating system, by partitioning an AIX operating system instance into multiple environments. Each environment within the AIX operating system instance is called a workload partition (WPAR). One WPAR can host applications and isolate the applications from applications executing in other WPARs. WPAR is a pure software solution and has no dependencies on hardware features.

Table 1-1 *(continued)*

IBM LPAR virtualization technology	Description
WPAR Manager	The WPAR manager allows an administrator to create, clone, and remove WPAR definitions, or start and stop WPARs. It enables Live Application Mobility which allows relocation of WPARs from one server to another without restarting the application. The WPAR Manager includes a policy engine to automate relocation of WPARs between systems based on system load and other metrics.
Application WPAR	An application Workload Partition (WPAR) is a lightweight partition in which individual applications run. An application WPAR can only run application processes, not system daemons such as `inetd cron`. An application WPAR is a temporary object which is removed when the application is completed.
System WPAR	A system Workload Partition (WPAR) has a private copy of many of the AIX OS parameters. If desired, it can have its own dedicated, completely writable file systems. Most OS daemons can run, and each system WPAR has its own user privilege space.
Live Partition Mobility	Live Partition Mobility enables greater control over the usage of resources in the data center by enabling the migration of a logical partition from one physical system to another. This feature enables the transfer of a configuration from source to destination without disrupting the hosted applications or the setup of the operating system and applications
Live Application Mobility	Live Application Mobility enables the planned migration of workloads from one system to another without interrupting the application and can be used to perform a planned firmware installation on a server.
Active Memory Sharing	Active Memory Sharing is a virtualization technology that enables multiple partitions to share a pool of physical memory. AMS increases system memory utilization and reduces the amount of physical memory that the system requires.
Active Memory Expansion	Active Memory Expansion relies on compression of in-memory data to increase the amount of data that can be placed into memory. This feature expands the effective memory capacity of a POWER7 system. The operating system manages the in-memory data compression, which is transparent to applications and users.

Table 1-1 *(continued)*

IBM LPAR virtualization technology	Description
Hardware Management Console (HMC)	Dedicated hardware/software to configure and administer a partition capable POWER server.
Integrated Virtual Manager	Management console which runs in the VIO for partition management of entry level systems.
Lx86	Supports x86 Linux applications running on POWER.

InfoScale products can be used in LPAR-based virtualization environments to provide advanced storage management, mission-critical clustering, and failover capabilities. This guide illustrates some reference configurations for the use of InfoScale products with IBM Power virtualization. These reference configurations can be customized to fit most implementations. An assumption is made that the reader understands the AIX operating system, including its architecture, as well as how to configure and manage LPARs using the management software already provided by AIX. There is also an expectation that the user is familiar with the basic InfoScale products software and is well versed with its administration and management utilities. Additional details regarding IBM AIX, LPARs, and InfoScale products software are available in the additional documentation section.

The InfoScale products support VIO clients that use memory from the Active Memory Sharing (AMS) pool. Arctera recommends that the ratio of the physical memory in the AMS pool should comply with the AIX guidelines.

Active Memory Expansion is configurable per logical partition (LPAR). Active Memory Expansion can be selectively enabled for one or more LPARs on a system. When Active Memory Expansion is enabled for an LPAR, the operating system compresses a portion of the LPAR's memory and leaves the remaining portion of memory uncompressed. The memory is effectively broken up into two pools – a compressed pool and an uncompressed pool. The operating system dynamically varies the amount of memory that is compressed, based on the workload and the configuration of the LPAR.

See the IBM Redpaper PowerVM Virtualization documents for the AIX guidelines.

About InfoScale products support for the AIX PowerVM environment

IBM PowerVM is a virtualization solution for AIX environments on IBM POWER technology. In the IBM PowerVM environment, multiple logical partitions (LPARs)

can be carved in a physical server. The physical system is also called the managed system. The LPARs can be assigned physical or virtual resources. The Virtual I/O Server is a dedicated partition and is a software appliance with which you can associate physical resources and that allows you to share these resources among multiple client logical partitions. The Virtual I/O Server can use both virtualized storage and network adapters. The managed system, LPARs, and the resources are managed using the Hardware Management Console (HMC) appliance sitting outside the physical frame.

InfoScale products are supported in the following components of the IBM PowerVM environment:

- IBM Virtual I/O Server (VIOS)
- IBM logical partition (LPAR)
- Workload Partition (WPAR)

Table 1-2 Supported IBM PowerVM components

InfoScale component or solution	VIOS	LPAR	WPAR
Dynamic Multi-Pathing (DMP)	Y	Y	Y
Storage Foundation (SF)	N	Y	Y
Cluster Server (VCS)	N	Y	Y
Storage Foundation Cluster File System High Availability (SFCFSHA)	N	Y	Y
Storage Foundation for Oracle RAC (SF Oracle RAC)	N	Y	Y
Replicator (VR)	N	Y	Y

About IBM LPARs with N_Port ID Virtualization (NPIV)

N_Port ID Virtualization or NPIV is a Fibre Channel (FC) industry standard technology that allows multiple N_Port IDs to share a single physical N_Port. NPIV provides the capability to take a single physical Fibre Channel HBA port and divide it such that it appears, to both the host and to the SAN, as though there are multiple World Wide Port Names (WWPNs).

NPIV provides direct access to the Fibre Channel adapters from multiple virtual machine (client partitions), simplifying zoning and storage allocation. Resources can be zoned directly to the individual virtual Fibre Channel client ports, each having its own World Wide Port Name (WWPN).

The use of NPIV with IBM VIO provides the capability to use a single Fibre Channel port and overlay multiple WWPNs so that it appears to the SAN as both the VIO server and client partitions have their dedicated Fibre Channel ports. NPIV enables the AIX VIO server to provision entire dedicated logical ports to client LPARs rather than individual LUNs. Client partitions with this type of logical port operates as though the partition has its own dedicated FC protocol adapter. To utilize the NPIV functionality, a new type of virtual Fibre Channel (VFC) adapter is defined on both the VIO and Client. A server VFC adapter can only be created on a VIO server partition; a client VFC adapter can only be created on client partitions. WWPNs are allocated to client VFC adapters when they are defined in the profile, based upon an assignment pool generated from the backing physical adapter.

There is always corresponding one-to-one mapping relationship between VFC adapters on client logical partitions and VFC on the VIOS. That is, each VFC that is assigned to a client logical partition must connect to only one VFC adapter on VIOS, and each VFC on VIOS must connect to only one VFC on the client logical partition.

Characteristics of a LUN through NPIV

- To the operating system, multi-pathing drivers and system tools, a LUN presented through NPIV has all the characteristics of a LUN presented through a dedicated HBA. Device inquiry and probing works as with physical HBAs. When a VFC interface is created, two World Wide Port Names (WWPNs) are assigned. This information is available in the HMC as part of the virtual HBA properties.
- All SCSI device inquiry operations work, allowing for array identification functions, visibility of LUN Device Identifiers, and discovery of such attributes as thin and thin re-claim capability. SCSI-3 persistent reservation functionality is also supported, enabling the use of SCSI-3 I/O Fencing if the underlying storage supports.
- When Zoning/LUN mapping operations occur, care should be taken to ensure that storage is assigned to both WWPNs. During normal operation, only one of the WWPN identifiers is in use, but during a LPAR live migration event, the WWPN identifier not previously used will be configured on the appropriate backing HBA on the target system, log into the SAN, and then become the active WWPN. The previously used WWPN will become inactive until the next Live Partition Mobility operation.

For SFHA Solutions support of NPIV:

See "Support for N_Port ID Virtualization (NPIV) in IBM Virtual I/O Server (VIOS) environments" on page 31.

For NPIV requirements:

See "About setting up logical partitions (LPARs) with InfoScale products" on page 32.

About Veritas Extension for Oracle Disk Manager

The Veritas Extension for Oracle Disk Manager (ODM) enhances file management and disk I/O throughput. The features of ODM are best suited for the Oracle databases that reside in a Veritas File System (VxFS). ODM allows users to improve database throughput for I/O intensive workloads with special I/O optimization.

The Veritas Extension for ODM must be installed and configured in the global environment of AIX. You can administer the Veritas Extension for ODM only from the global environment.

See "Enabling Veritas Extension for ODM file access from WPAR with VxFS" on page 48.

Virtualization use cases addressed by InfoScale

This section lists use cases where InfoScale components can improve the PowerVM environment:

Table 1-3 Virtualization use cases addressed by InfoScale in a PowerVM environment

Virtualization use case	Arctera solution	Use case implementation details
Simplified management	Dynamic Multi-Pathing (DMP) in the host	Configuring a Virtual I/O server for simplified management: how to manage virtual machines using the same command set, storage namespace, and environment as in a non-virtual environment: See "About simplified management" on page 56.
Application management and availability	Cluster Server (VCS) in the LPAR	How to manage application availability.
Virtual machine management and availability	VCS in the LPAR	How to manage virtual machine high availability (LPAR failover and migration). See "About virtual machine (logical partition) availability" on page 84.
Simplified management and high availability for IBM Workload partitions	VCS in the LPAR	Setting up Workload Partitions (WPARs) for operating system virtualization to reduce operating system images, and simplify management. See "About using IBM Workload Partitions (WPARs) with InfoScale products" on page 96.

Section 2

Implementation

Chapter 2

Setting up Storage Foundation and High Availability Solutions in AIX PowerVM virtual environments

This chapter includes the following topics:

Supported configurations for Virtual I/O servers (VIOS) on AIX

InfoScale products support various configurations in the VIOS-based virtual environment and are certified on AIX.

InfoScale products provide the following functionality for VIO:

- Storage visibility
- Storage management
- Replication support
- High availability
- Disaster recovery

The configurations profiled in the table below are the minimum required to achieve the storage and availability objectives listed. You can mix and match the use of SFHA Solutions products as needed to achieve the desired level of storage visibility, management, replication support, availability, and cluster failover for the Virtual I/O server (VIOS) and logical partitions (LPARS).

Table 2-1 InfoScale features in a VIO environment

Objective	Recommended InfoScale products configuration
Storage visibility for LPARs	Dynamic Multi-Pathing (DMP) in the LPAR See " Dynamic Multi-Pathing in the logical partition (LPAR)" on page 21.
Storage visibility for the VIOS	DMP in the VIOS See " Dynamic Multi-Pathing in the Virtual I/O server (VIOS)" on page 22.
Storage management features and replication support for LPARs	InfoScale products in the LPARs See "InfoScale products in the logical partition (LPAR)" on page 23.
Advanced storage management features and replication support for LPAR	Storage Foundation Cluster File System High Availability (SFCFSHA) in the LPARs See " Storage Foundation Cluster File System High Availability in the logical partition (LPAR)" on page 24.

Table 2-1 InfoScale features in a VIO environment *(continued)*

Objective	Recommended InfoScale products configuration
End-to-end storage visibility in the VIOS and LPARs	DMP in the VIOS and LPARs See " Dynamic Multi-Pathing in the Virtual I/O server (VIOS) and logical partition (LPAR)" on page 25.
Storage management features and replication support in the LPARs and storage visibility in in the VIOS	DMP in the VIOS and SF in the LPARs See " Dynamic Multi-Pathing in the Virtual I/O server (VIOS) and InfoScale products in the logical partition (LPAR)" on page 26.
Virtual machine monitoring and availability for managed LPARs	Cluster Server (VCS) in the management LPARs See " Cluster Server in the management LPAR" on page 28.
Application failover for LPARs	VCS in the LPARs See " Cluster Server in the logical partition (LPAR)" on page 28.
Application failover across LPARs and physical hosts	VCS in a cluster across LPARs and AIX physical host machines See " Cluster Server in a cluster across logical partitions (LPARs) and physical machines" on page 30.

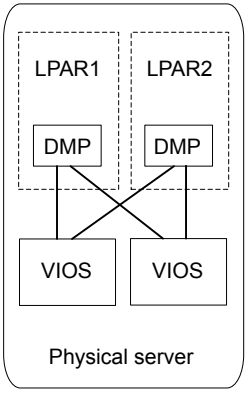
Each configuration has specific advantages and limitations.

Dynamic Multi-Pathing in the logical partition (LPAR)

Dynamic Multi-Pathing (DMP) can provide storage visibility in LPARs. DMP in the LPAR provides:

- Multi-pathing functionality for the operating system devices configured in the LPAR
- DMP metadevices (also known as DMP nodes) to represent all the device paths to the same physical LUN
- Support for enclosure-based naming
- Support for standard array types

Figure 2-1 Dynamic Multi-Pathing in the LPAR



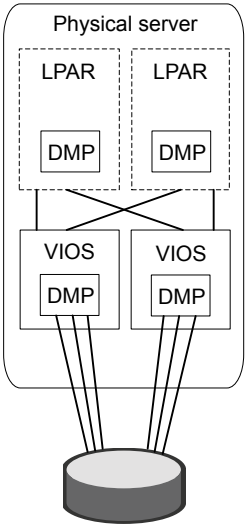
For more information on DMP features, see the *Dynamic Multi-Pathing Administrator's Guide*.

Dynamic Multi-Pathing in the Virtual I/O server (VIOS)

Dynamic Multi-Pathing (DMP) can provide storage visibility in the VIOS. Using DMP in the VIOS enables:

- Centralized multi-pathing functionality
- Enables active/passive array high performance failover
- Centralized storage path management
- Fast proactive failover
- Event notification

Figure 2-2 Dynamic Multi-Pathing in the VIOS



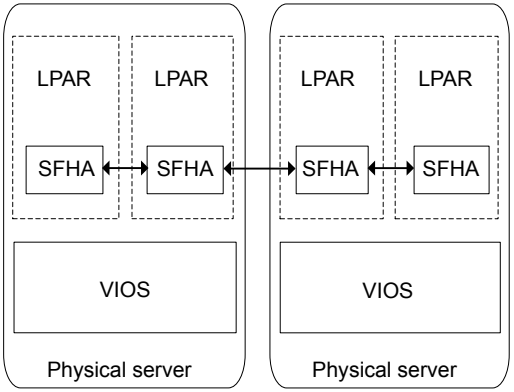
For more information on DMP features, see the *Dynamic Multi-Pathing Administrator's Guide*.

InfoScale products in the logical partition (LPAR)

InfoScale products in the LPAR provide storage management functionality for the LPAR resources. InfoScale products enable you to manage LPAR storage resources more easily by providing:

- Enhanced database performance
- SmartIO for caching data on Solid State Drives (SSDs)
- Point-in-time copy features for data back-up, recovery, and processing
- Options for setting policies to optimize storage
- Methods for migrating data easily and reliably
- Replication support
- High availability for applications and LPARs Disaster recovery for applications

Figure 2-3 Storage Foundation and High Availability in the LPAR



For more information on Storage Foundation features, refer to the *Storage Foundation Administrator's Guide*.

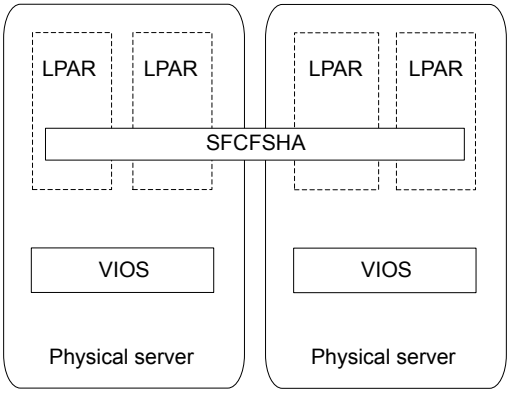
For more information on SmartIO, refer to the *InfoScale SmartIO for Solid-State Drives Solutions Guide*.

Storage Foundation Cluster File System High Availability in the logical partition (LPAR)

Storage Foundation Cluster File System High Availability (SFCFSHA) provides advanced storage management functionality for the LPAR. SFCFSHA enables you to manage your LPAR storage resources more easily by providing:

- Enhanced database performance
- SmartIO for caching data on Solid State Drives (SSDs)
- Point-in-time copy features for data back-up, recovery, and processing
- Options for setting policies to optimize storage
- Methods for migrating data easily and reliably
- Replication support
- High availability for applications
- LPARs Disaster recovery for applications
- Faster recovery of applications

Figure 2-4 Storage Foundation Cluster File System High Availability in the LPAR



For more information on Storage Foundation features, see the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

For more information on SmartIO, see the *InfoScale SmartIO for Solid-State Drives Solutions Guide*.

Dynamic Multi-Pathing in the Virtual I/O server (VIOS) and logical partition (LPAR)

Dynamic Multi-Pathing (DMP) can provide end-to-end storage visibility across both the VIOS and LPAR.

Using DMP in the VIOS enables:

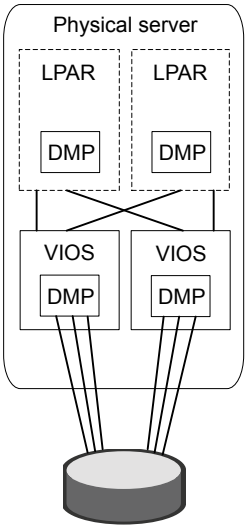
- Centralized multi-pathing functionality
- Enables active/passive array high performance failover
- Centralized storage path management
- Fast proactive failover
- Event notification

Using DMP in the LPAR provides:

- Multi-pathing functionality for the operating system devices configured in the LPAR
- DMP metadevices (also known as DMP nodes) to represent all the device paths to the same physical LUN
- Support for enclosure-based naming

- Support for standard array types

Figure 2-5 Dynamic Multi-Pathing in the VIOS and the LPAR



For more information on DMP features, see the *Dynamic Multi-Pathing Administrator's Guide*.

Dynamic Multi-Pathing in the Virtual I/O server (VIOS) and InfoScale products in the logical partition (LPAR)

Using Dynamic Multi-Pathing (DMP) in the VIOS in combination with Storage Foundation and High Availability (SFHA) in the LPAR provides storage management functionality for LPAR resources and storage visibility in the VIOS.

Using DMP in the VIOS provides:

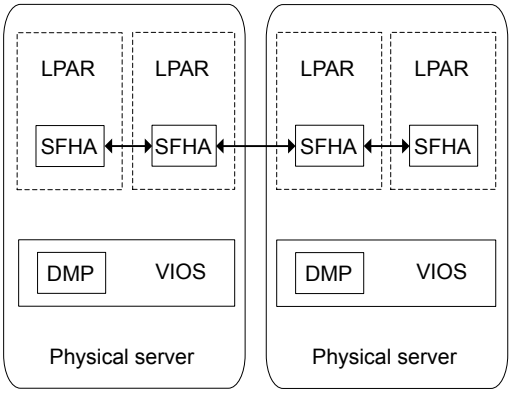
- Centralized multi-pathing functionality
- Active/passive array high performance failover
- Centralized storage path management
- Fast proactive failover
- Event notification

Using SFHA in the LPAR provides:

- Enhanced database performance
- SmartIO for caching data on Solid State Drives (SSDs)

- Point-in-time copy features for data back-up, recovery, and processing
- Options for setting policies to optimize storage
- Methods for migrating data easily and reliably
- Replication support
- High availability and disaster recovery for the applications

Figure 2-6 DMP in the VIOS and SFHA in the LPAR



The following clustering configurations are supported:

- LPAR to LPAR clustering and failover for application availability.
  VCS runs in the LPARs forming a cluster, and provides high availability of the applications running in the LPAR.
  See " Cluster Server in the logical partition (LPAR)" on page 28.
- LPAR to LPAR Clustering and Fail-over for LPAR (virtual machine) availability.
  VCS runs in one LPAR on each physical server, known as the management LPAR. The management LPAR behaves as a control point, and provides high-availability to the other LPARs running on the same physical server. The management LPAR views the LPARs that it manages as virtual machines but does not have visibility into the applications on the managed LPARs.
  You can create a cluster of the management LPARs on more than one physical server to provide failover of the managed LPARs on the different physical servers.
  See "VCS in the management LPAR" on page 84.

For more information on SFHA features, see the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

For more information on SmartIO, see the *InfoScale SmartIO for Solid-State Drives Solutions Guide*.

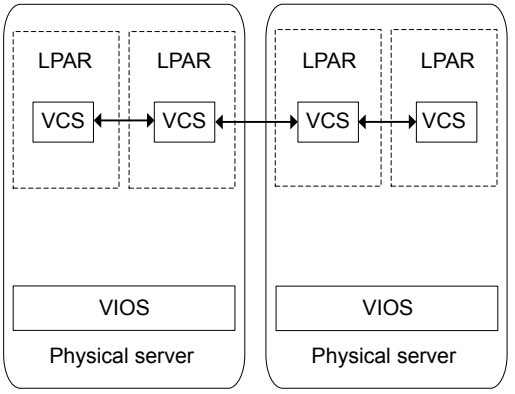
For more information on DMP features, see the *Dynamic Multi-Pathing Administrator's Guide*.

Cluster Server in the logical partition (LPAR)

Cluster Server (VCS) provides the following functionality for LPARs:

- Enables nodes to cooperate at the software level to form a cluster
- Enables other nodes to take predefined actions when a monitored application fails, for instance to take over and bring up applications elsewhere in the cluster

Figure 2-7 Cluster Server in the LPAR



For more information on Cluster Server features, see the *Cluster Server Administrator's Guide*.

Cluster Server in the management LPAR

The logical partition (LPAR) which is a Cluster Server (VCS) node and controls other LPARs on the physical server is referred to as a management LPAR (MPLAR).

VCS monitors and manages LPARs using LPAR resources.

VCS enables the following for management LPARs:

- Connects multiple LPARs to form a cluster for increased availability
- Redundant Hardware Management Console (HMC) support
- Multiple VIOS support
- Enables nodes to cooperate at the software level to form a cluster

- Enables other nodes to take predefined actions when a monitored application fails, for instance to take over and bring up applications elsewhere in the cluster
- Enables monitoring and failover of LPARs configured in VCS as LPAR resources

Figure 2-8 VCS in the management LPAR

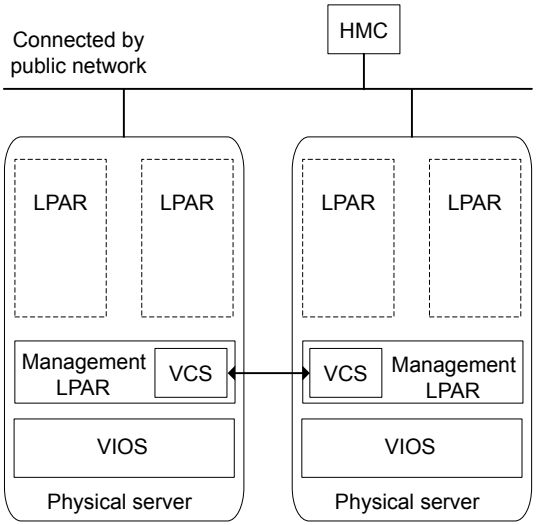
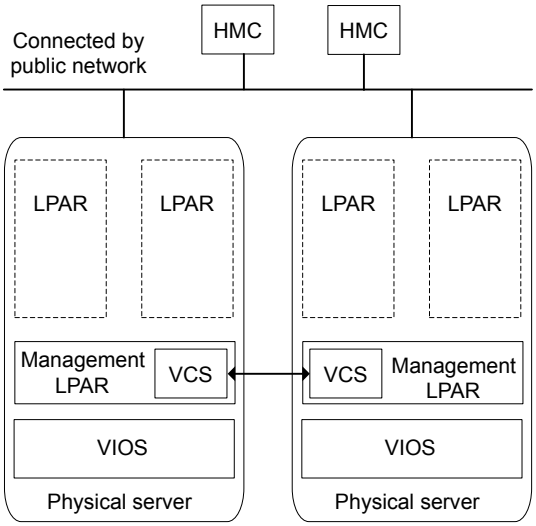
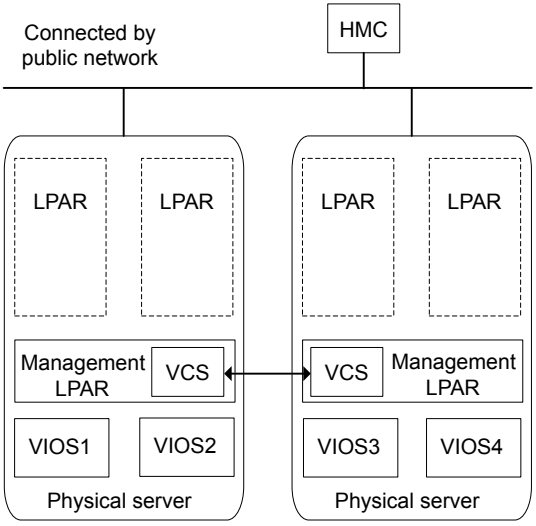


Figure 2-9 VCS in the management LPAR with redundant HMCs



The VCS LPAR agent supports redundant HMC configurations. VCS can use any HMC which is up and running to manage and monitor the LPARs.

Figure 2-10 VCS in the management LPAR with multiple VIOS



Multiple VIOS support provides high availability to LPARs in case of VIO server(s) crash: LPAR agent provides high availability against VIO server(s) crash for the managed LPARs. If all the VIO servers specified are down, managed LPARs are failed over to another host.

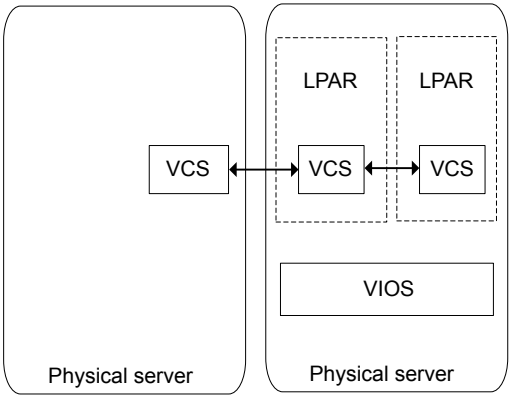
For more information on Cluster Server features, see the *Cluster Server Administrator's Guide*.

Cluster Server in a cluster across logical partitions (LPARs) and physical machines

Using Cluster Server (VCS) in both guests and hosts enables an integrated solution for resource management across virtual machines (VMs) and physical hosts. You can create a physical to virtual cluster combining VCS in the LPAR together with VCS running in the management LPAR on another physical host, enabling VCS to:

- Monitor applications running within the LPAR
- Fail the applications over to another physical host
- Failover an application running on a physical host to a LPAR

Figure 2-11 Cluster Server in a cluster across LPARs and physical machines



For more information on Storage Foundation features, see the *Cluster Server Administrator's Guide*.

Support for N_Port ID Virtualization (NPIV) in IBM Virtual I/O Server (VIOS) environments

SFHA Solutions support N_Port ID Virtualization (NPIV) in IBM Virtual I/O Server (VIOS) environments:

- The VIOS is configured with NPIV capable Fibre Channel (FC) adapters that are connected to a SAN switch that is NPIV capable.
- The LUNs mapped to the VIO client behave like an LPAR with a dedicated FC adapter.
- The devices in the VIO client appear as regular SCSI disks, which Storage Foundation can access. Unlike in the classic VIO environment without NPIV, Storage Foundation treats these devices as if they came from a regular SAN storage array.
- With NPIV, the VIO client environment is transparent to Storage Foundation. All of the Storage Foundation commands would have the same output as in a regular physical AIX server.
- You can create and import disk groups on NPIV devices, which provide access to volumes and file systems.
- Arctera has qualified NPIV support with Storage Foundation.
- Arctera has also qualified migration of storage used by Storage Foundation from the AIX physical server environment to the IBM VIO environment.

See "Migrating from Physical to VIO environment" on page 141.

Table 2-2 InfoScale products NPIV support

InfoScale components	NPIV support
Storage Foundation	Storage Foundation 9.0 supports all functionality available with dedicated HBAs when using LUNs presented through NPIV. All IBM supported NPIV enabled HBAs are supported by Storage Foundation.
Storage Foundation Cluster File System High Availability (SFCFS HA)	SFCFSHA is supported with NPIV.
Cluster Server (VCS)	VCS supports NPIV. With NPIV, the VIOS client environment is transparent to VCS and the LUNs are treated as regular SAN storage array LUNs. Since SCSI3 persistent reserve is available, I/O fencing is also supported.

About setting up logical partitions (LPARs) with InfoScale products

Before setting up your virtual environment, verify that your planned configuration will meet the system requirements, licensing and other considerations for installation with Storage Foundation and High Availability (SFHA) Solutions products.

- Licensing: Storage Foundation or Storage Foundation Cluster File System High Availability in an LPAR may be licensed either by licensing the entire server (for an unlimited number of LPARs) or licensing the maximum number of processors cores assigned to that LPAR.
- IBM Power virtualization requirements: See IBM documentation.
- InfoScale product requirements: See Table 2-4
- *Release Notes*: Each InfoScale product Release Notes contains last minute news and important details for each product, including updates to system requirements and supported software. Review the Release Notes for the latest information before you start installing the product.
  The product documentation is available on the Web at the following location:
  https://sort.veritas.com/documents

Table 2-3 IBM Power Virtualization system requirements

IBM Power Virtualization requirement	Description
Supported architecture	Power PC
Minimum system requirements	No specific requirements. See the Release Notes for your product.
Recommended system requirements	■ 6GB plus the required disk space recommended by the guest operating system per guest. For most operating systems more than 6GB of disk space is recommended ■ One processor core or hyper-thread for each virtualized CPU and one for the host ■ 2GB of RAM plus additional RAM for LPARs
IBM documentation for more information	See the IBM documentation.

Table 2-4 InfoScale product requirements

InfoScale product requirement	Description
Hardware	Full virtualization-enabled CPU Refer to the hardware compatibility list (HCL) applicable to your InfoScale version and platform.
Software	■ Dynamic Multi-pathing 9.0 Used for storage visibility on logical partitions (LPARs) and VIOS ■ Storage Foundation 9.0 Used for storage management on LPARs and VIOS ■ Storage Foundation HA 9.0 Storage Foundation and High Availability 9.0 Used for storage management and clustering on LPARs and VIOS ■ Cluster Server 9.0 Used for applications and managed LPARs monitoring and failover
Supported OS version in LPAR	Refer to the *InfoScale Release Notes*.
Supported OS management LPAR	Refer to the *InfoScale Release Notes*.

Table 2-4 InfoScale product requirements *(continued)*

InfoScale product requirement	Description
Storage	■ Shared storage for holding the LPAR image. (virtual machine failover) ■ Shared storage for holding the application data. (application failover)
Networking	■ Configure the LPAR for communication over the public network ■ Setup virtual interfaces for private communication.
Documentation: see the product release notes for the most current system requirements, limitations, and known issues:	■ *InfoScale Release Notes* ■ *InfoScale Operations Manager Release Notes* ■ Services and Operations Readiness Tools https://sort.veritas.com/documents
Installation, patching, and configuration requirements:	Arctera strongly recommends that you use InfoScale products with the latest patches. No other configuration is required. Refer to the following website for the latest patches for Storage Foundation 9.0 on AIX: https://sort.veritas.com/checklist/install/

To use VCS to manage LPARs as virtual machines, the following requirements must be met.

Table 2-5 VCS system requirements for the LPAR-supported configurations

VCS requirement	Description
VCS version	7.3.1 or later
Supported OS version in LPARs	AIX 7.2 TL4, TL5 AIX 7.3 TL0, TL1
Supported VIOS version	2.1.3.10-FP-23 and above
Supported HMC version	7.2.0.0, 8.1, 9.1 **Best practice**: all the physical servers that are part of a cluster are managed by the same HMC.
Supported hardware	Power 7, 8, 9, 10

Table 2-6 N_Port ID Virtualization (NPIV) requirements

NPIV requirement	Description
NPIV support	Included with PowerVM Express, Standard, and Enterprise Edition and supports AIX 7.2 and AIX 7.3.
VIO requirements	NPIV requires a minimum of Power6 systems, VIOS 2.1, and 8GB HBA adapters. NPIV also requires NPIV aware switches. The end storage devices need not be NPIV aware.
Hardware requirements	NPIV requires extended functionality on the HBA. Currently IBM sells this as an 8GB HBA, part number XXXXX. The SAN Switch ports must also support NPIV as well, Brocade and Cisco make products that provide this functionality.
Information for NPIV and how to configure an IBM VIO environment	See IBM documentation.

Installation, patching, and configuration requirements for NPIV support:

- No patches are needed at the time of release. No other configuration is required.
- Using 9.0 products with the latest patches when they become available is strongly recommended.
- For current information on patches, see: https://sort.veritas.com/checklist/install/

You can use Cluster Server (VCS) in a IBM PowerVM virtualization environment to provide mission-critical clustering and failover capabilities.

See "About IBM LPARs with N_Port ID Virtualization (NPIV)" on page 15.

See "Boot device management on NPIV presented devices" on page 82.

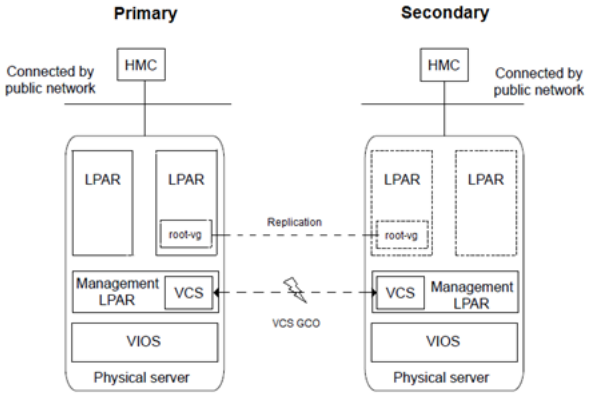
See "Additional documentation for AIX virtualization" on page 149.

Configuring IBM PowerVM LPAR guest for disaster recovery

The IBM PowerVM is configured for disaster recovery by replicating the boot disk by using the replication methods like Hitachi TrueCopy, EMC SRDF, IBM duplicating, cloning rootvg technology, and so on. The network configuration for

the LPAR on the primary site may not be effective on the secondary site, if the two sites are in the different IP subnets. To apply the different network configurations on the different sites, you will need to make additional configuration changes to the LPAR resource.

Figure 2-12 Across site disaster recovery of the managed LPAR using VCS in the management LPARs



To configure LPAR for disaster recovery, you need to configure VCS on both the sites in the management LPARs with the GCO option. See the *Cluster Server Administrator's Guide* for more information about the global clusters.

Perform the following steps to set up the LPAR guest (managed LPAR) for disaster recovery:

1 On the primary and the secondary site, create the PowerVM LPAR guest using the Hardware Management Console (HMC) with the ethernet and the client Fibre Channel (FC) virtual adapter's configuration.

Note: The installed OS in the LPAR guest is replicated using the IBM rootvg cloning technology or the DR strategy N_Port ID Virtualization (NPIV).

---

2 On the LPAR guest, copy and install the `VRTSvcsnr` fileset from the VCS installation media. This fileset installs the `vcs-reconfig` service in the LPAR guest. This service ensures that the site-specific-network parameters are applied when the LPAR boots. You can install the `VRTSvcsnr` fileset by performing the following steps:

```
# mkdir /<temp_dir>
# cp <media>/pkgs/VRTSvcsnr.bff /<tmp_dir>
# cd /<temp_dir>
# installp -a -d VRTSvcsnr.bff VRTSvcsnr
```

3 Create a VCS service group and add a VCS LPAR resource for the LPAR guest. Configure the DROpts attribute of the LPAR resource with the site-specific values for each of the following: IPAddress, Netmask, Gateway, DNSServers (nameserver), DNSSearchPath, Device, Domain, and HostName.

Set the value of the ConfigureNetwork attribute to 1 from the DROpts attribute to make the changes effective. The LPAR agent does not apply to the DROpts attributes for the guest LPAR, if the value of the ConfigureNetwork attribute is 0. For more info about DROpts attribute see the *Cluster Server Bundled Agents Reference Guide.*

4 [This step is optional:] To perform the rootvg replication using NPIV, the boot disk LUN is mapped directly to the guest LPARs via NPIV, and the source production rootvg LUN is replicated using the hardware technologies like Hitachi TrueCopy, EMC SRDF, and so on for the DR Site. Subsequently, add the appropriate VCS replication resource to the LPAR DR service group. Examples of hardware replication agents are SRDF for EMC SRDF, HTC for Hitachi TrueCopy, MirrorView for EMC MirrorView, and so on. VCS LPAR resource depends on the replication resource.

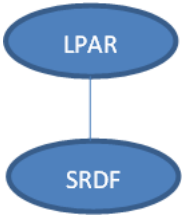
See Figure 2-13

For more information about the appropriate VCS replication agent that is used to configure the replication resource, you can visit our website at the following URL: https://sort.veritas.com/agents

The replication resource ensures that when the resource is online in a site, the underlying replicated devices are in the primary mode, and the remote devices are in the secondary mode. Thus, when the LPAR resource is online, the underlying storage is always in the read-write mode.

5 Repeat step 1 through step 4 on the secondary site.

Figure 2-13 Sample resource dependency diagram for NPIV base rootvg replication using the hardware replication technology



When the LPAR is online, the LPAR agent creates a private VLAN (with VLAN ID 123) between the management LPAR and the managed LPAR. The VLAN is used to pass the network parameters specified in the DROpts attribute to the managed LPAR. When the managed LPAR boots, it starts the `vcs-reconfig` service that requests for the network configuration from the management LPAR. As a result, the network configuration is resent, as a part of the response through the same VLAN. The `vcs-reconfig` service subsequently applies this configuration when the appropriate commands are run.

Installing and configuring Storage Foundation and High Availability (SFHA) Solutions in the logical partition (LPAR)

To set up an LPAR environment with SFHA solutions after installing AIX:

- Install the Storage Foundation and High Availability (SFHA) Solutions product on the required LPARs.
- Configure the SFHA Solutions product on the required LPARs.
- For SFHA Solutions product installation and configuration information:
  - *Storage Foundation Configuration and Upgrade Guide*
  - *InfoScale Installation Guide*
  - *Cluster Server Configuration and Upgrade Guide*
  - See "Additional documentation for AIX virtualization" on page 149.

The steps above apply for the following configurations:

Figure 2-14 Dynamic Multi-pathing in the LPAR

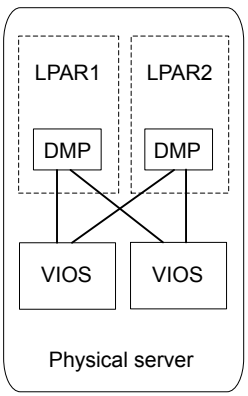


Figure 2-15 Storage Foundation in the LPAR

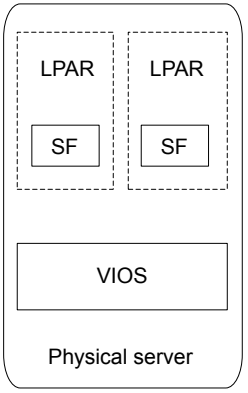


Figure 2-16 Storage Foundation High Availability in the LPAR

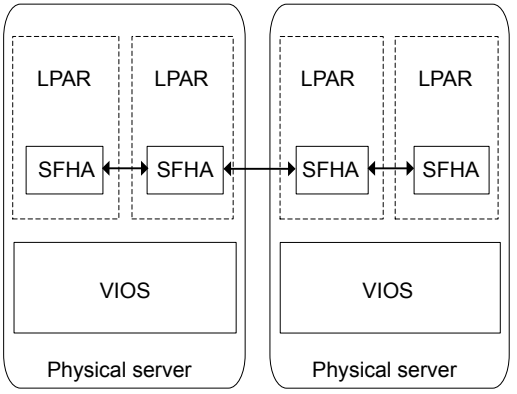
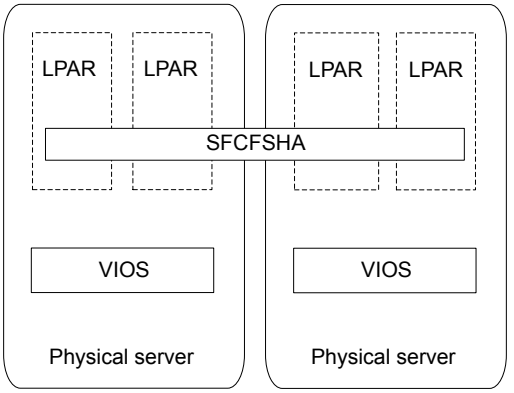


Figure 2-17 Storage Foundation for Cluster File System in the LPAR



Impact of over-provisioning on Storage Foundation and High Availability

Over-provisioning occurs when you select a high compression ratio. Select the compression ratio carefully, keeping in mind the minimum application requirement and the maximum load threshold.

Note that operating system commands present a different view of memory if Active Memory Expansion (AME) is enabled. The `prtconf -m` command displays actual physical memory. The `topas` command displays the relay memory considering AME.

About SmartIO in the AIX virtualized environment

When you install InfoScale products in the logical partition (LPAR), you can use SmartIO to cache data onto an SSD or any other supported fast device. The SSD used for the cache can be either a PCIe or SAS device, or an array-based SSD. The supported caching always happens at the LPAR level. In case the applications' I/O workload is running inside the WPARs, the I/O caching still works at the LPAR layer. SmartIO caching is not supported within the WPAR.

Storage Foundation for Oracle RAC is not supported in the WPAR.

If an array-based SSD is used, Live Partition Mobility (LPM) is supported with SmartIO caching. With direct attached devices (PCIe or SAS), LPM is not supported if SmartIO caching is enabled. If you need to perform LPM, you can use manual steps.

See "Performing LPM in the SmartIO environment" on page 42.

Table 2-7 shows how SmartIO can be used in the virtualized environment.

Table 2-7 AIX: SmartIO support in AIX virtual environments

Configuration in LPAR:	Configuration in VIOS:	Caching takes place:	VxVM read caching	VxFS read caching	VxFS writeback caching
SF	DMP (optional)	On LPAR	Yes	Yes	Yes
SFHA	DMP (optional)	On LPAR	Yes	Yes	Yes
SFCFS	DMP (optional)	On LPAR	Yes	Yes	Yes
SFRAC	DMP (optional)	On LPAR	Yes	Yes	No

For more information about using SmartIO, see the *InfoScale SmartIO for Solid-State Drives Solutions Guide*.

Performing LPM in the SmartIO environment

If an array-based SSD is used, Live Partition Mobility (LPM) is supported with SmartIO caching. With direct attached devices (PCIe or SAS), LPM is not supported if SmartIO caching is enabled. If you need to perform LPM, you can use manual steps.

To perform LPM in the SmartIO environment

1 To prepare the LPAR for the lLPM, perform the following steps:

- Offline the cache area that is created inside the LPAR.

```
Ldom1:/root# sfcache offline cachearea_name
```

- Delete the cache area.

```
Ldom1:/root# sfcache delete cachearea_name
```

2 Remove the SSD device from the VxVM configuration so that the device can be unexported from the LPAR.

```
Ldom1:/root# vxdisk rm ssd_device_name
```

3 Remove the SSD device from the operating system.

```
Ldom1:/root# rmdev -dl os_ssd_device_name -R

Ldom1:/root#  rmdev -dl flash_adapter_name
```

4 Using the IBM HMC, remove the Flash adapter from the partition (leaving only storage accessed by virtualized NPIV adapters). Use the below command from HMC:

```
hmc:~> chhwres  -r io -m lpar_name -o r -p client1 -l  21010121
```

5 Move the partition using Live Partition Mobility.

6 On the new server, attach a Flash adapter to the partition:

```
Lpar1:/root# cfgmgr
```

7 Discover the Flash adapter in VxVM:

```
Lpar1:/root# vxdisk scandisks
```

8 After the local PCIe device is available to the VxVM configuration, you can create the required SmartIO cache area.

9 To live migrate back the LPAR from target frame to source frame, follow step 1 to step 8.

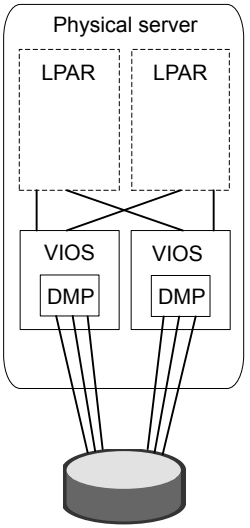
Installing and configuring storage solutions in the Virtual I/O server (VIOS)

To set up a VIOS with DMP after installing VIOS:

- Install Dynamic Multi-Pathing (DMP) on the required VIOS.
- Configure the DMP on the required VIOS.
- For DMP installation and configuration information:
  - *Storage Foundation Configuration and Upgrade Guide*
  - See "Additional documentation for AIX virtualization" on page 149.

The steps above apply for the following configuration:

Figure 2-18 Dynamic Multi-pathing in the VIOS



Installing and configuring Cluster Server for logical partition and application availability

To set up a logical partition (LPAR) environment with Cluster Server (VCS):

- Install VCS.
- Configure VCS.
- No additional VCS configuration is required to make it work inside the LPAR with or without VIOS.
- For installation and configuration information:
  Cluster Server Configuration and Upgrade Guide
  See "Additional documentation for AIX virtualization" on page 149.

The steps above apply for the following configurations:

- VCS in the LPAR
- VCS in the management LPAR

  Note: You must use VCS 6.0 or later in the management LPAR

- VCS in the management LPAR with redundant HMCs
- VCS in the management LPAR with multiple VIOS
- VCS in the management LPAR in the LPAR
- VCS in a cluster across LPARs

Figure 2-19 VCS in the LPAR

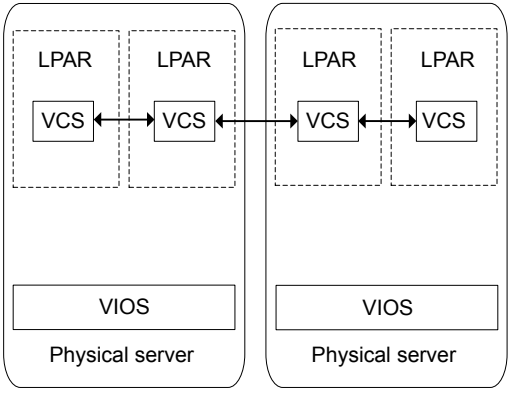


Figure 2-20 VCS in the management LPAR

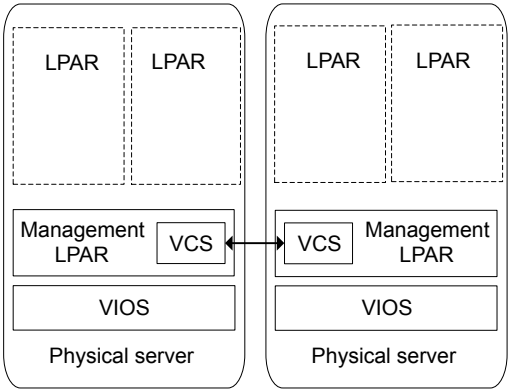


Figure 2-21 VCS in the management LPAR with redundant HMCs

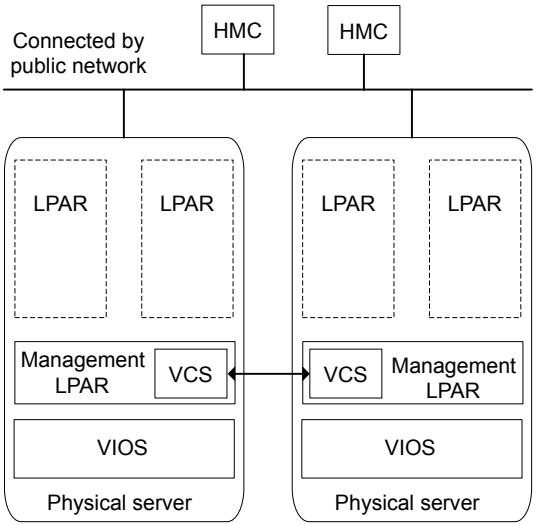


Figure 2-22 VCS in the management LPAR in the LPAR

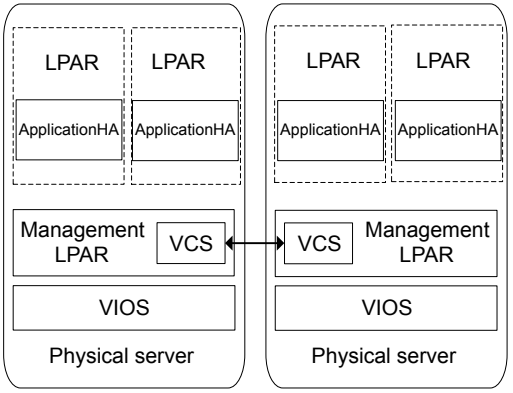
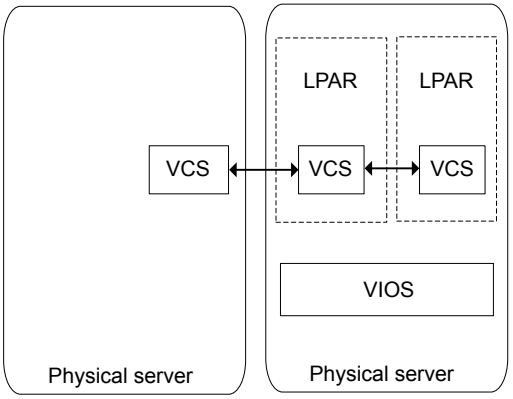


Figure 2-23 VCS in a cluster across LPARs



Recommendations for improved resiliency of InfoScale clusters in virtualized environments

Arctera recommends that you configure the following settings to improve the resiliency of InfoScale cluster configurations in virtualized environments:

- **Peerinact**: Set the default LLT tunable parameter peerinact to 32 seconds instead of 16 seconds. Doing so helps improve the stability of the cluster in virtualized environments, where multiple external factors as described further in this list, can affect the stability of the cluster.
- **Provisioning ratio**: The CPU and memory provisioning ratio affects the stability of the InfoScale cluster. To ensure maximum stability, set the ratio to the lowest value possible. For critical solutions that require maximum resiliency, the ratio must be set to **1:1**.

- **CPU load on host operating systems**: Although the provisioning ratio is low, the CPU load on the host operating systems still plays a part in cluster stability. If the load on the host operating system is very high, it can affect how vCPUs on the guest VMs are scheduled, because vCPUs are processes from the perspective of the host servers.
- **CPU requirement of the actual workload on guests**: When the total CPU requirement for workloads exceeds the available physical CPU capacity, it causes node evictions due to heartbeat timeouts.
- **External events**: External events like live migration of the guest VMs, virtualized disk backups, and so on, are known to add CPU load on the host servers. To reduce this additional load on the CPU, watch the stun duration in your environment caused by these events, and increase the peerinact value, if required. Increase the peerinact value only in these conditions and not in any other circumstances.
- **Hypervisor**: Always follow the best practices for the hypervisor.

How Cluster Server (VCS) manages logical partitions (LPARs)

High-level overview of how VCS manages LPARs.

- Management LPARs form a cluster.
  For information about installing VCS, see the *Cluster Server Configuration and Upgrade Guide*.
- CPU and memory resources are made available to create LPARs on all nodes in the cluster.
- VCS is installed on all the management LPARs to manage the LPARs.
- The operating system is installed on the LPAR on any one host.
- The LPAR is configured as an LPAR resource in VCS.

For detailed instructions on creating and configuring a PowerVM guest, see the IBM documentation.

To configure an LPAR for across physical servers, the following conditions apply:

- You must configure an LPAR on one node with the operating system installed on a shared storage accessible to all the VCS cluster nodes.
- Ensure that the image file resides on the shared storage so that the LPARs can fail over across cluster nodes.
- You can configure the first LPAR using the standard installation procedure:
  See "Setting up management LPAR" on page 86.

Bundled agents are included with VCS for managing many applications. The LPAR agent is included and can be used to manage and provide high availability for LPARs. For information on LPAR agent attributes, resource dependency and agent function, refer to the *Cluster Server Bundled Agents Reference Guide*.

Enabling Veritas Extension for ODM file access from WPAR with VxFS

Veritas Extension for ODM can be enabled from WPAR only when Storage Foundation (SF) is installed in the global environment. For information on installing SF, see the *InfoScale Installation Guide*.

To enable ODM file access from WPAR with VxFS

1 (Optional.) Create a system WPAR.

For example:

```
# mkwpar -n orawpar
```

For other options while creating WPARs, refer to the IBM Redbook for WPAR.

2 Start the system WPAR.

3 Create VxVM volumes and VxFS systems for the databases in the global environment.

For example:

```
/dev/vx/dsk/oradatadg/oradatavol,
/dev/vx/dsk/oradatadg/oraarcvol,
/dev/vx/dsk/orabindg/orabinvol
```

4 Make sure that the mount point directories are created inside the WPAR base directory from the global environment.

For example:

```
# mkdir -p /wpars/orawpar/oracle/oradata
# mkdir -p /wpars/orawpar/oracle/oraarc
# mkdir -p /wpars/orawpar/oracle/orabin
```

Note: While creating the directory, use the complete path of WPAR.

5 Mount the VxVM volumes with VxFS systems on the following mount points.

For example:

```
# mount -V vxfs /dev/vx/dsk/oradatadg/oradatavol\
/wpars/orawpar/oracle/oradata
# mount -V vxfs /dev/vx/dsk/oradatadg/oraarcvol\
/wpars/orawpar/oracle/oraarc
# mount -V vxfs /dev/vx/dsk/orabindg/orabinvol\
/wpars/orawpar/oracle/orabin
```

6 Make sure that the ODM pseudo device `/dev/odm` is mounted on the `/dev/odm`directory in the global environment.

For example:

```
# mount -V odm /dev/odm /dev/odm
```

7 Share the `/dev/odm` directory from the global environment using namefs mount. The mount point `dev/odm` directory must be created inside the WPAR base directory from the global environment. This step enables ODM file access from WPAR.

For example:

```
# mkdir -p /wpars/orawpar/dev/odm
# mount -V namefs /dev/odm /wpars/orawpar/dev/odm
# mount (only related with above commands)
```

```
Mounted                Mounted over                 vfs     date          options
/dev/odm                          /dev/odm                        vxodm   Jun 09 17:30  smartsync
/dev/vx/dsk/oradatadg/oradatavol /wpars/orawpar/oracle/oradata vxfs    Jun 10 14:08
rw,delaylog,suid,ioerror=mwdisable,qio,largefiles
/dev/vx/dsk/oradatadg/oraarcvol  /wpars/orawpar/oracle/oraarc  vxfs    Jun 10 14:08
rw,delaylog,suid,ioerror=mwdisable,qio,largefiles
/dev/vx/dsk/orabindg/orabinvol   /wpars/orawpar/oracle/orabin  vxfs    Jun 10 14:08
rw,delaylog,suid,ioerror=mwdisable,qio,largefiles
/dev/odm                          /wpars/orawpar/dev/odm        namefs    Jun 10 14:44  rw
```

Note: While creating the directory, use the complete path of WPAR.

8 Log on to WPAR.

For example:

```
# clogin orawpar
# df -kg

Filesystem GB blocks Free %Used Iused %Iused Mounted on
Global 2.00 1.94 3% 4670 2% /
Global 1.00 0.99 1% 5 1% /home
Global 5.00 3.81 24% 12095 2% /opt
Global - - - - - /proc
Global 2.00 1.99 1% 70 1% /tmp
Global 5.00 2.60 49% 52955 9% /usr
Global 2.00 1.98 1% 454 1% /var
Global 9.00 8.42 7% 6 1% /oracle/oradata
Global 9.00 8.42 7% 4 1% /oracle/oraarc
Global 9.00 1.86 80% 39074 8% /oracle/orabin
Global 0.00 0.00 -1% 6 100% /dev/odm
```

9 Install and configure Oracle Database server single instance in the system WPAR.

See the IBM and Oracle documentation for the detailed information on creating system WPAR and installing Oracle Database server single instance in a system WPAR on AIX.

10 Link the ODM libraries for Oracle inside the WPAR.

For example:

```
# ln -sf /opt/VRTSodm/lib/libodm64.so libodm11.so
# ls -lta |grep -i libodm11.so
lrwxrwxrwx 1 root system 28 Jun 10 14:51 libodm11.so -> /opt/VRTSodm/lib/libodm64.so
```

For information on configuring Oracle single instance database, see the *InfoScale Storage and Availability Management for Oracle Databases*.

Section 3

Use cases for AIX PowerVM virtual environments

Chapter 3

Application to spindle visibility

This chapter includes the following topics:

About application to spindle visibility using

Datacenters adopt virtualization technology to effectively use the IT-infrastructure and substantially reduce the capital and operational expenditures. If you have adopted virtualization technology in your datacenter, provides you an efficient way of discovering and managing your virtual storage and infrastructure assets.

In your datacenter, helps you view the following relationships:

- Applications in your datacenter that manages and the virtual hosts on which they are running.
- Physical storage in your datacenter that is exported to the virtual machines.

supports the following virtualization technologies:

- Logical partition (LPAR)

For logical partition (LPAR) discovery, can use Hardware Management Console (HMC), `VRTSsfmh` fileset that is installed on the LPAR client, or `VRTSsfmh` fileset installed as a part of DMP on the VIO server.

See "About discovering LPAR and VIO in Arctera InfoScale Operations Manager" on page 53.

For more information, see the documentation.

About discovering LPAR and VIO in Arctera InfoScale Operations Manager

You can use Arctera InfoScale Operations Manager to configure LPAR server, and discover the information that is related to LPARs, VIO clients, and VIO servers in your data center. Agentless discovery of client LPARs and VIO servers is not supported.

Note: The Arctera InfoScale Operations Manager supports only legitimate filename characters in an LPAR profile name. The special characters reserved for Operating System usage (for example space, "\", "$", "!", "&") are not supported. It is recommended to use upper and lower case alphabets, numeric values (0-9), "_" and "-" for the LPAR profile name.

LPAR discovery mechanisms can be grouped into the following categories:

- Discovery using the Hardware Management Console (HMC): The HMC server manages LPAR servers and lets you discover information related to VIO servers and VIO clients. You can use the virtualization management option on the Arctera InfoScale Operations Manager console to add the HMC server to Management Server.

  To add the HMC server to Arctera InfoScale Operations Manager, you need to install the Control Host add-on on the host where the HMC server should be added. Virtual SCSI disks on LPAR client are supported. However, NPIV, or virtual Fibre Channel disks are not supported. Currently, only Virtual SCSI disks backed by native or DMP devices are supported. By configuring HMC server only (without the `VRTSsfmh` package), you can discover information about the exported storage from the VIO server to the VIO clients and the devices that are given to the VIO server from the storage area network (SAN).

- Discovery using the `VRTSsfmh` package that is installed on the LPAR client: The presence of the `VRTSsfmh` package on LPAR client provides additional information about them. This information is correlated with the information that is discovered using the HMC server. Virtual SCSI device discovery, and Virtual SCSI device correlation with the source device in VIO server is also supported.

Note: Arctera InfoScale Operations Manager supports only native disks as the back-end devices for the VIO server. These disks can be controlled by Microsoft Multipath I/O (MPIO) and Dynamic Multi-Pathing (DMP). Disks that are controlled by any third-party multi-pathing software (or the logical volumes), when used as the backing devices, do not have end-to-end correlation available.

- Discovery using the `VRTSsfmh` package that is installed as a part of DMP on the VIO server: When a VIO server having DMP 6.0 is added, it provides the discovery of DMP backed exported storage along with the normal managed host discovery. For end-to-end correlation, DMP 6.0 on the VIO server is required. Storage mapping for DMP backed devices is available only if the VIO server (with DMP installed) is added to Arctera InfoScale Operations Manager Management Server.
- Storage Insight Add-on lets you discover complete information about arrays and LUNs from the SAN, which are allocated to the VIO server.

Note: When an array (consumed by the VIO server) is configured, or a VIO server (with DMP) is added to Arctera InfoScale Operations Manager Management Server, refreshing the corresponding HMC discovery is recommended to view the end-to-end correlation immediately in the Arctera InfoScale Operations Manager console.

See "Prerequisites for LPAR storage correlation support in Arctera InfoScale Operations Manager" on page 55.

About LPAR storage correlation supported in Arctera InfoScale Operations Manager

Arctera InfoScale Operations Manager provides the support for storage correlation of VIO servers and clients. The storage correlation support for VIO servers and clients provides the information that is related to VIO servers and clients storage consumption at each layer. The following VIO server and clients-related information is provided:

- Information about the assigned storage to the VIO client; whether the storage in the VIO client is directly assigned from the storage area network (SAN), or through a VIO server
- Discovery of VIO servers and correlation of VIO server storage with the SAN.

- Detail of the storage that is exported from the VIO server to the VIO client, and the mapping between the VIO server source device and the VIO client target device
- Information about which VIO server participates in storage allocation to the VIO clients, and how much storage is allocated.
- Information about how much storage is allocated to the VIO server, and how much storage is allocated to the VIO clients from that VIO server.
- Information about how much of the allocated storage is consumed by the VIO client for various applications, and file systems.

See "About discovering LPAR and VIO in Arctera InfoScale Operations Manager" on page 53.

Prerequisites for LPAR storage correlation support in Arctera InfoScale Operations Manager

Ensure that the following prerequisites are met before you begin the discovery of LPAR storage correlation:

- You have added the Hardware Management Console (HMC) to Arctera InfoScale Operations Manager Management Server. The HMC server is used to manage LPAR servers and LPARs.
- For end-to-end correlation, you must add the LPAR client (with `the VRTSsfmh` package installed) to Arctera InfoScale Operations Manager Management Server.
- You must have the required administrative privileges (hmcoperator role in HMC) to perform these tasks.
- For end-to-end correlation and discovery of exported storage backed by Dynamic Multi Pathing (DMP) on the VIO server, the VIO server managed host must be added to Arctera InfoScale Operations Manager Management Server.

See "About discovering LPAR and VIO in Arctera InfoScale Operations Manager" on page 53.

Chapter 4

Simplified storage management in VIOS

This chapter includes the following topics:

About simplified management

Independent of how an operating system is hosted, consistent storage management tools save an administrator time and reduce the complexity of the environment. InfoScale products in the guest provide the same command set, storage namespace, and environment as in a non-virtual environment.

The simplified management use case is about ease of use for provisioning virtual machines: setting up the LPARs using the same command set, storage namespace, and environment as in a non-virtual environment.

To implement simplified management use case, configure Dynamic Multi-Pathing (DMP) in the Virtual I/O server (VIOS) and implement DMP or Storage Foundation in the Virtual I/O clients or Logical partitions (LPARs).

- Consistent device naming (DMP)
- Provisioning virtual machines:
  See "About the partition migration process and simplified management" on page 117.
  See "Provisioning data LUNs in a mixed VxVM and LVM environment" on page 147.
- Boot disk management: How to use DMP for the rootvg to simplify system administration and system reliability
  See "Using DMP to provide multi-pathing for the root volume group (rootvg)" on page 81.

DMP is supported in VIOS servers and in LPARs:

- If DMP is installed on a VIOS server, the dmp_native_support tunable is enabled on VIOS by default. This tunable is required for DMP to work with VIOS CLIs and LVM. By default, Volume manager is disabled on VIOS.
- If DMP is installed on an LPAR, the dmp_native_support tunable is disabled by default. Volume manager functionality is fully available.
- If DMP in installed in both VIOS and LPAR, DMP in the LPAR displays extended attributes associated with corresponding vSCSI devices.

InfoScale products supported in LPARs:

- DMP
- Storage Foundation
- Storage Foundation High Availability
- Storage Foundation Cluster File System High Availability
- Storage Foundation for Oracle RAC

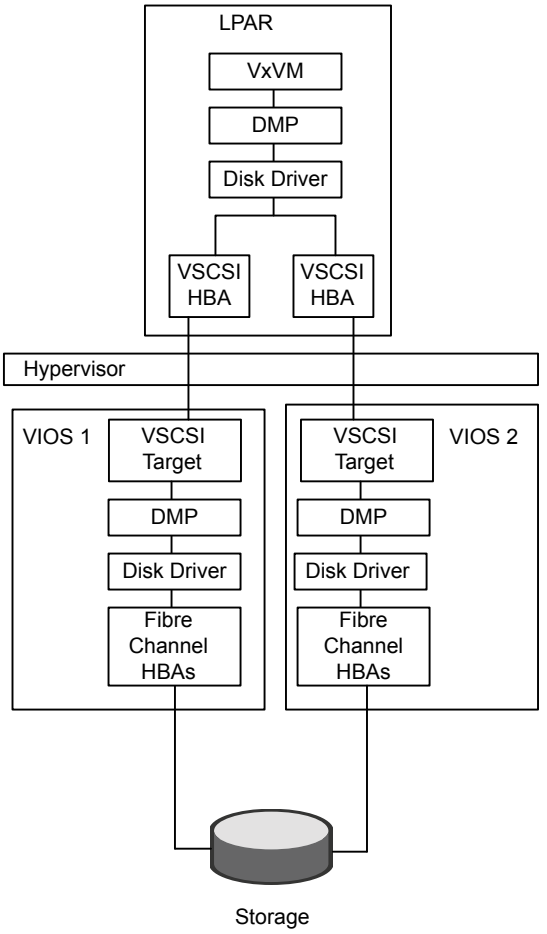
About Dynamic Multi-Pathing in a Virtual I/O server

The Virtual I/O (VIO) server virtualization technology from IBM is a logical partition (LPAR) that runs a trimmed-down version of the AIX operating system.

Virtual I/O servers have APV support, which allows sharing of physical I/O resources between virtual I/O clients.

Figure 4-1 illustrates DMP enablement in the Virtual I/O server.

Figure 4-1 Dynamic Multi-Pathing in the Virtual I/O serve



DMP is fully functional in the Virtual I/O server. DMP administration and management commands (`vxdmpadm`, `vxddladm`, `vxdisk`) must be invoked from the non-restricted root shell.

```
$ oem_setup_env
```

Some example commands:

```
dmpvios1$ vxdmpadm getsubpaths dmpnodename=ibm_ds8x000_0337

NAME     STATE[A] PATH-TYPE[M] CTLR-NAME ENCLR-TYPE ENCLR-NAME  ATTRS
================================================================================
hdisk21 ENABLED(A)  -          fscsi0   IBM_DS8x00  ibm_ds8x000  -
hdisk61 ENABLED(A)  -          fscsi0   IBM_DS8x00  ibm_ds8x000  -
hdisk80 ENABLED(A)  -          fscsi1   IBM_DS8x00  ibm_ds8x000  -
hdisk99 ENABLED(A)  -          fscsi1   IBM_DS8x00  ibm_ds8x000  -

dmpvios1$ vxdmpadm listenclosure all

ENCLR_NAME  ENCLR_TYPE ENCLR_SNO STATUS    ARRAY_TYPE LUN_COUNT FIRMWARE
================================================================================
disk        Disk       DISKS     CONNECTED Disk       1         -
ibm_ds8x000 IBM_DS8x00 75MA641   CONNECTED A/A        6         -
```

See the PowerVM wiki for more in-depth information about VIO server and virtualization:

http://www.ibm.com/developerworks/wikis/display/virtualization/VIO

For more information, see the *PowerVM Virtualization on IBM System p redbook*:

http://www.redbooks.ibm.com/redpieces/abstracts/sg247940.html

About the Volume Manager (VxVM) component in a Virtual I/O server

Volume Manager (VxVM) is a component of InfoScale whose functionality is disabled in Virtual I/O server (VIOS). VxVM commands that manage volumes or disk groups are disabled in the VIO server.

In the VIOS, VxVM does not detect disk format information, so the disk status for VxVM disks is shown as unknown. For example:

```
dmpvios1$ vxdisk list
DEVICE           TYPE     DISK    GROUP   STATUS
disk_0           auto     -       -       unknown
ibm_ds8x000_02c1 auto     -       -       unknown
ibm_ds8x000_0288 auto     -       -       unknown
ibm_ds8x000_029a auto     -       -       unknown
ibm_ds8x000_0292 auto     -       -       unknown
ibm_ds8x000_0293 auto     -       -       unknown
ibm_ds8x000_0337 auto     -       -       unknown
```

In the VIOS, VxVM displays an error if you run a command that is disabled, as follows:

```
dmpvios1$ vxdisk -f init ibm_ds8x000_0288
VxVM vxdisk ERROR V-5-1-5433 Device ibm_ds8x000_0288: init failed:
Operation not allowed. VxVM is disabled.

dmpvios1$ vxdg import datadg
VxVM vxdg ERROR V-5-1-10978 Disk group datadg: import failed:
Operation not allowed. VxVM is disabled.
```

Configuring Dynamic Multi-Pathing (DMP) on Virtual I/O server

You can install DMP in the virtual I/O server (VIOS). This enables the VIO server to export dmpnodes to the VIO clients. The VIO clients access the dmpnodes in the same way as any other vSCSI devices. DMP handles the I/O to the disks backed by the dmpnodes.

For support information concerning running Dynamic Multi-Pathing (DMP) in Virtual I/O server (VIOS), refer to the *InfoScale Release Notes*.

Dynamic Multi-Pathing (DMP) can operate in the Virtual I/O server. Install DMP on the Virtual I/O server.

To install DMP on the Virtual I/O server

1. Log into the VIO server partition.
2. Use the `oem_setup_env` command to access the non-restricted root shell.
3. Install Dynamic Multi-Pathing on the Virtual I/O server.

   See the *Storage Foundation Configuration and Upgrade Guide*.
4. Installing DMP on the VIO server enables the dmp_native_support tunable. Do not set the dmp_native_support tunable to off.

   ```
   dmpvios1$ vxdmpadm gettune dmp_native_support
   Tunable              Current Value  Default Value
   -----------------  --------------- -------------------
   dmp_native_support       on              off
   ```

Migration options for configuring multi-pathing on a Virtual I/O server:

- Migrate from other multi-pathing solutions to DMP on a Virtual I/O server

- Migrate from MPIO to DMP on a Virtual I/O server for a dual-VIOS configuration
- Migrate from PowerPath to DMP on Virtual I/O server for a dual-VIOS configuration

Migrating from other multi-pathing solutions to DMP on Virtual I/O server

DMP supports migrating from AIX MPIO and EMC PowerPath multi-pathing solutions to DMP on Virtual I/O server.

To migrate from other multi-pathing solutions to DMP on a Virtual I/O server

1. Before migrating, back up the Virtual I/O servers to use for reverting the system in case of issues.
2. Shut down all VIO client partitions that are serviced by the VIOS.
3. Log into the VIO server partition. Use the following command to access the non-restricted root shell. All subsequent commands in this procedure must be invoked from the non-restricted shell.

   ```
   $ oem_setup_env
   ```

4. Use commands like `lsdev` and `lsmap` to view the configuration.
5. Unconfigure all VTD devices from all virtual adapters on the system:

   ```
   dmpvios1$ rmdev -p vhost0
   ```

   Repeat this step for all other virtual adapters.
6. Migrate from the third-party device driver to DMP.

   Note that you do not need to do turn on the dmp_native_support again, because it is turned on for VIOS by default. You can use the `vxdmpadm gettune dmp_native_support` command to verify that the tunable parameter is turned on.

   For the migration procedure, see the *Storage Foundation Configuration and Upgrade Guide*.
7. Reboot the VIO Server partition.
8. Use the following command to verify that all Virtual SCSI mappings of TPD multi-pathing solution have been correctly migrated to DMP:

   ```
   dmpvios1$ /usr/ios/cli/ioscli lsmap -all
   ```

9 Repeat step 1 through step 8 for all of the other VIO server partitions of the managed system.

10 After all of the VIO Server partitions are successfully migrated to DMP, start all of the VIO client partitions.

Migrating from MPIO to DMP on a Virtual I/O server for a dual-VIOS configuration

This following example procedure illustrates a migration from MPIO to DMP on the Virtual I/O server, in a configuration with two VIO Servers.

Example configuration values:

```
Managed System:  dmpviosp6
VIO server1: dmpvios1
VIO server2: dmpvios2
VIO clients: dmpvioc1
SAN LUNs: IBM DS8K array
Current multi-pathing solution on VIO server: IBM MPIO

ODM definition fileset required to disable MPIO support
for IBM DS8K array LUNs:
devices.fcp.disk.ibm.rte
```

To migrate dmpviosp6 from MPIO to DMP

1 Before migrating, back up the Virtual I/O server to use for reverting the system in case of issues.

See the IBM website for information about backing up Virtual I/O server.

2 Shut down all of the VIO clients that are serviced by the VIO Server.

```
dmpvioc1$ halt
```

3 Log into the VIO server partition.Use the following command to access the non-restricted root shell. All subsequent commands in this procedure must be invoked from the non-restricted shell.

```
$ oem_setup_env
```

4 The following command shows `lsmap` output before migrating MPIO VTD devices to DMP:

```
dmpvios1$ /usr/ios/cli/ioscli lsmap -all
SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11     0x00000004

VTD               vtscsi0
Status            Available 8100000000000000
Backing device    hdisk21
LUN               0x
Physloc           U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L4
003403700000000

VTD               vtscsi1
Status            Available
LUN               0x8200000000000000
Backing device    hdisk20
Physloc           U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L4
00240C100000000

VTD               vtscsi2
Status            Available
LUN               0x8300000000000000
Backing device    hdisk18
Physloc           U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L4
002409A00000000
```

The VIO Server has MPIO providing multi-pathing to these hdisks. The following commands show the configuration:

```
dmpvios1$ lsdev -Cc disk | egrep "hdisk21|hdisk20|hdisk18"

hdisk18 Available 02-08-02 MPIO Other FC SCSI Disk Drive
hdisk20 Available 02-08-02 MPIO Other FC SCSI Disk Drive
hdisk21 Available 02-08-02 MPIO Other FC SCSI Disk Drive
```

5 Unconfigure all VTD devices from all virtual adapters on the system:

```
dmpvios1 $ rmdev -p vhost0
vtscsi0 Defined
vtscsi1 Defined
vtscsi2 Defined
```

Repeat this step for all other virtual adapters.

6 Migrate the devices from MPIO to DMP.

Unmount the file system and varyoff volume groups residing on the MPIO devices.

Display the volume groups (vgs) in the configuration:

```
dmpvios1$ lsvg
rootvg
brunovg

dmpvios1 lsvg -p brunovg

brunovg:
PV_NAME PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
hdisk19 active   511       501      103..92..102..102..102
hdisk22 active   511       501      103..92..102..102..102
```

Use the `varyoffvg` command on all affected vgs:

```
dmpvios1$ varyoffvg brunovg
```

Install the IBMDS8K ODM definition fileset to remove IBM MPIO support for IBM DS8K array LUNs.

```
dmpvios1$ installp -aXd . devices.fcp.disk.ibm.rte


+-----------------------------------------------------+
             Pre-installation Verification...
+-----------------------------------------------------+
Verifying selections...done
Verifying requisites...done
Results...
Installation Summary
--------------------
Name                        Level   Part  Event  Result
-------------------------------------------------------
devices.fcp.disk.ibm.rte  1.0.0.2 USR   APPLY  SUCCESS
devices.fcp.disk.ibm.rte  1.0.0.2 ROOT  APPLY  SUCCESS
```

7 Reboot VIO server1

```
dmpvios1$ reboot
```

8 After the VIO server1 reboots, verify that all of the existing volume groups on the VIO server1 and MPIO VTDs on the VIO server1 are successfully migrated to DMP.

```
dmpvios1 lsvg -p brunovg

brunovg:
PV_NAME          PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
ibm_ds8000_0292 active   511       501    103..92..102..102..102
ibm_ds8000_0293 active   511       501    103..92..102..102..102
```

Verify the vSCSI mappings of IBM DS8K LUNs on the migrated volume groups:

```
dmpvios1 lsmap -all

SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11     0x00000000
VTD                   vtscsi0
Status                Available
LUN                   0x8100000000000000
Backing device        ibm_ds8000_0337
Physloc

VTD                   vtscsi1
Status                Available
LUN                   0x8200000000000000
Backing device        ibm_ds8000_02c1
Physloc

VTD                   vtscsi2
Status                Available
LUN                   0x8300000000000000
Backing device        ibm_ds8000_029a
Physloc
```

9 Repeat step 1 through step 8 for VIO server2.

10 Start all of the VIO clients using HMC.

Migrating from PowerPath to DMP on a Virtual I/O server for a dual-VIOS configuration

This following example procedure illustrates a migration from PowerPath to DMP on the Virtual I/O server, in a configuration with two VIO Servers.

Example configuration values:

```
Managed System:  dmpviosp6
VIO server1: dmpvios1
VIO server2: dmpvios2
VIO clients: dmpvioc1
SAN LUNs: EMC Clariion array
Current multi-pathing solution on VIO server: EMC PowerPath
```

To migrate dmpviosp6 from PowerPath to DMP

1. Before migrating, back up the Virtual I/O server to use for reverting the system in case of issues.

   See the IBM website for information about backing up Virtual I/O server.

2. Shut down all of the VIO clients that are serviced by the VIO Server.

   ```
   dmpvioc1$ halt
   ```

3. Log into the VIO server partition.Use the following command to access the non-restricted root shell. All subsequent commands in this procedure must be invoked from the non-restricted shell.

   ```
   $ oem_setup_env
   ```

4 The following command shows `lsmap` output before migrating PowerPath VTD devices to DMP:

```
dmpvios1$ /usr/ios/cli/ioscli lsmap -all

SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11     0x00000004

VTD              P0
Status           Available
LUN              0x8100000000000000
Backing device   hdiskpower0
Physloc          U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L4
0034037
00000000

VTD              P1
Status           Available
LUN              0x8200000000000000
Backing device   hdiskpower1
Physloc          U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L40
0240C10
0000000

VTD              P2
Status           Available
LUN              0x8300000000000000
Backing device   hdiskpower2
Physloc          U789D.001.DQD04AF-P1-C5-T1-W500507630813861A-L40
02409A00000000
```

5 Unconfigure all VTD devices from all virtual adapters on the system:

```
dmpvios1$ rmdev -p vhost0
P0 Defined
P1 Defined
P2 Defined
```

Repeat this step for all other virtual adapters.

6 Migrate the devices from PowerPath to DMP.

Unmount the file system and varyoff volume groups residing on the PowerPath devices.

Display the volume groups (vgs) in the configuration:

```
dmpvios1$ lsvg
rootvg
brunovg

dmpvios1$ lsvg -p brunovg

brunovg:
PV_NAME     PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
hdiskpower3 active   511       501    103..92..102..102..102
```

Use the varyoffvg command on all affected vgs:

```
dmpvios1$ varyoffvg brunovg
```

Unmanage the EMC Clariion array from PowerPath control

```
# powermt unmanage class=clariion
hdiskpower0 deleted
hdiskpower1 deleted
hdiskpower2 deleted
hdiskpower3 deleted
```

7 Reboot VIO server1

```
dmpvios1$ reboot
```

8 After the VIO server1 reboots, verify that all of the existing volume groups on the VIO server1 and MPIO VTDs on the VIO server1 are successfully migrated to DMP.

```
dmpvios1$ lsvg -p brunovg

brunovg:
PV_NAME      PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
emc_clari0_138 active  511      501    103..92..102..102..102
```

Verify the mappings of the LUNs on the migrated volume groups:

```
dmpvios1$ lsmap -all

SVSA            Physloc                     Client Partition ID
--------------- --------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11    0x00000000
VTD               P0
Status            Available
LUN               0x8100000000000000
Backing device    emc_clari0_130
Physloc

VTD               P1
Status            Available
LUN               0x8200000000000000
Backing device    emc_clari0_136
Physloc

VTD               P2
Status            Available
LUN               0x8300000000000000
Backing device    emc_clari0_137
Physloc
```

9 Repeat step 1 to step 8 for VIO server2.

10 Start all of the VIO clients.

Configuring Dynamic Multi-Pathing (DMP) pseudo devices as virtual SCSI devices

DMP in the VIO server supports the following methods to export a device to the VIO client:

- DMP node method
  See "Exporting Dynamic Multi-Pathing (DMP) devices as virtual SCSI disks " on page 71.
- Logical partition-based method
  See "Exporting a Logical Volume as a virtual SCSI disk" on page 74.
- File-based method
  See "Exporting a file as a virtual SCSI disk" on page 76.

Exporting Dynamic Multi-Pathing (DMP) devices as virtual SCSI disks

DMP supports disks backed by DMP as virtual SCSI disks. Export the DMP device as a vSCSI disk to the VIO client.

To export a DMP device as a vSCSI disk

1 Log into the VIO server partition.

2 Use the following command to access the non-restricted root shell. All subsequent commands in this procedure must be invoked from the non-restricted shell.

```
$ oem_setup_env
```

3 The following command displays the DMP devices on the VIO server:

```
dmpvios1$ lsdev -t dmpdisk

ibm_ds8000_0287 Available  Veritas DMP Device
ibm_ds8000_0288 Available  Veritas DMP Device
ibm_ds8000_0292 Available  Veritas DMP Device
ibm_ds8000_0293 Available  Veritas DMP Device
ibm_ds8000_029a Available  Veritas DMP Device
ibm_ds8000_02c1 Available  Veritas DMP Device
ibm_ds8000_0337 Available  Veritas DMP Device
```

4 Assign the DMP device as a backing device. Exit from the non-restricted shell to run this command from the VIOS default shell.

```
dmpvios1$ exit

$ mkvdev -vdev ibm_ds8000_0288 -vadapter vhost0
vtscsi3 Available
```

5 Use the following command to display the configuration.

```
$ lsmap -all

SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11     0x00000000
VTD               vtscsi0
Status            Available
LUN               0x8100000000000000
Backing device    ibm_ds8000_0337
Physloc

VTD               vtscsi1
Status            Available
LUN               0x8200000000000000
Backing device    ibm_ds8000_02c1
Physloc

VTD               vtscsi2
Status            Available
LUN               0x8300000000000000
Backing device    ibm_ds8000_029a
Physloc V

TD                vtscsi3
Status            Available
LUN               0x8400000000000000
Backing device    ibm_ds8000_0288
Physloc
```

6 For a dual-VIOS configuration, export the DMP device corresponding to the same SAN LUN on the second VIO Server in the configuration. To export the DMP device on the second VIO server, identify the DMP device corresponding to the SAN LUN as on the VIO Server1.

- If the array supports the AVID attribute, the DMP device name is the same as the DMP device name on the VIO Server1.
- Otherwise, use the UDID value of the DMP device on the VIO Server1 to correlate the DMP device name with same UDID on the VIO Server2. On VIO Server1:

  ```
  $ oem_setup_env
  ```

```
dmpvios1$ lsattr -El ibm_ds8000_0288

attribute value            description        user_settable
dmpname   ibm_ds8x000_0288 DMP Device name    True
pvid      none             Physical volume identifier True
unique_id IBM%5F2107%5F75MA641%5F6005076308FFC61A000000000
0000288
Unique device identifier   True
```

On VIO Server2:

```
$ oem_setup_env

dmpvios2$ odmget -q "attribute = unique_id and
value = 'IBM%5F2107%5F75MA641%5F6005076308FFC61A000000000
0000288'" CuAt

CuAt:

   name = "ibm_ds8000_0288"
   attribute = "unique_id"
   value = "IBM%5F2107%5F75MA641%5F6005076308FFC61A00
   00000000000288"
   type = "R"
   generic = "DU"
   rep = "s"
   nls_index = 4
```

7 Use the DMP device name identified in step 6 to assign the DMP device as a backing device. Exit from the non-restricted shell to run this command from the VIOS default shell.

```
dmpvios1$ exit

$ mkvdev -vdev ibm_ds8000_0288 -vadapter vhost0
vtscsi3 Available
```

8 Use the following command to display the configuration.

```
$ lsmap -all

SVSA            Physloc                     Client Partition ID
--------------- --------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11  0x00000000
VTD               vtscsi0
Status            Available
LUN               0x8100000000000000
Backing device    ibm_ds8000_0337
Physloc

VTD               vtscsi1
Status            Available
LUN               0x8200000000000000
Backing device    ibm_ds8000_02c1
Physloc

VTD               vtscsi2
Status            Available
LUN               0x8300000000000000
Backing device    ibm_ds8000_029a
Physloc V

TD                vtscsi3
Status            Available
LUN               0x8400000000000000
Backing device    ibm_ds8000_0288
Physloc
```

Exporting a Logical Volume as a virtual SCSI disk

Dynamic Multi-Pathing (DMP) supports vSCSI disks backed by a Logical Volume. Export the Logical Volume as a vSCSI disk to the VIO client.

To export a Logical Volume as a vSCSI disk

1 Create the volume group.

```
$ mkvg -vg brunovg ibm_ds8000_0292 ibm_ds8000_0293
brunovg
```

The following command displays the new volume group:

```
$ lsvg -pv brunovg
brunovg:
PV_NAME          PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
ibm_ds8000_0292 active   494       494      99..99..98..99..99
ibm_ds8000_0293 active   494       494      99..99..98..99..99
```

2 Make a logical volume in the volume group.

```
$ mklv -lv brunovg_lv1 brunovg 1G
brunovg_lv1
```

The following command displays the new logical volume:

```
$ lsvg -lv brunovg
brunovg:
LV NAME       TYPE   LPs   PPs   PVs  LV STATE      MOUNT POINT
brunovg_lv1   jfs    256   256   1    closed/syncd  N/A
```

3 Assign the logical volume as a backing device.

```
$ mkvdev -vdev brunovg_lv1 -vadapter  vhost0
vtscsi4 Available
```

4 Use the following command to display the configuration.

```
$ lsmap -all

SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11  0x00000000
VTD                   vtscsi0
Status                Available
LUN                   0x8100000000000000
Backing device        ibm_ds8000_0337
Physloc

VTD                   vtscsi1
Status                Available
LUN                   0x8200000000000000
Backing device        ibm_ds8000_02c1
Physloc

VTD                   vtscsi2
Status                Available
LUN                   0x8300000000000000
Backing device        ibm_ds8000_029a
Physloc

VTD                   vtscsi3
Status                Available
LUN                   0x8400000000000000
Backing device        ibm_ds8000_0288
Physloc

VTD                   vtscsi4
Status                Available
LUN                   0x8500000000000000
Backing device        brunovg_lv1
Physloc
```

Exporting a file as a virtual SCSI disk

Dynamic Multi-Pathing (DMP) supports vSCSI disks backed by a file. Export the file as a vSCSI disk to the VIO client.

To export a file as a vSCSI disk

1 Create the storage pool.

```
$ mksp brunospool ibm_ds8000_0296
brunospool
0516-1254 mkvg: Changing the PVID in the ODM.
```

2 Create a file system on the pool.

```
$ mksp -fb bruno_fb -sp brunospool -size 500M
bruno_fb
File system created successfully.
507684 kilobytes total disk space.
New File System size is 1024000
```

3 Mount the file system.

```
$ mount


node       mounted   mounted over   vfs    date     options
---------- ----------------------   ----- --------------------
/dev/hd4       /        jfs2      Jul 02 14:47 rw,log=/dev/hd8
/dev/hd2      /usr      jfs2      Jul 02 14:47 rw,log=/dev/hd8
/dev/hd9var   /var      jfs2      Jul 02 14:47 rw,log=/dev/hd8
/dev/hd3      /tmp      jfs2      Jul 02 14:47 rw,log=/dev/hd8
/dev/hd1      /home     jfs2      Jul 02 14:48 rw,log=/dev/hd8
/dev/hd11admin /admin   jfs2      Jul 02 14:48 rw,log=/dev/hd8
/proc         /proc     procfs    Jul 02 14:48 rw
/dev/hd10opt  /opt      jfs2      Jul 02 14:48 rw,log=/dev/hd8
/dev/livedump /var/adm/ras/livedump jfs2 Jul 02 14:48 rw,log=
/dev/hd8
/dev/bruno_fb /var/vio/storagepools/bruno_fb jfs2 Jul 02 15:38
rw,log=INLINE
```

4 Create a file in the storage pool.

```
$ mkbdsp -bd bruno_fbdev -sp bruno_fb 200M
Creating file "bruno_fbdev" in storage pool "bruno_fb".
bruno_fbdev
```

5 Assign the file as a backing device.

```
$ mkbdsp -sp bruno_fb -bd bruno_fbdev -vadapter vhost0
Assigning file "bruno_fbdev" as a backing device.
vtscsi5 Available
bruno_fbdev
```

6 Use the following command to display the configuration.

```
$ lsmap -all

SVSA            Physloc                      Client Partition ID
--------------- ---------------------------- ------------------
vhost0          U9117.MMA.0686502-V2-C11     0x00000000
...
...
VTD                vtscsi5
Status             Available
LUN                0x8600000000000000
Backing device     /var/vio/storagepools/bruno_fb/bruno_fbdev
Physloc
```

Extended attributes in VIO client for a virtual SCSI disk

Using Dynamic Multi-Pathing (DMP) in the a Virtual I/O server enables the DMP in the VIO Client to receive the extended attributes for the LUN. This enables the client LPAR to view back-end LUN attributes such as thin, SSD, and RAID levels associated with the vSCSI devices.

For more information about extended attributes and the prerequisites for supporting them, see the following tech note:

https://www.veritas.com/support/en_US/article.TECH77062.html

Configuration prerequisites for providing extended attributes on VIO client for virtual SCSI disk

Dynamic Multi-Pathing (DMP) in VIO client will provide extended attributes information of backend SAN LUN. The following conditions are prerequisites for using extended attributes on the VIO client:

- VIO client has vSCSI disks backed by SAN LUNs.
- In the VIO Server partition, DMP is controlling those SAN LUNs.
- On VIO client, DMP is controlling the vSCSI disks.

Displaying extended attributes of virtual SCSI disks

When a VIO client accesses a virtual SCSI disk that is backed by a Dynamic Multi-Pathing (DMP) device on the a Virtual I/O server, the VIO client can access the extended attributes associated with the virtual SCSI disk.

The following commands can access and display extended attributes information associated with the vSCSI disk backed by DMP device on a Virtual I/O server.

- `vxdisk -e list`
- `vxdmpadm list dmpnodename=<daname>`
- `vxdmpadm -v getdmpnode dmpnodename=<daname>`
- `vxdisk -p list <daname>`

For example, use the following command on the VIO client dmpvioc1:

```
# vxdisk -e list
```

```
DEVICE           TYPE          DISK   GROUP  STATUS       OS_NATIVE_NAME  ATTR
ibm_ds8x000_114f auto:LVM      -      -      LVM          hdisk83         std
3pardata0_3968   auto:aixdisk  -      -      online thin  hdisk84         tp
```

```
# vxdmpadm list dmpnode dmpnodename=3pardata0_3968
```

```
dmpdev          = 3pardata0_3968
state           = enabled
enclosure       = 3pardata0
cab-sno         = 744
asl             = libvxvscsi.so
vid             = AIX
pid             = VDASD
array-name      = 3PARDATA
array-type      = VSCSI
iopolicy        = Single-Active
avid            = 3968
lun-sno         = 3PARdata%5FVV%5F02E8%5F2AC00F8002E8
udid            = AIX%5FVDASD%5F%5F3PARdata%255FVV%255F02E8%255F2AC00F8002E8
dev-attr        = tp
```

```
###path          = name state type transport ctlr hwpath aportID aportWWN attr
path            = hdisk84 enabled(a) - SCSI vscsi1 vscsi1 3 - -
```

Virtual IO client adapter settings for Dynamic Multi-Pathing (DMP) in dual-VIOS configurations

Arctera recommends the following Virtual I/O client (VIO client) adapter settings when using Dynamic Multi-Pathing (DMP) in dual-VIOS configurations:

- Set the vscsi_err_recov attribute to fast_fail.
  The virtual SCSI (vSCSI) adapter driver uses the vscsi_err_recov attribute, which is similar to the attribute fc_error_recov for physical fibre channel (FC) adapters. When this parameter is set to fast_fail, the VIO client adapter sends a FAST_FAIL datagram to the VIO server so that the I/O fails immediately, rather than being delayed.
- Enable the vscsi_path_to attribute.
  This attribute allows the virtual client adapter driver to determine the health of the VIO Server and improve path failover processing. The value of this attribute defines the number of seconds that the vSCSI client adapter waits for commands sent to the vSCSI server adapter to be serviced. If that time is exceeded, the vSCSI client adapter fails the outstanding requests. If DMP is present, another path to the disk will be tried to service the requests.
  A value of 0 (default) disables this functionality.

To set the VIO client adapter settings

1. Set the vscsi_err_recov attribute to fast_fail, and the vscsi_path_to attribute to a non-zero number. For example:

   ```
   # chdev -a vscsi_err_recov=fast_fail -a vscsi_path_to=30 -l vscsi0
   ```

2. Verify the settings.

   ```
   # lsattr -El vscsi0
   vscsi_err_recov        fast_fail
   vscsi_path_to          30
   ```

3. Repeat step 1 and step 2 for each vSCSI client adapter.

Using DMP to provide multi-pathing for the root volume group (rootvg)

In many cases, the use of MPIO for the rootvg creates a situation with dual multi-pathing tools. To simplify system administration and improve system reliability, use DMP to provide multi-pathing for the rootvg.

DMP is supported for the rootvg on vSCSI, NPIV, and physical HBAs. DMP is also supported for alternate root disks and root disks with multiple volumes.

To use DMP on the rootvg, DMP requires a vendor-specific ODM predefined fileset. Arctera includes the predefined filesets for vSCSI devices in the Arctera product distribution. For other devices, obtain and install the ODM predefined fileset from the storage vendor. For example, for the IBM DS array, install the `devices.fcp.disk.ibm.rte` fileset.

http://www-1.ibm.com/support/docview.wss?rs=540&context=ST52G7&dc=D400&q1=host+script&uid=ssg1S4000199&loc=en_US&cs=utf-8&lang=en

Rootability is achieved by using the `vxdmpadm` command, which uses the OS Native stack support feature internally.

To get help about rootability

- Run the following command:

```
# vxdmpadm help native

Manage DMP support for AIX boot volume group(rootvg)
Usage:
  vxdmpadm native { enable | disable } vgname=rootvg
  vxdmpadm native list [ vgname=rootvg ]
  vxdmpadm native { release | acquire } [ vgname=rootvg ]
where,
  enable   Enable DMP support for AIX boot volume group(rootvg)
  disable  Disable DMP support for AIX boot volume group(rootvg)
  list     List boot paths on which DMP support is enabled
  release  Giveback pvid to OS device paths corresponding to rootvg
  acquire  Takeover pvid from OS device paths corresponding to rootvg
```

To enable rootability

1. Run the following command:

```
# vxdmpadm native enable vgname=rootvg
```

2. Reboot the system to enable DMP support for LVM bootablity.

To disable rootability

1 Run the following command:

```
# vxdmpadm native disable vgname=rootvg
```

2 Reboot the system to disable DMP support for LVM bootability.

To monitor rootability

- Run the following command:

```
# vxdmpadm native list
PATH          DMPNODENAME
==========================
hdisk64       ams_wms0_302
hdisk63       ams_wms0_302
```

For more information about using DMP with rootvg, see the *Dynamic Multi-Pathing Administrator's Guide*.

Boot device management on NPIV presented devices

N_Port ID Virtualization(NPIV) is a Fibre Channel industry standard technology that provides the capability to assign a physical Fibre Channel adapter multiple unique world wide port names (WWPNs). NPIV enables the Virtual I/O Server (VIOS) to provision entire dedicated logical ports to client LPAR's rather than individual LUNs. A physical Fibre Channel HBA in VIOS can be shared across multiple guest operating systems in a virtual environment.

Dynamic Multi-Pathing (DMP) supports the NPIV presented devices for the rootvg, within the requirements outlined in the vendor support matrix.

Requirements for boot device management on NPIV-presented devices

Requirements for boot device management on NPIV-presented devices:

- Any computer with Power 6, Power 7, Power 8, or Power 9 architecture.
- SAN Switch & FC Adapters should be NPIV capable.
- At least one 8 GB PCI Express Dual Port FC Adapter in VIOS.
- VIOC Minimum OS-level:
  - AIX 7.2 TL4, TL5
  - AIX 7.3 TL0, TL1

- VIO Server Version 2.1 with Fix Pack 20.1 or later
- HMC 7.3.4 or later up to 9

Using Dynamic Multi-Pathing (DMP) on rootvg for boot device management

All the LUNs presented through NPIV for a client LPAR have the characteristics of a dedicated HBA. Therefore the procedure for using DMP on rootvg devices from NPIV presented devices is similar to using DMP on rootvg devices from physical HBA. Use of DMP on rootvg is supported through `vxdmproot native` command.

To use DMP on rootvg for boot device management

- For using DMP on rootvg:

  See "Using DMP to provide multi-pathing for the root volume group (rootvg)" on page 81.

Using NPIV for data volumes

The behavior of data volumes presented through NPIV devices is similar to those presented through physical HBA. All SCSI device inquiry operations work and SCSI-3 persistent reservation functionality is also supported, enabling the use of SCSI-3 I/O Fencing if the underlying storage supports.

To use NPIV for data volumes

- No special handling is required for data volumes.

Boot Devices on vSCSI, NPIV for data volumes

A hybrid solution is supported where the AIX rootvg is placed on vSCSI devices presented through a VIO pair with application data volumes presented through NPIV. This solution is often chosen to facilitate NPIV troubleshooting as well as presenting a consistent NIM installation profile.

Chapter 5

Virtual machine (logical partition) availability

This chapter includes the following topics:

About virtual machine (logical partition) availability

Cluster Server (VCS) provides high availability for logical partitions (LPARs). If an LPAR crashes, it will be automatically restarted or failed over to another physical server. An LPAR can crash due to an LPAR error, a Virtual I/O server (VIOS) error, or physical server error. A VCS cluster can comprise of LPARs. VCS running on these LPARs can manage other LPARs on the physical server that the VCS LPARs runs on. VCS LPARs managing other LPARs are referred to as management LPARs (MLPARs).

VCS in the management LPAR

VCS provides high availability for the AIX LPARs within a physical server. VCS is run in the control point which is an LPAR that is designated for management of other LPARs. The management LPARs on different physical servers form a VCS cluster.

VCS runs in one management LPAR on each physical server. The management LPAR provides high availability to the other LPARs on the same physical server, known as managed LPARs. Each managed LPAR is simply a resource that is managed and monitored by VCS running on the management LPAR, with the help of LPAR agent. This capability allows VCS to monitor the individual LPAR as an individual resource. VCS can restart the service group that has the LPAR resource on the same physical server or fail-over to another physical server.

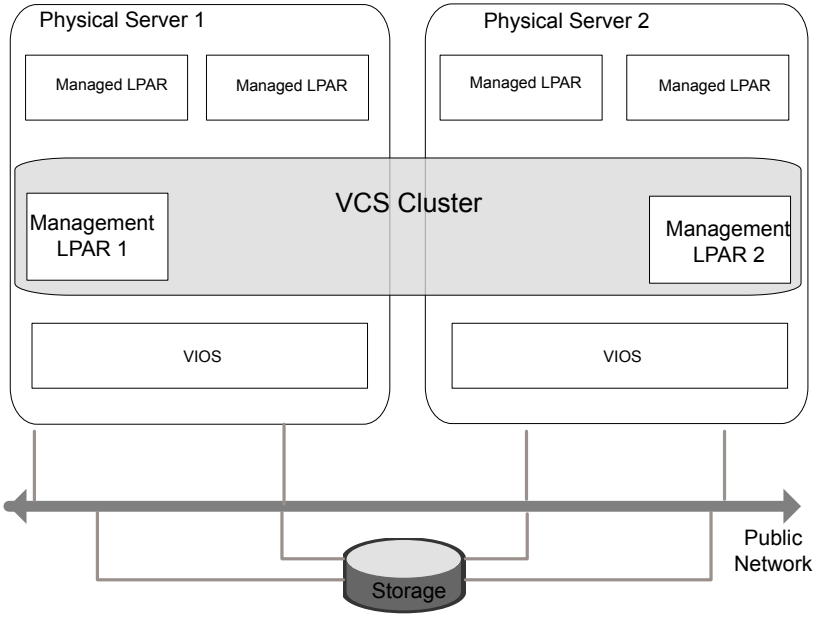
The management LPAR views the LPARs that it manages as virtual machines but does not have visibility into the applications on the managed LPARs. The management LPAR cluster does not monitor resources inside the managed LPARs.

A VCS cluster is formed among the management LPARs in this configuration. The VCS cluster provides failover for the managed LPARs between the management LPARs.

- Each physical server where you want VCS to manage LPARs should have one management server.
- VCS supports only one management LPAR per physical server.
- Each managed LPAR resource can have only one VCS system on one physical server in the system list.
- For a VCS configuration example:
  See "Configuring VCS service groups to manage the LPAR" on page 91.

Figure 5-1 provides an example of VCS in the management LPAR.

Figure 5-1 VCS in the management LPAR



This configuration also provides high availability for applications running on the management LPAR. The VCS cluster manages and controls the applications and services that run inside the management LPARs. Any faulted application or service is failed over to other management LPARs in the cluster.

Note: The managed LPARs cannot be in a cluster configuration.

Setting up management LPAR

Following is a high-level overview of the steps required for setting up the management LPARs.

Setting up management LPAR

1 Install VCS on all nodes of the management LPARs cluster. For information about installing VCS, refer to the *Cluster Server Configuration and Upgrade Guide*.

2 Make sure that the HMC is at the supported level.

3 Make sure that the VIOS are at the supported level.

4 Configure password-less SSH communication from the management LPAR to the HMC. This step is required for all nodes of the cluster even if a node has no LPAR resource.

See " Configuring password-less SSH communication between VCS nodes and HMC" on page 87.

5 The managed LPARs are managed via HMC. Ensure that the network connection between the physical servers and the HMC has redundancy.

6 Set auto restart of management LPAR to "on".

7 Ensure that PhysicalServer system level attribute has the physical server name.

Use the following command to retrieve the physical server name.

```
# lssyscfg -r sys -F name
```

8 Set up the managed LPARs.

See "Setting up managed LPARs" on page 88.

9 Configure the LPARs that need to be managed and monitored as VCS LPAR resources.

See "Bundled agents for managing the LPAR" on page 90.

See the *Cluster Server Bundled Agents Reference Guide*.

10 Set the Service Group level attribute SysDownPolicy = {"AutoDisableNoOffline"} for groups that have LPAR resources.

See "Managing logical partition (LPAR) failure scenarios" on page 92.

For more information on the Service Group level attribute SysDownPolicy, see the *Cluster Server User's Guide*.

Configuring password-less SSH communication between VCS nodes and HMC

To use remote command operations on the HMC, you must have SSH installed on the LPARs in the VCS cluster. You must configure the HMC to allow password-less SSH access from these LPARs. Refer to the appropriate IBM AIX documentation for information.

To verify that you have password-less SSH

- From each LPAR in the cluster, execute the following command to test if the password-less access works:

```
> ssh -l hscroot hmc2.veritas.com
   Last login:Thur Jun 16 22:46:51 2011 from 10.182.9.34
hscroot@hmc2:~>
```

Setting up managed LPARs

The following procedure provides a high-level overview of how to set up LPARs that VCS manages.

For detailed instructions on creating and configuring a LPAR, refer to the IBM PowerVM Guide.

To set up managed LPARs

1. Ensure CPU and memory resources are available to create managed LPARs on all physical servers in the cluster, where the managed LPAR can start.
2. Install VCS on all the management LPARs, to manage the LPAR.

   For information about installing VCS, see the *Cluster Server Configuration and Upgrade Guide*.
3. Create the LPAR profile on all the physical servers whose management LPAR is in the SystemList for the LPAR resource.

   See "Creating an LPAR profile " on page 88.
4. Set auto restart of the managed LPAR to "off" via HMC when VCS is managing the LPAR.
5. The boot disk should be shared and should be accessible from all the physical servers where the LPAR can fail over.
6. Verify if the LPAR can fail over to the other physical servers.
7. Configure the LPAR as a resource in VCS.

   See "Configuring VCS service groups to manage the LPAR" on page 91.

Creating an LPAR profile

The following steps describe how to create LPAR profile:

To create an LPAR profile

1 Identify the disk on which the AIX OS is to be installed for the LPAR. This disk should be on shared storage in order for LPAR to be capable of failover across physical servers. Change the following attributes of the disk on all VIOS of the physical servers that the LPAR will be configured to boot on.

```
vio1#chdev -a hcheck_cmd=inquiry -l hdisk7
vio1#chdev -a hcheck_interval=60 -l hdisk7 -P
vio1#chdev -a pv=yes -l hdisk7
vio1#chdev -a reserve_policy=no_reserve
```

2 Create the Virtual SCSI Host adapter on all VIOS on which the LPAR will be configured to boot on. Reboot the VIO, and then map the OS disk to this host adapter.

3 Log in to HMC and create the LPAR profile. The following example shows creating an LPAR profile.

```
hscadmin1@hmc2.veritas.com:~> lssyscfg -r sys -F name
 PServer1-SN100129A
 PServer2-SN100130A

hscadmin1@hmc2.veritas.com:~> lssyscfg -m PServer1-SN100129A -r lpar \
 -F name
Pserver1_VIO1

hscadmin1@hmc2.veritas.com:~> mksyscfg -m PServer1-SN100129A -r lpar \
-i name=lpar_test,lpar_env=aixlinux,profile_name=lpar_test,min_mem=512,\
desired_mem=512,max_mem=512,proc_mode=shared,sharing_mode=uncap,\
uncap_weight=128,min_proc_units=0.1,desired_proc_units=0.4,\
max_proc_units=2.0,min_procs=1,desired_procs=2,max_procs=4,\
lpar_io_pool_ids=none,max_virtual_slots=10,auto_start=1,\
boot_mode=norm,power_ctrl_lpar_ids=none,conn_monitoring=0,\
virtual_eth_adapters=2/1/1//0/1,virtual_scsi_adapters=3/client/1//10/1"

hscadmin1@hmc2.veritas.com:~> lssyscfg -m PServer1-SN100129A \
-r lpar -F name
 Pserver1_VIO1
 lpar_test
```

The virtual Ethernet adapter's VLAN ID should match that of VIO server in order for connectivity to outside network, the virtual scsi adapter's remote-lpar-ID/remote-lpar-name/remote-slot-number should match with that of VIO's partition ID, VIO's name and VIO's virtual SCSI Host adapter ID that has the OS disk mapped for this LPAR. Note: The VIO's virtual SCSI Host adapter that is assigned for this LPAR should have any partition and any slot option if this LPAR is capable and might be used for LPM in future (in addition to VCS failover capability).

4 Create the same profile on all physical servers where the LPAR can fail over.

5 Verify that the LPAR can boot on the physical servers where the profile has been created.

Bundled agents for managing the LPAR

The LPAR agent can be used to manage and provide high availability for LPARs.

The LPAR agent performs the following functions using HMC CLIs:

- Open: Blocks migration of the management LPAR. Get the information required by the LPAR agent.
- Monitor: Monitors the status of LPAR.
- Online: Starts the LPAR.
- Offline: Shuts down the LPAR.
- Clean: Stops the LPAR forcefully.
- Shutdown: Unblock migration of the management LPAR.
- Migrate: Migrates the LPAR.

Configuring VCS service groups to manage the LPAR

You must configure a VCS service group to manage the LPAR.

To configure LPAR service groups

1. Create a failover service group for LPAR.
2. Set the PhysicalServer attribute of all the systems (which are management LPARs) using the name of the physical server (managed system name).
3. Set SysDownPolicy = { "AutoDisableNoOffline" } for this group.
4. Configure all the cluster nodes (management LPARs) in the SystemList attribute where the managed LPAR can fail over.
5. Configure LPAR resource for the managed LPAR.

The sample main.cf for a VCS failover cluster for managed LPARs:

```
include "types.cf"

cluster cluster01 (
        )

system aixnode55mp1 (
        PhysicalServer = sys1
        )

system aixnode56mp1 (
        PhysicalServer = sys2
        )



group LPAR_aixnode5556mp2 (
```

```
SystemList = { aixnode55mp1 = 0, aixnode56mp1 = 1 }
SysDownPolicy = { AutoDisableNoOffline }
)

LPAR aixnode5556mp2 (
        LPARName = aixnode5556mp2
        MCUser = { hscroot, hscroot }
        MCName = { hmc6, hmc7 }
        VIOSName @aixnode55mp1 = { aixnode55vio1, aixnode55vio2 }
        VIOSName @aixnode56mp1 = { aixnode56vio1, aixnode56vio2 }
        RestartLimit = 1
        )
```

Managing logical partition (LPAR) failure scenarios

VCS handles the LPAR failures in the following cases.

Table 5-1 Failure scenarios and their resolutions

Failure scenario	Resolution
Physical server is down	When the physical server is down, the management LPAR as well as managed LPARs will be down. In this case, the managed LPARs which are running will be failed over to another system by VCS using the sysoffline trigger with the help of HMC. Ensure that HMC access is setup on all nodes of the cluster even if the node is not managing any LPAR. If the managed LPAR is configured for live migration, make sure that profile file for the managed LPAR is created on other management LPARs and its path is configured in ProfileFile attribute. For details on ProfileFile attribute and creation of profile file: See "Providing logical partition (LPAR) failover with live migration" on page 121.

Table 5-1 Failure scenarios and their resolutions *(continued)*

Failure scenario	Resolution
Management LPAR is down but physical server is up	When the management LPAR is down, the physical server may not be down. The managed LPARs might be running. In this case, it is not desirable to automatically failover the managed LPARs. To ensure that the managed LPAR is not automatically failed over, the group that has LPAR resource should have SysDownPolicy = { "AutoDisableNoOffline" }. With this the groups will remain autodisabled on system fault. You can online the LPAR on any other system by setting autoenable for the group, after ensuring that the LPAR is down on the faulted system.
VIO servers are down	When all the VIO servers which are providing virtual resources to the managed LPARs are down, then the managed LPARs are failed over to another host. Ensure that VIOSName attribute of the LPAR resources is populated with list of all VIO servers which are servicing that LPAR. If VIOSName is not populated, managed LPARs will not be failed over in case of VIO server(s) crash. If any one of the VIO servers specified in VIOSName attribute is running, LPAR agent won't failover the managed LPARs.
HMC is down	If the environment has redundant HMC, then even if one of the HMC goes down, LPAR agent can still manage the LPARs without any issues. For this, ensure that MCName and MCUser attributes are populated with both HMC details.

Chapter 6

Simplified management and high availability for IBM Workload Partitions

This chapter includes the following topics:

About IBM Workload Partitions

Workload Partitions allow administrators to virtualize the AIX operating system, by partitioning an AIX operating system instance into multiple environments. Each environment within the AIX operating system instance is called a workload partition (WPAR). One WPAR can host applications and isolate the applications from applications executing in other WPARs. WPAR is a pure software solution and has no dependencies on hardware features.

The WPAR solution allows for fewer operating system images on your IBM System p partitioned server. Prior to WPARs, you had to create a new Logical Partition (LPAR) for each new "isolated" environment. You can instead use multiple WPARs within one LPAR, in many circumstances.

In an LPAR environment, each LPAR requires its own operating system image and a certain number of physical resources. While you can virtualize many of these resources, some physical resources must be allocated to the system for each LPAR. Furthermore, you need to install patches and technology upgrades to each LPAR. Each LPAR requires its own archiving strategy and DR strategy. It also takes some time to create an LPAR; you also need to do this outside of AIX, through a Hardware Management Console (HMC) or the Integrated Virtualization Manager (IVM).

In contrast, WPARs are much simpler to manage and can be created from the AIX command line or through SMIT. WPARs allow you to avoid the biggest disadvantage of LPARs: maintaining multiple images, and therefore possibly over-committing expensive hardware resources, such as CPU and RAM. While logical partitioning helps you consolidate and virtualize hardware within a single box, operating system virtualization through WPAR technology goes one step further and allows for an even more granular approach of resource management.

The WPAR solution shares operating system images and is clearly the most efficient use of CPU, RAM, and I/O resources. Rather than a replacement for LPARs, WPARs are a complement to them and allow one to further virtualize application workloads through operating system virtualization. WPARs allow for new applications to be deployed much more quickly.

WPARs have no real dependency on hardware and can even be used on POWER4 systems that do not support IBM's PowerVM (formerly known as APV). For AIX administrators, the huge advantage of WPARs is the flexibility of creating new environments without having to create and manage new AIX partitions.

On the other hand, it's important to understand the limitations of WPARs. For example, each LPAR is a single point of failure for all WPARs that are created within the LPAR. In the event of an LPAR problem (or a scheduled system outage), all underlying WPARs are also affected.

The following sections describe the types of WPARs:

- System workload partition: the system WPAR is much closer to a complete version of AIX. The system WPAR has its own dedicated, completely writable file-systems along with its own inetd and cron. You can define remote access to the System workload partition.
- Application workload partition: application WPARs are lightweight versions of virtualized OS environments. They are extremely limited and can only run application processes, not system daemons such as inetd or cron. You cannot even define remote access to this environment. These are only temporarily objects; they actually disintegrate when the final process of the application partition ends, and as such, are more geared to execute processes than entire applications.

About using IBM Workload Partitions (WPARs) with InfoScale products

You can use WPARs when you need an isolated environment, especially if you do not want to create new LPARs because of the limitation of the available resources. Here are a few recommended scenarios:

- Application/workload isolation
- Quickly testing an application

WPARs share the global resources with other WPARs in the same LPAR, which limits the usefulness of WPARs in some situations.

We recommend not using WPARs in the following situations:

- Security: WPAR processes can be seen by the global environment from the central LPAR. If you are running a highly secure type of system, this may be a problem for you from a security standpoint. Further, the root administrator of your LPAR will now have access to your workload partition, possibly compromising the security that the application may require.
- Performance: Each WPAR within the LPAR uses the same system resources of the LPAR. You need to be more careful when architecting your system and also when stress testing the system.
- Physical devices: Physical devices are not supported within a WPAR. More details on WPAR administration can be found in the IBM red book on WPARs at
  http://www.redbooks.ibm.com/abstracts/sg247431.html

Limitations for using InfoScale products in AIX WPARs:

- VxFS inside an AIX WPARs is not a supported configuration.
- ODM inside an AIX WPARs is not a supported configuration
- VRTSvxfs and VRTSodm are not installed inside the WPAR for this reason.

Implementing InfoScale support for WPARs

This section describes File System (VxFS) support for workload partitions (WPARs). Currently, there are no VxVM operations available within a system WPAR, so any VxFS file system that is needed for data use must be created in the global environment, then set up so the WPAR can access it. The VxFS (local mount only) is supported inside the workload partition (WPAR) environment. Cluster mount is not yet supported inside a WPAR. WPAR can have both root and non-root partitions as VxFS file system.

In Storage Foundation, there is limited support for WPARs, as follows:

- All the Storage Foundation filesets must be installed and configured in the global partition of AIX.
- Storage Foundation can only be administered from the global partition.

There are two ways to use a local mount VxFS file system inside WPAR environment.

- Using a VxFS file system within a single system WPAR
- Using VxFS as a shared file system

Using a VxFS file system within a single system WPAR

The following procedure describes how to set up a WPAR with VxFS for non-root partition.

To set up WPAR with VxFS for non-root partition

1 Create a vxfs filesystem in the global environment:

```
# /opt/VRTS/bin/mkfs -V vxfs /dev/vx/rdsk/testvg/vol1
```

2 Create a WPAR. For example, use the following command.

```
# mkwpar -n wpar1
```

For other options while creating WPARs, refer to the IBM Redbook for WPAR.

3 List the WPAR.

```
# lswpar
Name              State Type Hostname         Directory
-------------------------------------------------------------------
wpar1 D     S    wpar1 /wpars/wpar1
```

4 The above output shows that WPAR does not have the devices. To get the vxfs file system in WPAR, create the file system in the global environment. Then mount it to the WPAR directories which are located at `/wpar/wparname/`

```
# mkdir /wpars/wpar1/vxfs_dir
# mount -V vxfs /dev/vx/dsk/testdg/vol1 \
/wpars/wpar1/vxfs_dir
```

5 Start the WPAR:

```
# startwpar -Dv wpar1 2>/startwpar_tl2
```

6 Log in to the WPAR.

```
# clogin hostname
```

For example, to log in to the WPAR wpar1:

```
# clogin wpar1
```

7 The following output shows the VxFS mount point in the WPAR.

```
# mount
```

```
node      mounted mounted over vfs      date            options
------  ------- ------------ ----   ------------ -------------
       Global   /               jfs2   Jun 23 03:15 rw,log=INLINE
            Global   /home        jfs2   Jun 23 03:15 rw,log=INLINE
       Global   /opt              namefs Jun 23 03:15 ro
       Global   /proc         namefs Jun 23 03:15 rw
             Global  /tmp             jfs2      Jun 23 03:15 rw,log=INLINE
       Global   /usr              namefs    Jun 23 03:15 ro
       Global   /var              jfs2      Jun 23 03:15 rw,log=INLINE
       Global   /vxfs_dir     vxfs      Jun 23 03:14 rw,delaylog,
                       suid,ioerror=mwdisable,qio,largefiles
```

8 To stop the WPAR, use the following command:

```
# stopwpar -Dv wpar1 2>/wpar1_tl2
```

WPAR with root (/) partition as VxFS

The / (root) partition of any WPAR can be created as vxfs. Previously, it was mandatory to have the root partition as JFS2. Other mount points appear as previously but root partition can be VxFS.

To set up WPAR with root (/) partition as VxFS

1 Create the / (root) partition of the WPAR as VxFS.

```
# mkwpar -n fsqawpar -M directory=/ dev=/dev/vx/rdsk/rootdg/vol2 vfs=vxfs
```

2 Start the WPAR.

```
# startwpar -v fsqawpar 2>/fsqawpar_tl2
```

3 Login to the WPAR.

```
# clogin fsqawpar
```

4 Other mount points appear as previously while root can be VxFS.

```
# mount
node        mounted     mounted over vfs         date             options
------  --------    ------------ ------ ------------ -------------
       Global     /                  vxfs       Jun 23 03:30 rw, delaylog

            Global   /home             jfs2         Jun 23 03:30 rw,log=I
       Global    /opt              namefs Jun 23 03:30 ro
       Global    /proc         namefs Jun 23 03:30 rw
       Global    /tmp              jfs2         Jun 23 03:30 rw,log=INLIN
       Global    /usr              namefs Jun 23 03:30 ro
       Global    /var              jfs2         Jun 23 03:30 rw,log=INLIN
```

Using VxFS as a shared file system

VxFS is also supported as “namefs” in the WPAR, so a VxFS file system can also be shared between the global environment and WPARs.

To use VxFS as a shared file system

1 Mount vxfs on some directory in the global environment.

```
# mount -V vxfs /dev/vx/dsk/testdg/vol1 /mnt
```

2 Mount that directory in /wpar/ wpar1/vxfs_dir.

```
# mount /mnt /wpars/wpar1/vxfs_dir/
```

3 Start the WPAR.

```
# startwpar -Dv wpar1 2>/wpar1_tl2
```

4 Login to the WPAR.

```
# clogin wpar1
```

5 After login to wpar1, /vxfs_dir will appear as namefs.

```
# mount
node        mounted mounted over vfs        date              options
-----   -------------------- -----  ------------ -------------
       Global   /               jfs2       Jun 23 03:30 rw,log=INLINE
       Global   /home         jfs2      Jun 23 03:30 rw,log=INLINE
       Global   /opt              namefs Jun 23 03:30 ro
       Global   /proc         namefs Jun 23 03:30 rw
       Global   /tmp              jfs2      Jun 23 03:30 rw,log=INLINE
             Global /usr              namefs Jun 23 03:30 ro
             Global /var              jfs2      Jun 23 03:30 rw,log=INLIN
       Global   /vxfs_dir     namefs Jun 23 03:29 rw
```

How Cluster Server (VCS) works with Workload Patitions (WPARs)

VCS provides application management and high availability to applications that run in WPARs. VCS supports only system WPARs; application WPARs are not supported.

You can use VCS to perform the following:

- Start, stop, monitor, and failover a WPAR.
- Start, stop, monitor, and failover an application that runs in a WPAR.

Tasks for installing and configuring WPARs in VCS environments

- Install and configure the WPAR.
- Create the VCS service group with the standard application resource types (application, storage, networking) that need to be run inside the WPAR, and the WPAR resource.
  VCS represents the WPAR and its state using the WPAR resource.

Running VCS, its resources, and your applications

VCS and the necessary agents run in the global environment. For applications that run in a WPAR, the agents can run some of their functions (entry points) inside the WPAR. If any resource faults, VCS fails over the service group with the WPAR to another node.

About the ContainerInfo attribute

The ContainerInfo attribute has the Name key, Type key, and Enabled key. The Name key defines the name of the WPAR. The Type key lets you select the type of container that you plan to use (WPAR). The Enabled key enables the WPAR-aware resources within the service group. To configure the ContainerInfo attribute, use the `hawparsetup.pl` command.

You can specify a per-system value for the ContainerInfo attribute. For more information, refer to the *Cluster Server Administrator's Guide*.

About the ContainerOpts attribute

The ContainerOpts attribute has the RunInContainer key and PassCInfo key. If the resource type has the RunInContainer and PassCInfo keys defined in ContainerOpts, the resource type is WPAR-aware. WPAR-aware indicates that VCS can monitor and control a resource of that type inside a WPAR.

The ContainerOpts attribute determines the following:

- The RunInContainer key determines whether the entry points of a WPAR-aware resource type can run in the WPAR.
- The PassCInfo key determines whether the container information is passed to the entry points of the resource type. The container information is defined in the service group's ContainerInfo attribute. An example use of the PassCInfo key is to pass the agent the name of the WPAR.
  For more information, refer to the *Cluster Server Administrator's Guide*.

Note: Arctera recommends that you do not modify the value of the ContainerOpts attribute with the exception of the Mount agent.

The following are the ContainerOpts attribute default values for resource types. WPAR-aware resources have predefined default values for the ContainerOpts attribute.

Table 6-1 ContainerOpts attribute default values for resource types

Resource Type	RunInContainer	PassCInfo
Application	1	0
DB2	1	0
IP	0	1
IPMultiNICB	0	1

Table 6-1 ContainerOpts attribute default values for resource types *(continued)*

Resource Type	RunInContainer	PassCInfo
Netlsnr	1	0
Mount	0	0
Oracle	1	0
Process	1	0
WPAR	0	1

You may need to modify the ContainerOpts values for the Mount resource in certain situations. Refer to the *Cluster Server Bundled Agents Reference Guide* for more information.

About the WPAR agent

The WPAR agent monitors WPARs, brings WPARs online, and takes them offline.

For more information about the agent, see the *Cluster Server Bundled Agents Reference Guide*.

This agent is IMF-aware and uses asynchronous monitoring framework (AMF) kernel driver for IMF notification. For more information about the Intelligent Monitoring Framework (IMF) and intelligent resource monitoring, refer to the *Cluster Server Administrator's Guide*.

The agent requires a user account with group administrative privileges to enable communication between the global environment and the WPAR. To create a user account, use the `hawparsetup.pl` command to configure the service group.

See "Configuring the service group for the application" on page 107.

In secure clusters, the agent renews the authentication certificate before the certificate expires.

Configuring VCS in WPARs

Configuring VCS in WPARs involves the following tasks:

- Review the prerequisites.
  See "Prerequisites for configuring VCS in WPARs" on page 103.

- Decide on the location of the WPAR root, which is either on local storage or NFS. The WPAR root is the topmost directory in a section of the file system hierarchy in which the WPAR is configured.
  See "Deciding on the WPAR root location" on page 104.
- Install the application in the WPAR.
  See "Installing the application" on page 107.
- Create the application service group and configure its resources.
  See "Configuring the service group for the application" on page 107.

Prerequisites for configuring VCS in WPARs

- In a WPAR configuration, all nodes that host applications must run the same version of the operating system.
- The WPAR root must be installed on JFS, JFS2, NFS, or VxFS.
- Mounts must meet one of the following two conditions:
  - Use a namefs file system. All mounts that the application uses must be part of the WPAR configuration and must be configured in the service group. For example, you can create a WPAR, w_ora, and define the file system containing the application's data to have the mount point as /oradata. When you create the WPAR, you can define a path in the global environment, for example /export/home/oradata, which maps to the mount directory in the WPAR. The MountPoint attribute of the Mount resource for the application is set to /export/home/oradata.
  - Use a direct mount file system. All file system mount points that the application uses that run in a WPAR must be set relative to the WPAR's root. For example, if the Oracle application uses /oradata, and you create the WPAR with the WPAR path as /w_ora, then the mount must be /w_ora/oradata. The MountPoint attribute of the Mount resource must be set to this path.

  For more information about how to configure Mount resource inside WPAR, see the *Cluster Server Bundled Agents Reference Guide*.

About using custom agents in WPARs

- If you use custom agents to monitor applications running in WPARs, make sure the agents use script-based entry points. VCS does not support running C++ entry points inside a WPAR.
- If the custom agent monitors an application that runs in a WPAR, add the resource type to the APP_TYPES environment variable. If the custom agent

monitors an application running in the global environment, add the resource type to the SYS_TYPES environment variable.

Note: This step is required only for `hawparsetup`.

- If you want the custom agent to monitor an application in the WPAR, for the custom agent type, set the following values for the ContainerOpts attribute: RunInContainer = 1 and the PassCInfo = 0.
- If you do not want the custom agent to monitor an application in the WPAR, for the custom agent type, set the following values for the ContainerOpts attribute: RunInContainer = 0 and the PassCInfo= 0.

Deciding on the WPAR root location

Each WPAR has its own section of the file system hierarchy in the WPAR root directory. Processes that run in the WPAR can access files only within the WPAR root.

You can set the WPAR root in the following two ways:

- WPAR root on local storage.
  In this configuration, you must create a WPAR on each node in the cluster.
- WPAR root on NFS.
  In this configuration, create a WPAR on the NFS storage. You need to duplicate the WPAR configuration across all the nodes in the cluster.
  When you set the WPAR root on NFS, install the WPAR from one node only. The WPAR root can fail over to the other nodes in the cluster. The system software, including the patches, must be identical on each node during the existence of the WPAR.

Creating a WPAR root on local disk

Use the following procedure to create a WPAR root on the local disk on each node in the cluster.

To create a WPAR root on local disks on each node in the cluster

1 Create the actual WPAR root directory.

2 Use the `mkwpar` command to create the WPAR.

```
mkwpar -n wpar -h host -N ip_info -d wroot -o /tmp/wpar.log
```

Use the following information to replace the appropriate variables:

wpar	The name of the WPAR.
host	The hostname for the WPAR being created.
ip_info	The information to set the virtual IP address of the system to be the IP address of the WPAR. This value also defines the device name for the NIC associated with the IP address. If you do not specify the value of the interface or netmask, the global partition's values are used. Use the following format to replace ip_info: `interface=interface netmask=netmask address=IPaddress` Example: `interface='en0' address='172.16.0.0' netmask='255.255.255.0'`
wroot	The location of the WPAR root directory. For example: /wpar1.

To install application for DB2 under WPAR, you may want to create a detached WPAR where `/opt` and `/usr` are writable under WPAR. An example on how to create detach WPAR:

```
# mkwpar -l -n db2wpar -h db2wpar -d /wpars/shirish -r -N \
address=10.209.87.132 netmask=255.255.252.0
```

3 Repeat the command in step 2 to create the WPAR on each system in the service group's SystemList.

4 Start the WPAR.

5 On one of the systems in the SystemList, mount the shared file system containing the application data.

Creating WPAR root on shared storage using NFS

Use the following procedure to create a WPAR root on shared storage using NFS.

To create WPAR root on shared storage using NFS

1 Create a file system on NFS storage for the WPAR root. The file system that is to contain the WPAR root may be in the same file system as the file system containing the shared data.

2 Type the following `mkwpar` command to create the WPAR:

```
mkwpar -n wpar -h host -N ip_info -r -M r_fs -M v_fs -M h_fs -M
t_fs -d wroot
```

Use the following information to replace the appropriate variables:

Attribute	Description
wpar	The name of the WPAR.
host	The hostname of the WPAR being created.
ip_info	The information to set the virtual IP address of the system to be the IP address of the WPAR. This value also defines the device name for the NIC associated with the IP address. Use the following format to replace ip_info: `interface=interface netmask=netmask address=IPaddress` For example: `interface='en0' address='172.16.0.0' netmask='255.255.255.0'` If you do not specify the value of the interface or netmask, the global partition's values are used.
r_fs	The information to specify the NFS volume to use for the root private file system for the WPAR. For example: `directory=/ vfs=nfs host=host123 dev=/root01`
v_fs	The information to specify the NFS volume to use for the /var private file system for the WPAR. For example: `directory=/var vfs=nfs host=host123 dev=/var01`
h_fs	The information to specify the NFS volume to use for the /home private file system for the WPAR. For example: `directory=/home vfs=nfs host=host123 dev=/home01`
t_fs	The information to specify the NFS volume to use for the /tmp private file system for the WPAR. For example: `directory=/tmp vfs=nfs host=host123 dev=/tmp01`
wroot	The location of the WPAR root directory, for example, /wpar1.

3 Use the `lswpar` command to display information about the WPAR's properties and their values.

4 On the system where you created the WPAR, run the command:

```
mkwpar -w -o config_file_name -e wparname_just_created
```

5 On all the other systems copy the configuration file, run the command:

```
mkwpar -p -f config_file_name -n wparname_just_created
```

6 List the WPAR.

7 Start the WPAR.

8 On one system, mount the shared file system containing the application data.

9 Make sure the WPAR created from the first system is in the D state on all other systems in the service group's System List.

Installing the application

Install the application in the WPAR. Perform the following:

- If you have created WPARs on each node in the cluster, install the application identically on all nodes. If you are installing an application that supports a High Availability agent, see the installation and configuration guide for the agent.
- Install the agent. Agent filesets are installed in the global environment and the currently existing WPARs. The operating system installs the agents in future WPARs when they are created.
- In the WPAR, configure all mount points used by the application.
  - If you use namefs mounts, verify the global directories are properly mounted inside the WPAR.
  - If you use a direct mount, verify the mount points used by the application have been mounted relative to the WPAR's root. For example, if a WPAR w_ora needs to use /oracle, mount the drive at /wpars/w_ora/oracle.

Configuring the service group for the application

The following diagrams illustrates different examples of resource dependencies. In one case the WPAR root is set up on local storage. In the other, WPAR root is set up on shared storage.

Figure 6-1 WPAR root on local disks (with direct mount file system)

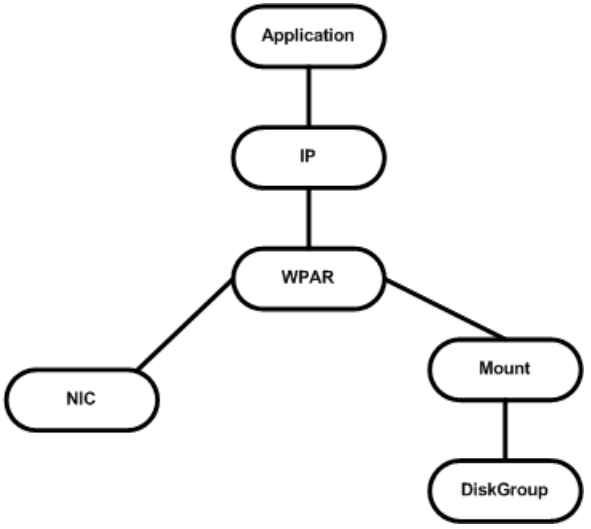


Figure 6-2 WPAR root on local disks (file system mounted from inside WPAR)

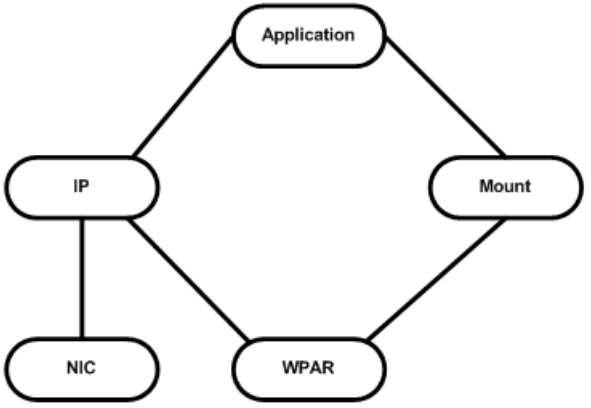
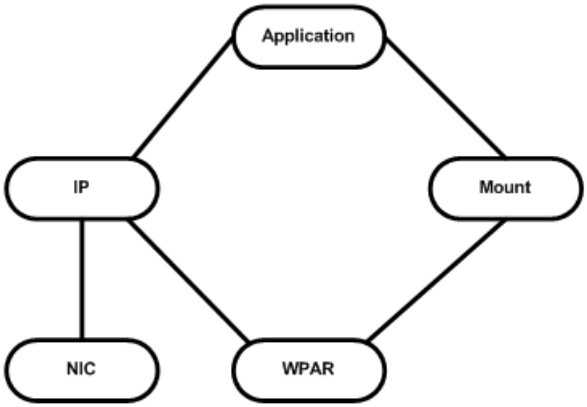


Figure 6-3 WPAR root on shared storage (with namefs file system)



Modifying the service group configuration

Perform the following procedure to modify a service group's configuration.

To add a service group or modify the service group configuration

1 Run the `hawparsetup.pl` command to set up the WPAR configuration.

```
# /opt/VRTSvcs/bin/hawparsetup.pl servicegroup_name WPARres_name \
WPAR_name password systems
```

servicegroup_name	Name of the application service group.
WPARres_name	Name of the resource configured to monitor the WPAR.
WPAR_name	Name of the WPAR.
password	Password to be assigned to VCS or Security (Veritas Product Authentication Service—VxAT) user created by the command.
systems	List of systems on which the service group will be configured. Use this option only when creating the service group.

The command adds a resource of type WPAR to the application service group. It also creates a user account with group administrative privileges to enable WPAR to global communication.

If the application service group does not exist, the command creates a service group.

2 Modify the resource dependencies to reflect your WPAR configuration. See the resource dependency diagrams for more information.

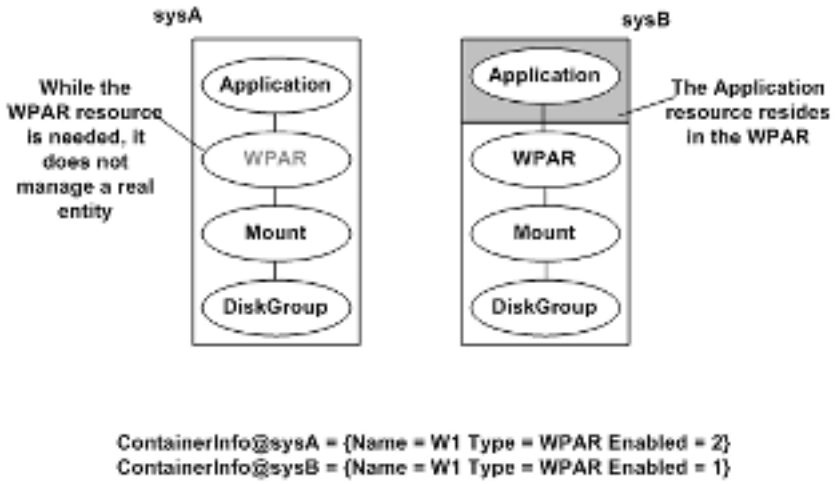
3 Save the service group configuration and bring the service group online.

About configuring failovers

An application can be failed over from an LPAR to a WPAR running on a different LPAR. You can configure VCS to fail over from a physical system to a virtual system and vice versa. A physical to virtual failover gives an N + N architecture in an N + 1 environment. For example, several physical servers with applications can fail over to containers on another physical server. On AIX, a container is a WPAR.

In this configuration, you have two LPARs. One node runs AIX 7.3 (sysA) and another node that runs AIX 7.2 (sysB). The node that runs AIX 7.2 has WPARs configured.

Figure 6-4 An application service group that can fail over onto a WPAR



In the `main.cf` configuration file, define the container name, type of container, and whether it is enabled or not. The following is an example of the ContainerInfo lines in the `main.cf` file:

```
ContainerInfo@sysA = {Name = W1, Type = WPAR, Enabled = 2}

ContainerInfo@sysB = {Name = W1, Type = WPAR, Enabled = 1}
```

On sysA, you set the value of Enabled to 2 to ignore WPARs so that the application runs on the physical system. When an application running on sysA fails over to sysB, the application runs inside the WPAR after the failover because Enabled is set to 1 on sysB. The application can likewise fail over to sysA from sysB.

IMF must be disabled on the node where Enabled is set to 2 (sysA in this example). To disable IMF, set the mode to 0.

On a Workload Partition (WPAR) where the WPAR is ignored to run the application on the physical system, you can disable the IMF for the WPARs.

To disable IMF monitoring

- Set the Mode key of IMF attribute to 0:

```
# hares -override <wpar_res> IMF
# hares -local <wpar_res> IMF
# hares -modify <wpar_res> IMF Mode 0 MonitorFreq 5 RegisterRetryLimit 3
-sys sysA
```

Verifying the WPAR configuration

Run the `hawparverify.pl` command to verify the WPAR configuration. The command verifies the following requirements:

- The systems hosting the service group have the required operating system to run WPARs.
- The service group does not have more than one resource of type WPAR.
- The dependencies of the WPAR resource are correct.

To verify the WPAR configuration

1. If you use custom agents make sure the resource type is added to the APP_TYPES or SYS_TYPES environment variable.

   See "About using custom agents in WPARs" on page 103.

2. Run the `hawparverify.pl` command to verify the WPAR configuration.

   # `/opt/VRTSvcs/bin/hawparverify` *`servicegroup_name`*

Maintenance tasks

Perform the following maintenance tasks when you use WPARs:

- Whenever you make a change that affects the WPAR configuration, you must run the `hawparsetup` command to reconfigure the WPARs in VCS.

  See "Configuring the service group for the application" on page 107.
- Make sure that the WPAR configuration files are consistent on all the nodes at all times.
- When you add a patch or upgrade the operating system on one node, make sure to upgrade the software on all nodes.
- Make sure that the application configuration is identical on all nodes. If you update the application configuration on one node, apply the same updates to all nodes.

Troubleshooting information

Symptom	Recommended Action
VCS HA commands do not work.	Verify that the VCS filesets are installed. Run the `hawparsetup` command to set up the WPAR configuration. Run the `hawparverify` command to verify the configuration. Run the `halogin` command from the WPAR. For more information, refer to the *Cluster Server Administrator's Guide*. Verify your VCS credentials. Make sure the password is not changed. Verify the VxSS certificate is not expired.
Resource does not come online in the WPAR.	Verify VCS and the agent filesets are installed correctly. Verify the application is installed in the WPAR. Verify the configuration definition of the agent. Make sure to define the Name and Type keys in the ContainerInfo attribute.

Configuring AIX WPARs for disaster recovery using VCS

AIX workload partitions (WPARs) can be configured for disaster recovery by replicating the base directory using replication methods like Hitachi TrueCopy, EMC SRDF, Volume Replicator (VVR), and so on. The network configuration for the WPAR in the primary site may not be effective in the secondary site if the two sites are in different IP subnets. Hence, you need to make these additional configuration changes to the WPAR resource.

To configure the WPAR for disaster recovery, you need to configure VCS on both the sites in the logical partitions (LPARs) with the GCO option.

Refer to the *Cluster Server Administrator's Guide* for more information about global clusters.

To set up the WPAR for disaster recovery

1 On the primary site, create the WPAR and configure its network parameters.

2 On the primary site, start the WPAR and configure the DNS settings.

3 On the primary site, shut down the WPAR.

4 Use replication-specific commands to fail over the replication to the secondary site from the primary site.

5 Repeat step 1 on the secondary site.

6 Perform step 7, step 8, step 9, and step 10 on both the primary cluster and secondary clusters.

7 Create a VCS service group with a VCS WPAR resource for the WPAR.

Refer to the *Cluster Server Bundled Agents Reference Guide* for more information about the WPAR resource.

Configure the DROpts association attribute on the WPAR resource with the following keys and site-specific values for each: DNSServers, DNSSearchPath, and DNSDomain.

8 Add the appropriate Mount resources and DiskGroup resources for the file system and disk group on which the WPAR's base directory resides.

Add a resource dependency from the WPAR resource to the Mount resource and another dependency from the Mount resource to the Diskgroup resource.

9 Add the appropriate VCS replication resource in the service group.

Examples of hardware replication agents are SRDF for EMC SRDF, HTC for Hitachi TrueCopy, MirrorView for EMC MirrorView, etc.

Refer the appropriate VCS replication agent guide for configuring the replication resource.

For VVR-based replication, add the appropriate RVGPrimary resource to the service group.

Refer to the following manuals for more information:

- For information about configuring VVR-related resources, see the *InfoScale Replication Administrator's Guide*.
- For information about the VVR-related agents, see the *Cluster Server Bundled Agents Reference Guide*.

10 Add a dependency from the DiskGroup resource to the replication resource.

Figure 6-5 Sample resource dependency diagram for hardware replication based WPARs

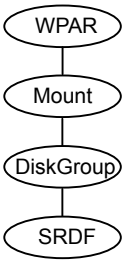
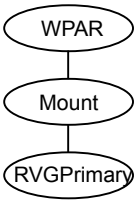


Figure 6-6 Sample resource dependency diagram for VVR replication-based WPARs



When the replication resource is online in a site, the replication agent makes sure of the following:

- The underlying replicated devices are in primary mode and hence the underlying storage and eventually the WPAR's base directory is always in read-write mode.
- The remote devices are in secondary mode.

When the WPAR resource goes online the resource modifies the appropriate files inside the WPAR to apply the disaster recovery-related parameters to the WPAR.

Chapter 7

High availability and live migration

This chapter includes the following topics:

About Live Partition Mobility (LPM)

You can use Live Partition Mobility feature from IBM for greater control over the usage of resources in the data center.

Live Partition Mobility enables:

- Migration of an entire logical partition from one physical system to another.
- The transfer of a configuration from source to destination without disrupting the hosted applications or the setup of the operating system and applications.
- A level of reconfiguration that was previously impossible due to complexity or service level agreements that did not allow an application to be stopped for an architectural change.

The migration process can be performed in the following ways:

- Inactive migration
  The logical partition is powered off and moved to the destination system.
- Active migration
  The migration of the partition is performed while service is provided, without disrupting user activities. During an active migration, the applications continue to handle their normal workload. Disk data transactions, running network connections, user contexts, and the complete environment are migrated without any loss and migration can be activated any time on any production partition.

About the partition migration process and simplified management

The partition migration, either inactive or active, is divided into the following stages:

- Preparing the infrastructure to support Live Partition Mobility.
- Checking the configuration and readiness of the source and destination systems.
- Transferring the partition state from the source to destination. The same command is used to launch inactive and active migrations. The HMC determines the appropriate type of migration to use based on the state of the mobile partition.
- Completing the migration by freeing unused resources on the source system and the HMC.

For performance considerations, consult IBM PowerVM Live Partition Mobility documentation.

About Storage Foundation and High Availability (SFHA) Solutions support for Live Partition Mobility

All SFHA Solutions products support Live Partition Mobility (LPM). A few requirements apply, as listed below.

Some limitations for LPM apply when VCS is configured to manage high availability of LPARs.

See "Limitations and unsupported LPAR features" on page 127.

The requirements for supporting the migration of a logical partition with high availability are:

- Generic requirements for logical partitions
  See "About setting up logical partitions (LPARs) with InfoScale products" on page 32.
- NPIV disks for fencing

See the IBM documentation for the detailed information on the LPM requirements and LPM process.

Providing high availability with live migration in a Cluster Server environment

You can use Live Partition Mobility to perform a stateful migration of a logical partition (LPAR) in a Cluster Server (VCS) environment. VCS supports LPAR live migration in two ways:

- LPAR migration outside of VCS control
- LPAR migration through VCS commands

VCS initiated LPAR migration

- To migrate a managed LPAR:

```
# hagrp -migrate <lpar_service_group> \
-to <target_vcs_node>
```

- Requirements for high availability support for live migration through VCS commands:
  - The `ProfileFile` attribute must contain correct information. If it does not, the LPAR creation or deletion fails. VCS cannot guarantee the correctness of the `ProfileFile` attribute.
  - The `ProfileFile` for an LPAR resource must contain valid VIOS mappings. If it does not, and the LPAR resource fails, then VCS is not able to delete VIOS mappings. This leaves the LPAR configuration in an intermediate state.
  - The `ProfileFile` for an LPAR must be recreated for the specific physical server if it is live migrated to a physical server. The live migration might assign mapping information which is not the same as earlier `ProfileFile`.

If VCS encounters an error while creating or deleting an LPAR configuration or VIOS mappings, then online or offline of LPAR resource stops immediately and does not recover from the intermediate state. Administrative intervention is

required when an LPAR configuration or VIOS mappings creation or deletion fails.

Some limitations for LPM apply when VCS is configured to manage high availability of LPARs.

See "Limitations and unsupported LPAR features" on page 127.

For more information, refer to the *Cluster Server Administrator's Guide*.

LPAR migration outside of VCS control

- VCS can detect LPAR migration initiated outside of VCS. During this period, you may see notifications if the migrating node is unable to heartbeat with its peers within LLT's default peer inactive timeout. You can reset the default LLT `peerinact timeout` value to enable the migrating node to heartbeat with its peers within LLT's default peer inactive timeout. For the example procedure below, the sample value is set to 90 seconds.

To avoid false failovers for LPAR migration outside of VCS control

1. Determine how long the migrating node is unresponsive in your environment.
2. If that time is less than the default LLT peer inactive timeout of 16 seconds, VCS operates normally.

   If not, increase the peer inactive timeout to an appropriate value on all the nodes in the cluster before beginning the migration.

   For example, to set the LLT `peerinact` timeout to 90 seconds, use the following command:

   ```
   # lltconfig -T peerinact:9000
   ```

   The value of the `peerinact` command is in .01 seconds.

3 Verify that `peerinact` has been set to 90 seconds:

```
# lltconfig -T query

Current LLT timer values (.01 sec units):
 heartbeat    = 50
 heartbeatlo = 100
 peertrouble = 200
 peerinact    = 9000
 oos          = 10
 retrans      = 10
 service      = 100
 arp          = 30000
 arpreq          = 3000
 Current LLT flow control values (in packets):
 lowwater  = 40
```

4 Repeat steps 1 to 3 on other cluster nodes.

5 Reset the value back to the default after the migration is complete.

To make LLT peerinact value persistent across reboots:

- Append the following line at the end of /etc/llttab file:

  ```
  set-timer peerinact:9000
  ```

  After appending the above line, /etc/llttab file should appear similar to the following:

  ```
  # cat /etc/llttab
  set-node host1
  set-cluster 1234
  link en2 en-00:15:17:48:b5:80 - ether - -
  link en3 en-00:15:17:48:b5:81 - ether - -
  set-timer peerinact:9000
  ```

Some limitations for Live Partition Mobility (LPM) apply when VCS is configured to manage high availability of LPARs.

See "Limitations and unsupported LPAR features" on page 127.

For more information on VCS commands, see the *Cluster Server Administrator's Guide*.

For attributes related to migration, see the *Cluster Server Bundled Agents Reference Guide*.

To migrate the managed LPAR without ProfileFile support

1 From the source managed system, back up the LPAR profile. After migration completes, the LPAR and its profiles are automatically deleted from the source.

   For VCS to manage the LPAR, the profile is required on the managed physical system of the management VCS that is part of the system list of the LPAR resource.

2 On the destination system, rename the LPAR profile that was created during initial configuration of LPAR as a resource on all systems. LPM validation fails if it finds the profile with same LPAR name on the destination managed physical system

3 Migrate the managed LPAR.

4 Perform one of the following:

   - If migration succeeds, the profile on source is removed. Restore and rename the LPAR profile from the backup that was taken in step 1. Remove the renamed LPAR profile on the destination.
   - If migration fails, remove the backup profile on the source. On the destination, rename the renamed LPAR profile to original LPAR profile.

Providing logical partition (LPAR) failover with live migration

This section describes how to create a profile file and use the ProfileFile attribute to automate LPAR profile creation on failback after migration.

For more information on manage the LPAR profile on source system after migration:

See " Live partition mobility of managed LPARs" on page 128.

Live migration of a managed LPAR deletes the LPAR profile and mappings of adapters from VIO servers from the source physical server. Without the LPAR configuration and VIOS adapter mappings of the physical server the LPAR cannot be brought online or failed over to the Cluster Server (VCS) node from which it has been migrated.

If an LPAR is to be made highly available, the LPAR profile file must be created using the steps provided below on all the VCS nodes on which the LPAR is to be made highly available. The VIO server names for an LPAR are different for each physical server and the adapter ids for the LPAR on each of the physical servers

also might be different, therefore the profile file must be created for each of the VCS nodes separately.

When bringing an LPAR online on another node, VCS performs the following actions:

- Checks if the LPAR configuration exists on that node.
- If the LPAR configuration does not exist and if the `ProfileFile` attribute is specified, VCS tries to create an LPAR configuration and VIOS mappings as specified in the file specified by ProfileFile.
- If creation of the LPAR configuration and VIOS mappings is successful, VCS brings LPAR online.
- If the `ProfileFile` attribute is not configured and if the LPAR configuration does not exist on the physical server, the LPAR resource cannot be brought online.

The `ProfileFile` attribute is used to specify path of LPAR profile file.If the `ProfileFile` attribute for a VCS node is configured and the `RemoveProfileOnOffline` attribute is set to 1, VCS performs the following on offline or clean:

- Deletes the LPAR configuration from the physical server.
- Deletes the adapter mappings from the VIO servers.

For more information on attributes `RemoveProfileOnOffline` and `ProfileFile`, see the *Cluster Server Bundled Agent Reference Guide*.

To create the profile file for an LPAR

1 Run the following command on HMC:

```
$ lssyscfg -r lpar -m physical-server-name --filter \
lpar_names=managed-lpar-name
```

2 From the output of above command, select the following fields in key-value pairs:

```
name,lpar_id,lpar_env,work_group_id,shared_proc_pool_util_auth,\
allow_perf_collection,power_ctrl_lpar_ids,boot_mode,auto_start,\
redundant_err_path_reporting,time_ref,lpar_avail_priority,\
suspend_capable,remote_restart_capable,affinity_group_id --header
```

Delete the remaining attributes and their values.

3 The remaining attributes are obtained from any profile associated with the managed LPAR.The name of the profile which you want to create is managed-lpar-profile-name.

Run the following command on HMC to get the remaining attribute.

```
$ lssyscfg -r prof -m physical-server-name --filter \
lpar_names=managed-lpar-name,profile_names=managed-lpar-profile-name
```

From the output of above command, select the following fields in key-value pairs:

```
name,all_resources,min_mem,desired_mem,max_mem,mem_mode,\
mem_expansion,hpt_ratio,proc_mode,min_procs,desired_procs,max_procs,\
sharing_mode,io_slots,lpar_io_pool_ids,max_virtual_slots,\
virtual_serial_adapters,virtual_scsi_adapters,virtual_eth_adapters,\
vtpm_adapters,virtual_fc_adapters,hca_adapters,conn_monitoring,\
auto_start,power_ctrl_lpar_ids,work_group_id,bsr_arrays,\
lhea_logical_ports,lhea_capabilities,lpar_proc_compat_mode
```

4 Rename the `name` attribute in the above output to `profile_name`.

5 Concatenate outputs from 1 and 3 with a comma and write it in a single line to a text file. This is the configuration file required for VCS to create or delete LPAR configuration. The absolute path of this file should be given in `ProfileFile` attribute.

Note: If an error occurs while creating a partition from the LPAR profile file, make sure that all the missing attributes are populated in the profile data file. For more information on the error, see the `LPAR_A.log` file.

Following example procedure illustrates the profile file generation for lpar05 which is running on physical_server_01. The LPAR resource which monitors lpar05 LPAR is lpar05_resource. The VCS node that manages the lpar05_resource on physical server physical_server_01 is lpar101 and on physical_server_02 is lpar201.

To generate a file profile for lpar05 on on physical_server_01

1 To get the LPAR details from the HMC, enter:

```
$ lssyscfg -r lpar -m physical_server_01 --filter \
lpar_names=lpar05
```

The output of this command is the following:

```
name=lpar05,lpar_id=15,lpar_env=aixlinux,state=Running,\
resource_config=1,os_version=AIX 7.1 7100-00-00-0000,\
logical_serial_num=06C3A0PF,default_profile=lpar05,\
curr_profile=lpar05,work_group_id=none,\
shared_proc_pool_util_auth=0,allow_perf_collection=0,\
power_ctrl_lpar_ids=none,boot_mode=norm,lpar_keylock=norm,\
auto_start=0,redundant_err_path_reporting=0,\
rmc_state=inactive,rmc_ipaddr=10.207.111.93,time_ref=0,\
lpar_avail_priority=127,desired_lpar_proc_compat_mode=default,\
curr_lpar_proc_compat_mode=POWER7,suspend_capable=0,\
remote_restart_capable=0,affinity_group_id=none
```

2 Select the output fields as explained in the procedure above.

See "To create the profile file for an LPAR" on page 122.

The key value pairs are the following:

```
name=lpar05,lpar_id=15,lpar_env=aixlinux,work_group_id=none,\
shared_proc_pool_util_auth=0,allow_perf_collection=0,\
power_ctrl_lpar_ids=none,boot_mode=norm,auto_start=0,\
redundant_err_path_reporting=0,time_ref=0,lpar_avail_priority=127,\
suspend_capable=0,remote_restart_capable=0
```

3 To get the profile details from the HMC, enter:

```
$ lssyscfg -r lpar -m physical_server_01 --filter \
lpar_names=lpar05,profile_names=lpar05
```

The output of this command is the following:

```
name=lpar05,lpar_name=lpar05,lpar_id=15,lpar_env=aixlinux,\
all_resources=0,min_mem=512,desired_mem=2048,max_mem=4096,\
min_num_huge_pages=null,desired_num_huge_pages=null,\
max_num_huge_pages=null,mem_mode=ded,mem_expansion=0.0,\
hpt_ratio=1:64,proc_mode=ded,min_procs=1,desired_procs=1,\
max_procs=1,sharing_mode=share_idle_procs,\
affinity_group_id=none,io_slots=none,lpar_io_pool_ids=none,\
max_virtual_slots=1000,\
"virtual_serial_adapters=0/server/1/any//any/1,\
1/server/1/any//any/1",\
"virtual_scsi_adapters=304/client/2/vio_server1/4/1,\
404/client/3/vio_server2/6/1",\
"virtual_eth_adapters=10/0/1//0/0/ETHERNET0//all/none,\
11/0/97//0/0/ETHERNET0//all/none,\
12/0/98//0/0/ETHERNET0//all/none",vtpm_adapters=none,\
"virtual_fc_adapters=""504/client/2/vio_server1/8/c050760431670010,\
c050760431670011/1"","",""604/client/3/vio_server2/5/c050760431670012,\
c050760431670013/1""",hca_adapters=none,boot_mode=norm,\
conn_monitoring=1,auto_start=0,power_ctrl_lpar_ids=none,\
work_group_id=none,redundant_err_path_reporting=null,bsr_arrays=0,\
lhea_logical_ports=none,lhea_capabilities=none,\
lpar_proc_compat_mode=default,electronic_err_reporting=null
```

4 After selection of the fields and renaming name to profile_name, the output is as follows:

```
profile_name=lpar05,all_resources=0,min_mem=512,desired_mem=2048,\
max_mem=4096,mem_mode=ded,mem_expansion=0.0,hpt_ratio=1:64,\
proc_mode=ded,min_procs=1,desired_procs=1,max_procs=1,\
sharing_mode=share_idle_procs,affinity_group_id=none,io_slots=none,\
lpar_io_pool_ids=none,max_virtual_slots=1000,\
"virtual_serial_adapters=0/server/1/any//any/1,1/server/1/any//any/1",\
"virtual_scsi_adapters=304/client/2/vio_server1/4/1,\
404/client/3/vio_server2/6/1",\
"virtual_eth_adapters=10/0/1//0/0/ETHERNET0//all/none,\
11/0/97//0/0/ETHERNET0//all/none,\
12/0/98//0/0/ETHERNET0//all/none",vtpm_adapters=none,\
"virtual_fc_adapters=""504/client/2/vio_server1/8/c050760431670010,\
c050760431670011/1"","""604/client/3/vio_server2/5/c050760431670012,\
c050760431670013/1""",hca_adapters=none,\
boot_mode=norm,conn_monitoring=1,auto_start=0,\
power_ctrl_lpar_ids=none,work_group_id=none,bsr_arrays=0,\
lhea_logical_ports=none,lhea_capabilities=none,\
lpar_proc_compat_mode=default
```

5 Concatenate these two outputs with comma, which is as follows:

```
name=lpar05,lpar_id=15,lpar_env=aixlinux,work_group_id=none,\
shared_proc_pool_util_auth=0,allow_perf_collection=0,\
power_ctrl_lpar_ids=none,boot_mode=norm,auto_start=0,\
redundant_err_path_reporting=0,time_ref=0,lpar_avail_priority=127,\
suspend_capable=0,remote_restart_capable=0,profile_name=lpar05,\
all_resources=0,min_mem=512,desired_mem=2048,max_mem=4096,\
mem_mode=ded,mem_expansion=0.0,hpt_ratio=1:64,proc_mode=ded,\
min_procs=1,desired_procs=1,max_procs=1,sharing_mode=share_idle_procs,\
affinity_group_id=none,io_slots=none,lpar_io_pool_ids=none,\
max_virtual_slots=1000,"virtual_serial_adapters=0/server/1/any//any/1,\
1/server/1/any//any/1",\
"virtual_scsi_adapters=304/client/2/vio_server1/4/1,\
404/client/3/vio_server2/6/1",\
"virtual_eth_adapters=10/0/1//0/0/ETHERNET0//all/none,\
11/0/97//0/0/ETHERNET0//all/none,12/0/98//0/0/ETHERNET0//all/none",\
vtpm_adapters=none,\
"virtual_fc_adapters=""504/client/2/vio_server1/8/c050760431670010,\
c050760431670011/1"",""604/client/3/vio_server2/5/c050760431670012,\
c050760431670013/1""",hca_adapters=none,boot_mode=norm,\
conn_monitoring=1,auto_start=0,power_ctrl_lpar_ids=none,\
work_group_id=none,bsr_arrays=0,lhea_logical_ports=none,\
lhea_capabilities=none,lpar_proc_compat_mode=default
```

6 Write this output to a text file. Assuming that the absolute location of profile file thus generated on lpar101 is /configfile/lpar05_on_physical_server_01.cfg, execute the following commands to configure the profile file in VCS.

```
$ hares -local lpar05_res ProfileFile
$ hares -modify lpar05_res ProfileFile \
/configfile/lpar05_on_physical_server_01 -sys lpar101
```

7 Repeat steps 1-6 to create the profie file for lpar05 for physical_server02.

Limitations and unsupported LPAR features

The following limitations apply to VCS support for LPARs:

- Live partition mobility of management LPARs is not supported.
  If LPARs are managed by VCS running in the management LPAR, then the live partition mobility of the management LPAR is blocked by VCS.
  If you need to migrate the management LPAR, follow the recommended steps.

See " Live partition mobility of management LPARs" on page 128.

- If the LPAR agent crashes, the migration of the management LPAR will remain blocked even if it is not managing any LPARs. To unblock, you can perform the following:

  ```
  # /usr/sbin/drmgr -u vcs_blockmigrate.sh
  ```

Live partition mobility of management LPARs

Live partition mobility is not supported if VCS is running in the management LPAR on a physical system.

If LPARs are managed by VCS running in the management LPAR, then the live partition mobility of the management LPAR is blocked by VCS. If you need to migrate the management LPAR, use the following procedure.

To migrate the management LPAR

1. Migrate or fail over the managed LPARs to another physical server before migrating the management LPAR.
2. Stop the LPAR agent.
3. Migrate the management LPAR.
4. When management LPAR is back on the source physical server (which matches with the value of PhysicalServer in the VCS system), start the LPAR agent.

   See the *Cluster Server Bundled Agents Reference Guide* for more information on the LPAR agent.

Live partition mobility of managed LPARs

When LPAR is managed by VCS, set the virtual SCSI adapter with any partition and any slot on the VIO. Map the virtual SCSI adapter to the correct SCSI adapter on the managed LPAR. This step needs to be part of the initial configuration on all physical hosts. Otherwise, reboot the VIO so that the configuration takes effect before you perform the migration.

Chapter 8

Multi-tier business service support

This chapter includes the following topics:

About Virtual Business Services

The Virtual Business Services feature provides visualization, orchestration, and reduced frequency and duration of service disruptions for multi-tier business applications running on heterogeneous operating systems and virtualization technologies. A virtual business service represents the multi-tier application as a consolidated entity that helps you manage operations for a business service. It builds on the high availability and disaster recovery provided for the individual tiers by Arctera InfoScale™ products such as Cluster Server.

Application components that are managed by Cluster Server or Microsoft Failover Clustering can be actively managed through a virtual business service.

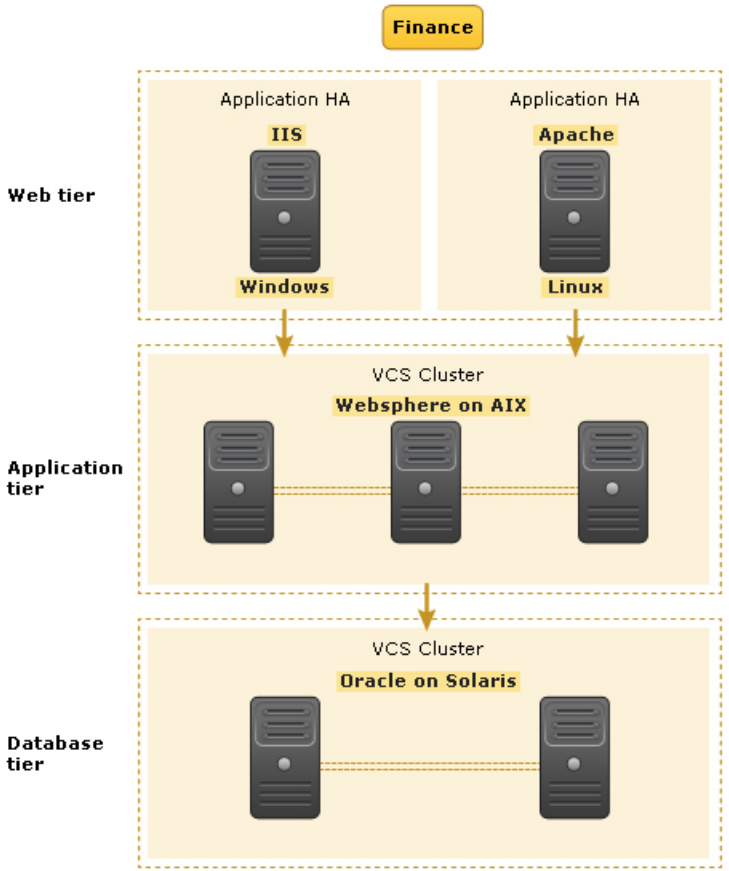
You can use the Arctera InfoScale Operations Manager Management Server console to create, configure, and manage virtual business services.

Sample virtual business service configuration

This section provides a sample virtual business service configuration comprising a multi-tier application. Figure 8-1 shows a Finance application that is dependent on components that run on three different operating systems and on three different clusters.

- Databases such as Oracle running on Solaris operating systems form the database tier.
- Middleware applications such as WebSphere running on AIX operating systems form the middle tier.
- Web applications such as Apache and IIS running on Windows and Linux virtual machines form the Web tier.
  Each tier can have its own high availability mechanism. For example, you can use Cluster Server for the databases and middleware applications for the Web servers.

Figure 8-1 Sample virtual business service configuration



Each time you start the Finance business application, typically you need to bring the components online in the following order – Oracle database, WebSphere, Apache and IIS. In addition, you must bring the virtual machines online before you start the Web tier. To stop the Finance application, you must take the components offline in the reverse order. From the business perspective, the Finance service is unavailable if any of the tiers becomes unavailable.

When you configure the Finance application as a virtual business service, you can specify that the Oracle database must start first, followed by WebSphere and the Web servers. The reverse order automatically applies when you stop the virtual business service. When you start or stop the virtual business service, the components of the service are started or stopped in the defined order.

For more information about Virtual Business Services, refer to the *Virtual Business Service–Availability User's Guide*.

Chapter 9

Server consolidation

This chapter includes the following topics:

About IBM LPARs with virtual SCSI devices

This discussion of vSCSI devices applies only to SAN-based LUNs presented through VIO. Internal devices, volumes, and files presented by VIO as vSCSI devices are not recommended for use with Storage Foundation.

Virtual SCSI uses a client/server model. A Virtual I/O server partition owns the physical I/O devices, and exports the devices as virtual SCSI (vSCSI) resources to the client partitions. The Virtual I/O client is a logical partition that has a virtual client adapter node defined in its device tree. The VIO client uses the vSCSI resources provided by the Virtual I/O Server partition to access the storage devices.

If redundant SAN connections exist to the VIO server, the VIO server provides multi-pathing to the array. Client partitions can also perform multi-pathing between VIO servers in an active/standby configuration. This configuration provides extended protection from VIO configuration and maintenance. Redundant VIO servers are recommended for production workloads.

A virtual SCSI (vSCSI) disk is a resource which can be a SCSI disk, or a volume or file in a VIO Server (VIOS) that is exported to a virtual IO client (VIOC). IBM

vSCSI LUNs implement a sub-set of the SCSI protocol. The two main limitations are:

- Persistent reservations (SCSI3 – PGR) are not implemented.
  The lack of SCSI reservations means that I/O Fencing is not supported. Storage Foundation Cluster File System High Availability (SFCFSHA) and Storage Foundation for Oracle RAC (SFRAC) do not support vSCSI disks, because SFCFSHA and SFRAC require I/O fencing.
- Device inquiry limitations.
  Storage Foundation (SF) cannot directly fetch the inquiry data, as is done from a physical SCSI disk. However, if the vSCSI disk in VIOC is backed by a dmpnode in VIOS, then all the inquiry data that can be fetched from a physical disk can be fetched.
  Cross-platform data sharing (CDS) functionality is supported.

Using Storage Foundation in the logical partition (LPAR) with virtual SCSI devices

Storage Foundation provides support for virtual SCSI (vSCSI) devices on the VIO client. You can create and manage Volume Manager (VxVM) volumes on vSCSI devices, as on any other devices. Storage Foundation provides Dynamic Multi-Pathing (DMP) for vSCSI devices, by default. Storage Foundation can also co-exist with MPIO for multi-pathing. If you choose to use MPIO to multipath the vSCSI devices, DMP works in pass-through mode.

Use the `vxddladm` utility and the `vxdmpadm` utility to administer DMP for vSCSI devices. The `vxddladm` utility controls enabling and disabling DMP on vSCSI devices, adding and removing supported arrays, and listing supported arrays. The `vxdmpadm` utility controls the I/O policy and the path policy for vSCSI devices.

Using Storage Foundation with virtual SCSI devices

Storage Foundation identifies the vSCSI LUNs through the array properties of the LUNs. Otherwise, the devices in the VIO client or logical partition (LPAR) appear as regular SCSI disks.

Storage Foundation supports vSCSI disks with version 5.1MP1 and later.

Portable Data Containers (disk type CDS) are supported. With extensions included in Storage Foundation 5.1, CDS type devices are now supported.

Storage Foundation can be deployed in the following ways:

- Use DMP in the VIO server to provide multi-pathing to the array. DMP presents a dmpnode as a vSCSI device to the LPAR.
- Use Storage Foundation in the LPAR to provide volume management on the vSCSI devices, and multi-pathing through the VIO servers with DMP.
- Use Storage Foundation in the LPAR to provide volume management on the vSCSI devices, and use MPIO to provide multi-pathing.

Setting up DMP for vSCSI devices in the logical partition (LPAR)

DMP is enabled on LPARs by default. After you install or upgrade Storage Foundation in the LPAR, any vSCSI devices are under DMP control and MPIO is disabled.

If you have already installed or upgraded Storage Foundation in the Virtual I/O client, use the following procedure to enable DMP support for vSCSI devices. This procedure is only required if you have previously disabled DMP support for vSCSI devices.

To enable vSCSI support within DMP and disable MPIO

1 Enable vSCSI support.

```
# vxddladm enablevscsi
```

2 You are prompted to reboot the system, if required.

DMP takes control of the devices, for any array that has DMP support to use the array for vSCSI devices. You can add or remove DMP support for vSCSI for arrays.

See "Adding and removing DMP support for vSCSI devices for an array" on page 136.

About disabling DMP for vSCSI devices in the logical partition (LPAR)

DMP can co-exist with MPIO multi-pathing in the Virtual I/O client or logical partition (LPAR). To use MPIO for multi-pathing, you can override the default behavior which enables Dynamic Multi-Pathing (DMP) in the LPAR.

There are two ways to do this:

- Before you install or upgrade Storage Foundation in the Virtual I/O client
  See "Preparing to install or upgrade Storage Foundation with DMP disabled for vSCSI devices in the logical partition (LPAR)" on page 135.
- After Storage Foundation is installed in the Virtual I/O client

See "Disabling DMP multi-pathing for vSCSI devices in the logical partition (LPAR) after installation or upgrade" on page 135.

Preparing to install or upgrade Storage Foundation with DMP disabled for vSCSI devices in the logical partition (LPAR)

Before you install or upgrade Storage Foundation, you can set an environment variable to disable DMP use for the vSCSI devices. Storage Foundation is installed with DMP in pass-through mode. MPIO is enabled for multi-pathing.

Note: When you upgrade an existing VxVM installation that has DMP enabled, then DMP remains enabled regardless of whether or not the environment variable __VXVM_DMP_VSCSI_ENABLE is set to no.

To disable DMP before installing or upgrading SF in the LPAR

1 Before you install or upgrade VxVM, set the environment variable __VXVM_DMP_VSCSI_ENABLE to no.

```
# export __VXVM_DMP_VSCSI_ENABLE=no
```

Note: The environment variable name __VXVM_DMP_VSCSI_ENABLE begins with two underscore (_) characters.

---

2 Install Storage Foundation, as described in the *InfoScale Installation Guide*.

Disabling DMP multi-pathing for vSCSI devices in the logical partition (LPAR) after installation or upgrade

After VxVM is installed, use the `vxddladm` command to switch vSCSI devices between MPIO control and DMP control.

To return control to MPIO, disable vSCSI support with DMP. After DMP support has been disabled, MPIO takes control of the devices. MPIO implements multi-pathing features such as failover and load balancing; DMP acts in pass-through mode.

To disable vSCSI support within DMP and enable MPIO

1 Disable vSCSI support.

```
# vxddladm disablevscsi
```

2 You are prompted to reboot the system, if required.

Adding and removing DMP support for vSCSI devices for an array

Dynamic Multi-Pathing (DMP) controls the devices for any array that has DMP support to use the array for vSCSI devices.

To add or remove DMP support for an array for use with vSCSI devices

1. To determine if DMP support is enabled for an array, list all of the arrays that DMP supports for use with vSCSI devices:

   ```
   # vxddladm listvscsi
   ```

2. If the support is not enabled, add support for using an array as a vSCSI device within DMP:

   ```
   # vxddladm addvscsi array_vid
   ```

3. If the support is enabled, you can remove the support so that the array is not used for vSCSI devices within DMP:

   ```
   # vxddladm rmvscsi array_vid
   ```

4. You are prompted to reboot the system, if required.

How DMP handles I/O for vSCSI devices

On the VIO client, DMP uses the Active/Standby array mode for the vSCSI devices. Each path to the vSCSI device is through a VIO server. One VIO server is Active and the other VIO servers are Standby. An Active/Standby array permits I/O through a single Active path, and keeps the other paths on standby. During failover, I/O is scheduled on one of the standby paths. After failback, I/Os are scheduled back onto the original Active path.

The following command shows the vSCSI enclosure:

```
# vxdmpadm listenclosure all
ENCLR_NAME ENCLR_TYPE ENCLR_SNO STATUS    ARRAY_TYPE LUN_COUNT FIRMWARE
============================================================================
ibm_vscsi0 IBM_VSCSI  VSCSI     CONNECTED VSCSI      9         -
```

The following command shows the I/O policy for the vSCSI enclosure:

```
# vxdmpadm getattr enclosure ibm_vscsi0 iopolicy
ENCLR_NAME     DEFAULT        CURRENT
============================================
ibm_vscsi0     Single-Active  Single-Active
```

For vSCSI devices, DMP balances the load between the VIO servers, instead of balancing the I/O on paths. By default, the `iopolicy` attribute of the vSCSI array is set to `lunbalance`. When `lunbalance` is set, the vSCSI LUNs are distributed so that the I/O load is shared across the VIO servers. For example, if you have 10 LUNs and 2 VIO servers, 5 of them are configured so that VIO Server 1 is Active and VIO Server 2 is Standby. The other 5 are configured so that the VIO Server 2 is Active and VIO Server 1 is Standby. To turn off load sharing across VIO servers, set the `iopolicy` attribute to `nolunbalance`.

DMP dynamically balances the I/O load across LUNs. When you add or remove disks or paths in the VIO client, the load is rebalanced. Temporary failures like enabling or disabling paths or controllers do not cause the I/O load across LUNs to be rebalanced.

Setting the vSCSI I/O policy

By default, DMP balances the I/O load across VIO servers. This behavior sets the I/O policy attribute to `lunbalance`.

To display the current I/O policy attribute for the vSCSI array

- Display the current I/O policy for a vSCSI array:

```
# vxdmpadm getattr vscsi iopolicy
VSCSI     DEFAULT      CURRENT
============================================
IOPolicy  lunbalance   lunbalance
```

To turn off the LUN balancing, set the I/O policy attribute for the vSCSI array to `nolunbalance`.

To set the I/O policy attribute for the vSCSI array

- Set the I/O policy for a vSCSI array:

```
# vxdmpadm setattr vscsi iopolicy={lunbalance|nolunbalance}
```

Note: The DMP I/O policy for each vSCSI device is always Single-Active. You cannot change the DMP I/O policy for the vSCSI enclosure. Only one VIO server can be Active for each vSCSI device.

Using VCS with virtual SCSI devices

Cluster Server (VCS) supports disk groups and volume groups created on virtual SCSI devices. The VCS DiskGroup agent supports disk groups. The VCS LVMVG agent supports volume groups.

Due to lack of SCSI3 persistent reservations, I/O Fencing is not supported with virtual SCSI devices.

If fencing is enabled and private diskgroups created on virutal SCSI devices are configured under VCS, then Reservation attribute should be defined and set to NONE for these diskgroups.

The command to set the value of Reservation attribute to None is:

```
#hares -modify <dg_res> Reservation "None"
```

About server consolidation

You can consolidate workloads from multiple physical servers to a physical server with virtual machines.

About IBM Virtual Ethernet

Virtual Ethernet enables communication between inter-partitions on the same server, without requiring each partition to have a physical network adapter. You can define in-memory connections between partitions that are handled at the system level (for example, interaction between POWER Hypervisor and the operating systems). These connections exhibit characteristics similar to physical high-bandwidth Ethernet connections and support the industry standard protocols (such as IPv4, IPv6, ICMP, or ARP). Virtual Ethernet also enables multiple partitions to share physical adapters for access to external networks using Shared Ethernet Adapter (SEA).

Shared Ethernet Adapter (SEA)

A Shared Ethernet Adapter is a layer-2 network bridge to securely transport network traffic between virtual Ethernet networks and physical network adapters. The SEA also enables several client partitions to share one physical adapter. The SEA is hosted in the Virtual I/O Server.

To bridge network traffic between the internal virtual network and external networks, configure the Virtual I/O Server with at least one physical Ethernet

adapter. Multiple virtual Ethernet adapters can share one SEA. Each virtual Ethernet adapter can support multiple VLANs.

The SEA has the following characteristics:

- Virtual Ethernet MAC addresses of virtual Ethernet adapters are visible to outside systems (using the `arp -a` command).
- Supports unicast, broadcast, and multicast. Protocols such as Address Resolution Protocol (ARP), Dynamic Host Configuration Protocol (DHCP), Boot Protocol (BOOTP), and Neighbor Discovery Protocol (NDP) can work across an SEA.

Chapter 10

Physical to virtual migration (P2V)

This chapter includes the following topics:

About migration from Physical to VIO environment

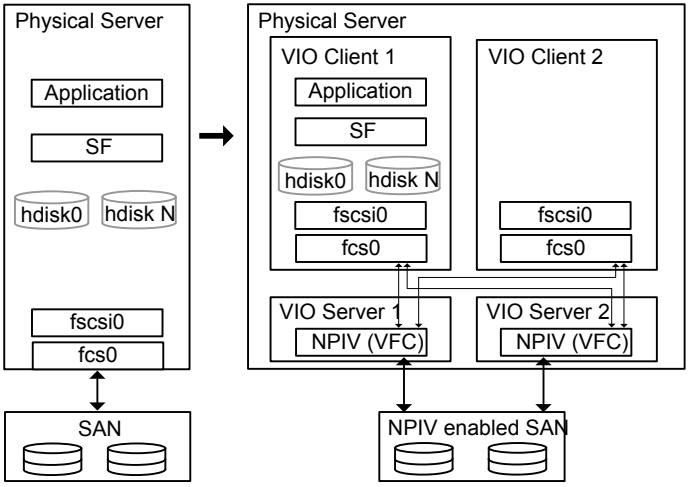
Arctera has qualified migration of storage that is used by Storage Foundation from the physical environment to IBM VIO environment.

Storage Foundation provides the PDC (Portable Data Container) feature, which enables migrating storage from other platforms (Solaris or Linux) to AIX VIO environment. You can also use PDC feature to migrate the storage consumed by a AIX physical server to a AIX VIO environment. NPIV helps you migrate the applications along with storage from a AIX physical environment to AIX VIO environment and vice-versa.

When storage is consumed by SF, Volume Manager (VxVM) initializes the storage LUNs as CDS (Cross-platform Data Sharing) type disks by default. A CDS disk group can be imported in a VIO client which has access to LUN's that are mapped through VFC Adapter on the client.

As part of the migration qualification, an application's storage is migrated from physical server to VIO environment (VIO client 1) which has NPIV capable FC adapter connected to it. This allows the application to access the storage in VIO client 1. With NPIV capable FC adapter at VIOS, the devices presented to the VIO client would appear as regular AIX hdisk devices. Figure 10-1 shows this migration.

Figure 10-1 SF migration from a physical environment to AIX VIO environment



Migration is an offline task.

Migrating from Physical to VIO environment

Migration of storage from a physical environment to the VIO environment is an offline task. The migration procedure involves stopping the application, unmounting the file systems and deporting the disk group on the physical server. Prior to being deported, you can take a space optimized snapshot, to facilitate fail-back.

Verify that the devices are visible on VIO client and the VFC adapter mapping between VIOS and VIO client is set up correctly. Refer to the IBM documentation for details. After all the required devices are accessible in VIO client 1, import the disk group in the client, mount the file system, and then start the application on the VIO client 1.

Refer to *IBM* documentation on how to configure the VFC adapter mappings between the VIO partition and the Client Partition.

Section 4

Reference

Appendix A

How to isolate system problems

This appendix includes the following topics:

About VxFS trace events

AIX trace facility lets you isolate system problems by monitoring selected system events or selected processes.

Events that can be monitored include:

- Entry and exit to selected subroutines
- Kernel routines
- Kernel extension routines
- Interrupt handlers

Trace can also be restricted to tracing a set of running processes or threads, or it can be used to initiate and trace a program.

When the trace facility is active, information is recorded in a system trace log file. The default trace log file from which the system generates a trace report is the `/var/adm/ras/trcfile` file. You can specify an alternate log file using the `-o Name`, this overrides the `/var/adm/ras/trcfile` default trace log file and writes trace data to a user-defined file.

See the `trcrpt` command that formats a report from the trace log.

You can specify your own trace log file path. Otherwise, the `/var/adm/ras/trcfile` is the default path.

The trace facility includes commands or subroutines for:

- Activating traces - The `trace` command or `trcstart` subroutine.
- Controlling traces - The `trcstop` command or the `trcstop` subroutine for stopping the tracing. While active, tracing can be suspended or resumed with the `trcoff` and `trcon` commands, or the `trcoff` and `trcon` subroutines.
- Generating trace reports - The `trcrpt` command.

The trace report can be generated from trace event data in an already defined format. Applications and kernel extensions can use several subroutines to record additional events.

You can specify your own trace format file path. Otherwise, the `/etc/trcfmt` is the default path.

VxFS uses the tracing facility to trace file read-write, inode cache operation, and low memory scenario. Table A-1 shows the details of trace hooks identifier (a three- or four-digit hexadecimal number that identifies an event being traced) used for these events:

Table A-1 Events and trace hook identifiers

Event	Hook Identifier
Read-write	0E1
Inode Cache	0E4
Low Memory	0E5

Tracing file read-write event

Trace events also support tracing of read-write operation. Trace can provide information such as file offset, file data size, file segment, and time taken for read-write operation.

In this example, the trace hook identifier for VxFS file read-write operation is 0E1. For tracing only file read-write operation, enter:

```
# trace -a -j 0E1 &
# trcon
# cat script.sh  (file on VxFS file system)
```

```
# trcoff
# trcstop
# trcrpt > trace.out
```

Sample output:

```
0E1 9.285786109 0.223881 VxFS rdwr
(vp,ip)=(F1000A02A334B500,F1000A02A334AF20)
0E1 9.285796646 0.010537 VxFS read  offset=00001000, seg=829B85,
bcount=1000, ip=F1000A02A334AF20
```

Tracing Inode cache event

For any inode, an iget() call increments its vcount by 1, and an iput() call decrements the vcount if it is not equal to one. The trace entries containing 'iget' or 'iput' would correspond to inode cache operations. The entries also provide details such as inode number and device ID, whenever inode cache operations are performed in VxFS. You can use tracing to collect the data and analyze it further.

In this example, the trace hook identifier for VxFS inode cache operations is 0E4. For tracing only inode cache operations, enter:

```
# trace -a -j 0E4 &
# trcon
# cat script.sh  (file on VxFS file system)
# trcoff
# trcstop
# trcrpt > trace.out
```

Sample output :

```
0E1 0.028564739 0.089402 VxFS VxFS iget: vp = F10001180CCDEDF0,
dev = 8000002D00004E20, fsindex = 03E7, iltype = 0000, vcount = 0001,
inode = 0002, getcaller = D60063

0E1 0.031875091 0.001512 VxFS VxFS iput: vp = F10001180CCDEDF0,
dev = 8000002D00004E20, fsindex = 03E7, iltype = 0000, vcount = 0001,
inode = 0002, getcaller = D60063
```

Tracing Low Memory event

Trace events have also been added to trace memory pressure situations. It helps to identify pressure on page cache or pinned memory heap.

While performing read/write operations on page cache, 'lowmem' thresholds is calculated. If the number of client free pages goes below this threshold, VxFS starts its own pager. The trace entries provide details about free client pages and whether the low memory condition is detected.

In this example, the trace hook identifier for VxFS memory pressure situations is 0E5. For tracing memory pressure:

```
# trace -a -j 0E5 &
# trcon
# cat script.sh (file on VxFS file system)
# trcoff
# trcstop
# trcrpt > trace.out
```

Sample output with entries for low page cache:

```
0E1 0.000619916 VxFS lowpgcache: numclientframes AA96, trigger_limit CC880,
lowmem detected 0000
```

VxFS checks the available pinned heap and calculates the limitalloc, limitsuff, and limitfree thresholds. If pinnable memory goes below the 'limitalloc' threshold, pinned memory allocation for new inodes and buffers is allowed in certain scenarios only. The 'limitsuff' threshold is limit to decide whether in-sufficient pinned memory situation has reached. If 'limitfree' threshold is hit then VxFS starts freeing allocations which are not in use currently. Basically there are pools or caches for various types of data structures. Whenever VxFS needs any of these data structures, memory from theses pools or caches are taken instead of asking the operating system every time. When VxFS is done with the data structure, the memory is returned back to pool or cache as the case may be. If memory pressure is noticed and this threshold is hit then VxFS looks at these pools, caches and see whether any memory can be given back to system. The trace entries provide details about these thresholds and pinnable memory left.

The trace entries provide details about these thresholds and pinnable memory left.

Sample output with entries for low pinned heap:

```
0E1 0.926113776 0.118139 VxFS lowpinnedheap: limitalloc = 17B8E,
limitsuff = 14C1C, limitfree = 11CAA, pinnable_left = 71E55
```

Appendix B

Provisioning data LUNs

This appendix includes the following topics:

Provisioning data LUNs in a mixed VxVM and LVM environment

The standard method for identifying which data LUNs are controlled by VxVM or LVM is to use the 'vxdisk list' command and review the output.

To identify data LUNs controlled by VxVM or LVM

- Use the `vxdisk list` command to identify disks that are under LVM or VxVM control:

```
Disks that are under LVM control
disk_0       auto:LVM        -            -            LVM
v_xiv0_108a  auto:LVM        -            -            LVM

Disks that are under VxVM control
ibm_ds8x000_06cd auto:aixdisk    -        -           online
ibm_ds8x000_06cf auto:cdsdisk    -        -           online

Disks that are not under any volume manager control.
ibm_ds8x000_06ce auto:none       -        -           online invalid
```

Once you identify the LUNs controlled by LVM you can select the ones you want to release for the available storage pool. You must use the native LVM commands `rmlv` and `reducevg` to remove LVM volumes and volume groups created on those LUNs as illustrated below.

To release LVM disks and bring them under VxVM control

1 Remove any LVM volumes that reside on the LVM disk.

Example:

```
# rmlv -B -f logical_volume_name
```

2 Remove the disk from the volume group.

Example:

```
# reducevg volume_group_name LVM_disk_name
```

3 Use `vxdiskunsetup` to clear out any stale header information.

4 Use `vxdisksetup` to bring them under VxVM control.

To release VxVM disks for LVM control

1 Make sure that disk is not under any disk group. Remove the VxVM header if present.

Example:

```
# /etc/vx/bin/vxdiskunsetup -Cf disk_name
```

2 Remove the disk from VxVM control.

Example:

```
# vxdisk rm disk_name
```

For more details on managing volumes, see the *Storage Foundation Administrator's Guide*.

Appendix C

Where to find more information

This appendix includes the following topics:

InfoScale documentation

The latest documentation is available on the Services and Operations Readiness Tools (SORT) website in the Adobe Portable Document Format (PDF).

See the release notes for information on documentation changes in this release.

Make sure that you are using the current version of documentation. The document version appears on page 2 of each guide. The publication date appears on the title page of each document. The documents are updated periodically for errors or corrections.

https://sort.veritas.com/documents

You need to specify the product and the platform and apply other filters for finding the appropriate document.

Additional documentation for AIX virtualization

For IBM documentation:

See http://www-03.ibm.com/systems/power/software/virtualization

Service and support

To access the self-service knowledge base, go to the following URL:

https://www.veritas.com/support/en_US.html

About Services and Operations Readiness Tools (SORT)

Services and Operations Readiness Tools (SORT) is a website that automates and simplifies some of the most time-consuming administrative tasks. SORT helps you manage your datacenter more efficiently and get the most out of your Arctera products.

SORT can help you do the following:

Prepare for your next installation or upgrade	■ List product installation and upgrade requirements, including operating system versions, memory, disk space, and architecture. ■ Analyze systems to determine if they are ready to install or upgrade Arctera products. ■ Download the latest patches, documentation, and high availability agents from a central repository. ■ Access up-to-date compatibility lists for hardware, software, databases, and operating systems.
Manage risks	■ Get automatic email notifications about changes to patches, array-specific modules (ASLs/APMs/DDIs/DDLs), and high availability agents from a central repository. ■ Identify and mitigate system and environmental risks. ■ Display descriptions and solutions for hundreds of Arctera error codes.
Improve efficiency	■ Find and download patches based on product version and platform. ■ List installed Arctera products and license keys. ■ Tune and optimize your environment.

Note: Certain features of SORT are not available for all products. Access to SORT is available at no extra cost.

To access SORT, go to:

https://sort.veritas.com